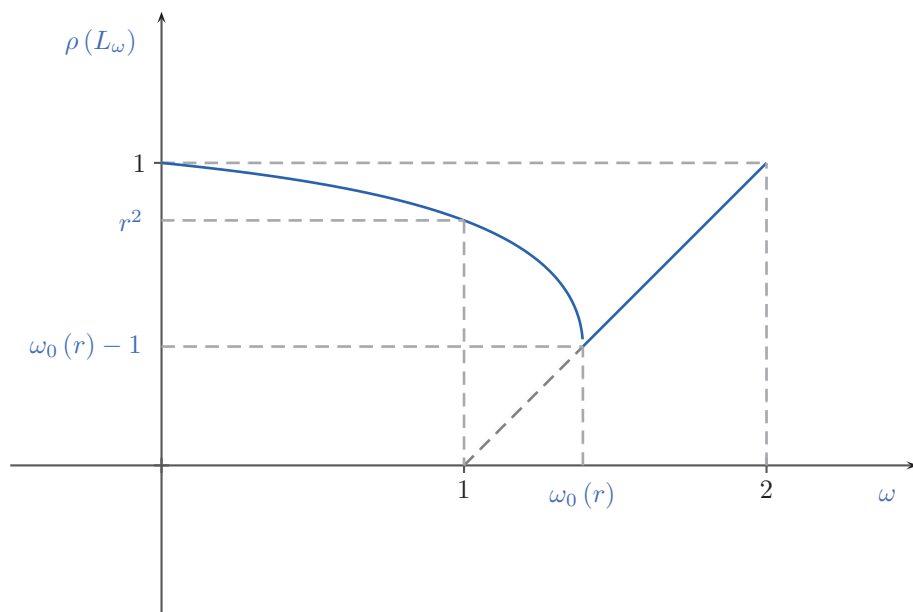




Jean-Étienne Rombaldi

ANALYSE MATRICIELLE

COURS ET

EXERCICES RÉSOLUS

2^e édition

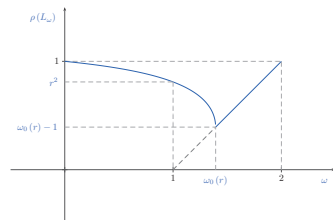
Jean-Étienne Rombaldi

Cette deuxième édition du livre « Analyse matricielle » est corrigée et augmentée d'un chapitre sur les matrices réelles positives et stochastiques.

Cet ouvrage est consacré à l'étude de l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ des matrices carrées à coefficients réels ou complexes du point de vue algébrique et topologique, préalable nécessaire à tout cours d'analyse numérique. La synthèse réalisée par l'auteur permet aux étudiants d'approfondir leurs connaissances sur les espaces vectoriels normés et l'algèbre linéaire, des notions de base en algèbre linéaire et en topologie étant suffisantes pour la lecture de ce livre.

Le public visé est celui des candidats à l'agrégation (interne et externe) et également celui des étudiants de licence et maîtrise de mathématiques. Chaque chapitre est suivi d'une série d'exercices corrigés. Les résultats classiques sont illustrés par des exemples qui peuvent trouver leur place dans les leçons d'oral des concours.

Jean-Étienne Rombaldi est professeur agrégé de mathématiques, son dernier poste étant à l'Institut Fourier de Grenoble (Université Grenoble-Alpes). Il a longtemps été préparateur, à l'université et pour le compte du CNED à l'agrégation interne et externe de mathématiques.



Jean-Étienne ROMBALDI

Analyse matricielle

Cours et exercices résolus

2^e édition

Dans la même collection

Éléments d'analyse réelle, 2^e édition

Jean-Étienne Rombaldi

2019, ISBN : 978-2-7598-2339-0

Thèmes pour l'agrégation de mathématiques, 2^e édition

Jean-Étienne Rombaldi

2019, ISBN : 978-2-7598-2340-6

Imprimé en France

ISBN (papier) : 978-2-7598-2341-3 - ISBN (ebook) : 978-2-7598-2419-9

Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.

© EDP Sciences, 2019

Table des matières

Avant-propos	v
1 Polynômes minimal et caractéristique. Sous espaces caractéristiques	1
1.1 Définitions et premières propriétés	2
1.2 Localisation des valeurs propres d'une matrice complexe	7
1.3 Matrice compagnon d'un polynôme	10
1.4 Le théorème de Cayley-Hamilton	13
1.5 Méthodes de calcul du polynôme caractéristique d'une matrice complexe	14
1.6 Sous espaces caractéristiques	17
1.7 Exercices	21
2 Réduction des endomorphismes et des matrices	31
2.1 Trigonalisation	31
2.2 Diagonalisation	33
2.3 Espaces vectoriels euclidiens	34
2.4 Réduction des matrices orthogonales	40
2.5 Réduction des matrices symétriques réelles	42
2.6 Tridiagonalisation des matrices symétriques réelles. Méthode de Householder	44
2.7 Espaces vectoriels hermitiens	46
2.8 Réduction des matrices normales	49
2.9 Forme réduite de Jordan	52
2.10 Exercices	56
3 L'espace vectoriel normé $\mathcal{M}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)	73
3.1 Norme matricielle induite par une norme vectorielle	73
3.2 Le groupe topologique $GL_n(\mathbb{K})$	77
3.3 Propriétés topologiques de l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$	83
3.4 Rayon spectral d'une matrice complexe	86
3.5 Conditionnement d'une matrice	94
3.6 Quotient de Rayleigh-Ritz et Hausdorffien	96
3.7 Conditionnement des problèmes de valeurs propres	99
3.8 Exercices	102

4	Matrices positives et irréductibles	123
4.1	Matrices positives	123
4.2	Matrices strictement positives et théorème de Perron-Frobenius . .	128
4.3	Matrices irréductibles	134
4.4	Matrices primitives	139
4.5	Matrices stochastiques et bistochastiques	141
4.6	Exercices	154
5	Systèmes linéaires	161
5.1	Position des problèmes et notations	161
5.2	Problèmes numériques liés à la résolution des systèmes linéaires . .	162
5.3	Cas des matrices triangulaires	164
5.4	Matrices de dilatation et de transvection. Opérations élémentaires	164
5.5	Méthode des pivots de Gauss	168
5.6	Résolution des systèmes linéaires à coefficients entiers	170
5.7	Décomposition LR ou méthode de Crout	171
5.8	Décomposition LD^tL des matrices symétriques réelles	174
5.9	Décomposition de Cholesky des matrices symétriques réelles définies positives	175
5.10	Méthode d'élimination de Gauss-Jordan	176
5.11	Méthodes itératives de résolution des systèmes linéaires	177
5.12	Méthode de Jacobi	178
5.13	Méthode de Gauss-Seidel	179
5.14	Méthode de relaxation	181
5.15	Méthodes de descente et de gradient	188
5.16	Exercices	196
6	Calcul approché des valeurs et vecteurs propres	209
6.1	Introduction	209
6.2	Méthode de la puissance itérée	209
6.3	Méthode de Jacobi pour les matrices symétriques	213
6.4	La méthode de Givens et Householder	218
6.5	Exercices	223
7	Systèmes différentiels linéaires et exponentielle d'une matrice	229
7.1	Systèmes différentiels linéaires à coefficients constants	229
7.2	L'exponentielle d'une matrice	233
7.3	Un algorithme de calcul de l'exponentielle d'une matrice	239
7.4	Equations différentielles linéaires d'ordre n à coefficients constants	240
7.5	Systèmes différentiels linéaires à coefficients non constants	242
7.6	Méthode de variation des constantes	245
7.7	Surjectivité et injectivité de l'exponentielle matricielle	247
7.8	Exercices	251

Avant-propos

Cet ouvrage, qui pourrait s'intituler « Matrices réelles et complexes, propriétés algébriques et topologiques, applications » est consacré à l'étude de l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n à coefficients réels ou complexes du point de vue algébrique et topologique. Cette étude est un préalable important à tout bon cours d'analyse numérique.

Des connaissances de base en algèbre linéaire et en topologie sont amplement suffisantes pour la lecture de cet ouvrage.

Le public visé est celui des étudiants du deuxième cycle universitaire et des candidats à l'Agrégation externe et interne de Mathématiques.

La synthèse proposée est un bon moyen de réviser ses connaissances sur les espaces vectoriels normés et l'algèbre linéaire. Les candidats à l'agrégation trouveront tout au long de cet ouvrage de nombreux exemples d'applications des résultats classiques souvent proposés dans les leçons d'oral. Par exemple, si dans une leçon sur le groupe orthogonal on pense à mentionner la compacité de $\mathcal{O}_n(\mathbb{R})$ il faut avoir réfléchi à quelques exemples d'applications de ce résultat. En suivant cette idée, je me suis efforcé de faire suivre chaque résultat classique et important d'un certain nombre d'applications.

Chaque chapitre est suivi d'une liste d'exercices corrigés. Une bonne utilisation de ces exercices consiste bien évidemment à les chercher au préalable, puis à confronter les résultats obtenus aux solutions proposées.

L'étude des propriétés topologiques de l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ et l'application aux méthodes itératives de résolution des systèmes linéaires et de recherche des valeurs et vecteurs propres utilisent quelques résultats de base sur les espaces vectoriels normés de dimension finie. On pourra se reporter à [18] pour l'étude des espaces vectoriels normés. En particulier, le théorème du point fixe de Banach est utilisé dans l'étude des systèmes différentiels linéaires.

Les chapitres 1 et 2 sont consacrés à l'étude des valeurs et vecteurs propres des matrices réelles ou complexes. Les résultats importants sont le théorème de décomposition des noyaux et les divers théorèmes de réduction à la forme triangulaire ou diagonale.

C'est au chapitre 3 qu'on aborde l'étude des propriétés topologiques de l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$. On y introduit les notions de norme matricielle induite par une norme vectorielle et on démontre quelques résultats classiques de densité et de connexité.

Pour ce qui est des applications de ce chapitre, je me suis limité à l'analyse numérique linéaire. Pour une application aux groupes de Lie, le lecteur intéressé pourra consulter l'ouvrage de Mnéimné et Testard [12].

Le chapitre 4, qui n'était pas présent dans la première édition, est consacré à l'étude des matrices à coefficients positifs ou strictement positifs avec pour application une étude des matrices stochastiques et doublement stochastiques qui interviennent en théorie des probabilités.

Les chapitres 5 et 6 sont deux chapitres importants de l'analyse numérique linéaire. On s'intéresse aux méthodes directes et itératives de résolution des systèmes linéaires et aux méthodes de calcul approché des valeurs et vecteurs propres d'une matrice carrée réelle ou complexe.

Enfin le chapitre 7 est une application à l'étude des systèmes différentiels linéaires à coefficients constants ou non et à l'exponentielle d'une matrice. L'exponentielle d'une matrice y est définie à partir de l'étude des systèmes différentiels linéaires à coefficients constants.

Cette deuxième édition diffère de la première par la suppression du premier chapitre sur les espaces vectoriels normés et l'ajout d'un chapitre sur les matrices réelles positives. On renvoie à [18], publié chez le même éditeur, pour les résultats sur les espaces vectoriels normés utilisés dans cet ouvrage.

Je tiens à remercier les éditions EDP Sciences pour la confiance qu'ils m'accordent en publiant une deuxième édition de ce travail.

Chapitre 1

Polynômes minimal et caractéristique. Sous espaces caractéristiques

Pour ce chapitre, $\mathbb{K}$ est un corps commutatif et E un $\mathbb{K}$ -espace vectoriel de dimension $n \geq 1$.

Pour toute partie non vide X de E on désigne par $\text{Vect}(X)$ le sous espace vectoriel de E engendré par X , soit l'ensemble de toutes les combinaisons linéaires (finies) d'éléments de X .

On note $\mathcal{L}(E)$ l'algèbre des endomorphismes de E , $\mathcal{M}_n(\mathbb{K})$ l'algèbre des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$ et $GL_n(\mathbb{K})$ le groupe multiplicatif des éléments inversibles de $\mathcal{M}_n(\mathbb{K})$.

On note Id [resp. I_n] l'endomorphisme [resp. la matrice] identité.

Pour tous i, j compris entre 1 et n , on note E_{ij} la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) qui vaut 1. La famille $(E_{i,j})_{1 \leq i, j \leq n}$ est la une base canonique de $\mathcal{M}_n(\mathbb{K})$.

Le choix d'une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E permet de réaliser un isomorphisme d'algèbres de $\mathcal{L}(E)$ sur $\mathcal{M}_n(\mathbb{K})$. Cet isomorphisme est réalisé de la façon suivante : à tout endomorphisme u de E , on associe sa matrice $A = ((a_{ij}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$ dans la base $\mathcal{B}$ définie par :

$$\forall j \in \{1, \dots, n\}, u(e_j) = \sum_{i=1}^n a_{ij} e_i$$

À toute matrice A dans $\mathcal{M}_n(\mathbb{K})$ est associé l'endomorphisme de $\mathbb{K}^n$, que nous noterons encore A :

$$\begin{aligned} A : \quad \mathbb{K}^n &\rightarrow \mathbb{K}^n \\ x = (x_i)_{1 \leq i \leq n} &\mapsto Ax = \left(\sum_{j=1}^n a_{ij} x_j \right)_{1 \leq i \leq n} \end{aligned}$$

On désigne par $\mathbb{K}[X]$ l'anneau des polynômes à coefficients dans $\mathbb{K}$ et par $\mathbb{K}(X)$ son corps des fractions rationnelles.

Un polynôme non nul est dit unitaire si son coefficient dominant est égal à 1.

On rappelle que $\mathbb{K}[X]$ est un anneau euclidien, donc principal et factoriel. Un résultat qui nous sera utile est le théorème de Bézout qui nous dit que deux polynômes A et B sont premiers entre eux dans $\mathbb{K}[X]$ si, et seulement si, il existe deux polynômes U et V dans $\mathbb{K}[X]$ tels que $AU + BV = 1$.

Pour tout endomorphisme u de E , on note $u^0 = Id$ et on définit les puissances successives de u par la relation de récurrence $u^{k+1} = u^k u$ pour tout $k \in \mathbb{N}$, ce qui nous permet de définir, pour tout polynôme $P = \sum_{k=0}^p a_k X^k \in \mathbb{K}[X]$, l'endomorphisme

$P(u) = \sum_{k=0}^p a_k u^k$. La sous algèbre de $\mathcal{L}(E)$ engendrée par u est constituée

des endomorphismes $v = P(u)$ où P est dans $\mathbb{K}[X]$. On note naturellement $\mathbb{K}[u]$ cette algèbre et il est facile de vérifier qu'elle est commutative. Précisément on a :

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

On définit de manière analogue la sous algèbre $\mathbb{K}[A]$ de $\mathcal{M}_n(\mathbb{K})$ engendrée par une matrice $A \in \mathcal{M}_n(\mathbb{K})$. Si A est la matrice de u dans une base $\mathcal{B}$ de E , la matrice de $P(u)$ dans $\mathcal{B}$ est alors $P(A)$.

1.1 Définitions et premières propriétés

L'espace vectoriel $\mathcal{L}(E)$ étant de dimension n^2 , on en déduit que pour tout endomorphisme u de E la famille $(u^k)_{0 \leq k \leq n^2}$ est liée, ce qui se traduit en disant qu'il existe un polynôme P non nul dans $\mathbb{K}[X]$ tel que $P(u) = 0$. Il en résulte que l'ensemble $I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\}$ n'est pas réduit au polynôme nul. Cet ensemble qui est le noyau du morphisme d'algèbres $P \mapsto P(u)$, est un idéal de $\mathbb{K}[X]$. L'anneau $\mathbb{K}[X]$ étant principal on peut donner la définition suivante.

Définition 1.1. Pour tout endomorphisme u de E , on appelle idéal annulateur de u l'idéal I_u et polynôme minimal de u le générateur unitaire de cet idéal. On note π_u ce polynôme.

On a donc $I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\} = \mathbb{K}[X]\pi_u$ et π_u est le polynôme unitaire de plus petit degré annulant u .

On définit de manière analogue le polynôme minimal d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$.

Si $u \in \mathcal{L}(E)$ a pour matrice A dans une base de E , il a alors le même polynôme minimal que A .

Définition 1.2. Soit u un endomorphisme de E . On dit que λ dans $\mathbb{K}$ est une valeur propre de u s'il existe un vecteur non nul x dans E tel que $u(x) = \lambda x$. On dit alors que x est un vecteur propre de u associé à la valeur propre λ et que le sous espace vectoriel de $E_\lambda = \ker(u - \lambda Id)$ de E est le sous espace propre associé à λ . L'ensemble des valeurs propres de u est appelé le spectre de u et noté $\text{Sp}(u)$.

Définition 1.3. Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que λ dans $\mathbb{K}$ est valeur propre de A s'il existe un vecteur non nul x dans $\mathbb{K}^n$ tel que $Ax = \lambda x$. On dit alors que x est un vecteur propre de A associé à la valeur propre λ et que le sous espace vectoriel $E_\lambda = \ker(A - \lambda I_n)$ de $\mathbb{K}^n$ est le sous espace propre associé à λ . L'ensemble des valeurs propres de A est appelé le spectre de A et noté $\text{Sp}(A)$.

Si $u \in \mathcal{L}(E)$ a pour matrice A dans une base de E , un scalaire $\lambda \in \mathbb{K}$ est valeur propre de u si, et seulement si, il est valeur propre de A .

Pour $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$, on a les équivalences :

$$\begin{aligned} (\lambda \in \text{Sp}(A)) &\Leftrightarrow (\ker(A - \lambda I_n) \neq \{0\}) \Leftrightarrow (A - \lambda I_n \notin GL_n(\mathbb{K})) \\ &\Leftrightarrow (\det(A - \lambda I_n) = 0) \end{aligned}$$

La matrice $XI_n - A$ étant un élément de $\mathcal{M}_n(\mathbb{K}[X]) \subset \mathcal{M}_n(\mathbb{K}(X))$, on peut considérer son déterminant $\chi_A(X) = \det(A - XI_n)$ qui est un élément de $\mathbb{K}[X]$. Ce déterminant est le polynôme caractéristique de la matrice A . C'est un polynôme unitaire de degré n .

Une matrice et sa transposée ayant même déterminant, on en déduit qu'elles ont le même polynôme caractéristique.

Pour toute matrice $P \in GL_n(\mathbb{K})$, les matrices $A - XI_n$ et $P^{-1}AP - XI_n = P^{-1}(A - XI_n)P$ sont semblables dans $\mathcal{M}_n(\mathbb{K}(X))$, donc $\chi_A = \chi_{P^{-1}AP}$. On peut donc définir le polynôme caractéristique d'un endomorphisme $u \in \mathcal{L}(E)$ par $\chi_u(X) = \chi_A(X)$, où A est la matrice de u dans une quelconque base de E . On peut noter $\chi_u(X) = \det(u - XI_d)$.

Avec ces notations, le spectre de u [resp. de A] est l'ensemble des racines de son polynôme caractéristique. C'est donc une partie finie de $\mathbb{K}$ ayant au plus n éléments. Ce spectre peut être vide (par exemple pour $\mathbb{K} = \mathbb{R}$) ou pas (par exemple pour $\mathbb{K} = \mathbb{C}$ d'après le théorème de d'Alembert-Gauss).

On rappelle que la trace d'une matrice $A = ((a_{ij}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$ est $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$. Pour A, B dans $\mathcal{M}_n(\mathbb{K})$, on a :

$$\begin{aligned} \text{Tr}(AB) &= \sum_{k=1}^n (AB)_{kk} = \sum_{k=1}^n \sum_{i=1}^n a_{ik} b_{k,i} = \sum_{i=1}^n \sum_{k=1}^n b_{k,i} a_{ik} \\ &= \sum_{i=1}^n (BA)_{ii} = \text{Tr}(BA) \end{aligned}$$

Il en résulte que deux matrices semblables ont même trace. En effet si $B = P^{-1}AP$ avec $P \in GL_n(\mathbb{K})$, on a alors :

$$\text{Tr}(B) = \text{Tr}(P^{-1}(AP)) = \text{Tr}((AP)P^{-1}) = \text{Tr}(A)$$

On peut donc définir la trace de $u \in \mathcal{L}(E)$ comme la trace de sa matrice dans n'importe quelle base de E .

Théorème 1.1.

Si $A = ((a_{ij}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$ admet n valeurs propres $\lambda_1, \dots, \lambda_n$ distinctes ou confondues dans $\mathbb{K}$, on a alors :

$$\det(A) = \prod_{k=1}^n \lambda_k \text{ et } \text{Tr}(A) = \sum_{k=1}^n \lambda_k$$

Preuve. Le développement du déterminant dans $\mathcal{M}_n(\mathbb{K}(X))$ nous donne :

$$\chi_A(X) = \begin{vmatrix} X - a_{11} & \cdots & -a_{1,n} \\ \vdots & \ddots & \vdots \\ -a_{n,1} & \cdots & X - a_{n,n} \end{vmatrix} = X^n - \text{Tr}(A) X^{n-1} + \cdots + (-1)^n \det(A)$$

et dans le cas où χ_A est scindé sur $\mathbb{K}$, on a aussi :

$$\chi_A(X) = \prod_{k=1}^n (X - \lambda_k) = X^n - \left(\sum_{k=1}^n \lambda_k \right) X^{n-1} + \cdots + (-1)^n \prod_{k=1}^n \lambda_k$$

ce qui nous donne, par identification des coefficients de X^{n-1} et des coefficients constants, les égalités $\text{Tr}(A) = \sum_{k=1}^n \lambda_k$ et $\det(A) = \prod_{k=1}^n \lambda_k$. $\square$

Théorème 1.2.

Soient $u \in \mathcal{L}(E)$ et $P \in \mathbb{K}[X]$. Pour toute valeur propre $\lambda \in \text{Sp}(u)$, $P(\lambda)$ est valeur propre de $P(u)$. Pour $\mathbb{K}$ algébriquement clos, on a $\text{Sp}(P(u)) = \{P(\lambda) \mid \lambda \in \text{Sp}(u)\}$.

Preuve. Si $\lambda \in \mathbb{K}$ est une valeur propre de u et $x \in E \setminus \{0\}$ un vecteur propre associé, on vérifie alors facilement que pour tout polynôme $P \in \mathbb{K}[X]$, on a $P(u)x = P(\lambda)x$. En effet, de $u(x) = \lambda x$, on déduit par récurrence sur $k \geq 0$ que $u^k(x) = \lambda^k x$ pour tout $k \in \mathbb{N}$, puis par linéarité, il en résulte que $P(u)x = P(\lambda)x$ pour tout $P \in \mathbb{K}[X]$. Ce qui signifie que x est un vecteur propre de $P(u)$ associé à la valeur propre $P(\lambda)$.

Si $P(X) = a_0$ est un polynôme constant, $P(u) = a_0 \text{Id}$ a alors pour unique valeur propre a_0 , l'espace propre associé étant E .

On suppose que $\mathbb{K}$ est algébriquement clos et que P est non constant ($\text{Sp}(u)$ et $\text{Sp}(P(u))$ sont donc non vides). On a vu que $\{P(\lambda) \mid \lambda \in \text{Sp}(u)\} \subset \text{Sp}(P(u))$. Si $\mu \in \text{Sp}(P(u))$, en notant $Q(X) = P(X) - \mu$, l'endomorphisme $Q(u)$ est non

injectif et en écrivant que $Q(X) = \alpha \prod_{i=1}^p (X - \lambda_i)^{m_i}$ ($\mathbb{K}$ est algébriquement clos),

on en déduit qu'il existe un indice i tel que l'endomorphisme $u - \lambda_i \text{Id}$ soit non injectif ce qui signifie que λ_i est une valeur propre de u , puis de $Q(\lambda_i) = 0$, on déduit que $\mu = P(\lambda_i)$. En définitive on a $\text{Sp}(P(u)) = \{P(\lambda) \mid \lambda \in \text{Sp}(u)\}$. $\square$

Pour $\mathbb{K}$ non algébriquement clos, l'inclusion $\{P(\lambda) \mid \lambda \in \text{Sp}(u)\} \subset \text{Sp}(P(u))$ peut être stricte. Par exemple pour $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ et $P(X) = X^2$, on a $A^2 = -I_2$ et $\text{Sp}(A) = \emptyset$, donc l'inclusion est stricte.

Lemme 1.1 *Soit u un endomorphisme non nul de E . Si F est un sous espace vectoriel de E stable par u , le polynôme caractéristique de la restriction de u à F divise alors celui de u .*

Preuve. Soit $\mathcal{B}_1$ une base de F complétée en une base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ de E . Dans cette base la matrice de u est $A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$ où A_1 est la matrice, dans la base $\mathcal{B}_1$, de la restriction de u à F (F est stable par u) et le polynôme caractéristique de u s'écrit $\chi_u(X) = \det(A_1 - XI_{n_1}) \det(A_3 - XI_{n_3})$. Il en résulte que χ_u est un multiple du polynôme caractéristique de la restriction de u à F . $\square$

Théorème 1.3.

Soient u un endomorphisme de E et $\lambda \in \mathbb{K}$ une valeur propre de u . Si λ a pour multiplicité α en tant que racine du polynôme caractéristique de u , on a alors $1 \leq \dim(\ker(u - \lambda Id)) \leq \alpha$.

Preuve. Pour $\lambda \in \text{Sp}(u)$, le sous-espace vectoriel $E_\lambda = \ker(u - \lambda Id)$ n'est pas réduit au vecteur nul et sa dimension est supérieure ou égale à 1. Ce sous-espace vectoriel étant stable par u , le polynôme caractéristique χ_λ de la restriction de u à E_λ divise le polynôme caractéristique χ_u de u (lemme 1.1). En remarquant que $\chi_\lambda(X) = (\lambda - X)^\delta$ où δ est la dimension de E_λ , on en déduit que $\chi_u(X) = (\lambda - X)^\delta Q(X)$ et la racine λ de χ_u étant de multiplicité α , on a nécessairement $\delta \leq \alpha$. $\square$

Théorème 1.4.

Soit u un endomorphisme de E . Les valeurs propres de u sont les racines de son polynôme minimal.

Preuve. Si $\lambda \in \mathbb{K}$ est une valeur propre de u et x un vecteur propre (non nul) associé, de l'égalité $0 = \pi_u(u)(x) = \pi_u(\lambda)x$ avec $x \neq 0$, on déduit que $\pi_u(\lambda) = 0$, c'est-à-dire que λ est racine de π_u . Réciproquement si λ est racine de π_u , on a alors $\pi_u(X) = (X - \lambda)Q(X)$ et avec $\pi_u(u) = (u - \lambda Id) \circ Q(u) = 0$ et du caractère minimal de π_u on déduit que $u - \lambda Id$ est non inversible, ce qui équivaut à dire que λ est une valeur propre de u . $\square$

Définition 1.4. *Soit u un endomorphisme de E . La multiplicité d'une valeur propre de u en tant que racine de son polynôme minimal est appelée l'indice de cette valeur propre.*

Définition 1.5. On dit qu'un endomorphisme $u \in \mathcal{L}(E)$ [resp. une matrice $A \in \mathcal{M}_n(\mathbb{K})$] est nilpotent [resp. nilpotente] s'il existe un entier r strictement positif tel que $u^{r-1} \neq 0$ et $u^r = 0$. [resp. $A^{r-1} \neq 0$ et $A^r = 0$]. On dit que r est l'ordre de nilpotence de u [resp. de A].

Il est facile de vérifier que 0 est la seule valeur propre d'un endomorphisme nilpotent.

Lemme 1.2 Si $u \in \mathcal{L}(E)$ est nilpotent, on a alors $\text{Tr}(u^k) = 0$ pour tout k compris entre 1 et n . Pour $\mathbb{K}$ de caractéristique nulle, un endomorphisme $u \in \mathcal{L}(E)$ est nilpotent si, et seulement si, $\text{Tr}(u^k) = 0$ pour tout k compris entre 1 et n .

Preuve.

1. On vérifie tout d'abord par récurrence sur la dimension $n \geq 1$ de E , qu'un endomorphisme nilpotent est de trace nulle. Pour $n = 1$, l'unique endomorphisme nilpotent est l'endomorphisme nul et sa trace est nulle. Supposons le résultat acquis pour les espaces vectoriels de dimension au plus égale à $n - 1 \geq 1$ et soit $u \in \mathcal{L}(E)$ nilpotent d'ordre $r \geq 1$ avec E de dimension $n \geq 2$. Comme 0 est valeur propre de u ($u^{r-1} \neq 0$, donc il existe $y \in E$ tel que $x = u^{r-1}(y) \neq 0$ et on a $u(x) = u^r(y) = 0$), il existe un vecteur non nul e_1 dans le noyau de u et complétant ce vecteur en une base $\mathcal{B}_1$ de E , la matrice de u dans $\mathcal{B}_1$ est de la forme $A = \begin{pmatrix} 0 & \alpha \\ 0 & B \end{pmatrix}$ où $\alpha \in \mathcal{M}_{1,n-1}(\mathbb{K})$ et $B \in \mathcal{M}_{n-1}(\mathbb{K})$. Avec $A^r = \begin{pmatrix} 0 & \alpha B^{r-1} \\ 0 & B^r \end{pmatrix} = 0$, on déduit que B est nilpotente et en conséquence $\text{Tr}(B) = 0$ (l'hypothèse de récurrence nous donne le résultat sur $\mathcal{M}_{n-1}(\mathbb{K})$), ce qui entraîne que $\text{Tr}(u) = \text{Tr}(A) = \text{Tr}(B) = 0$.
2. Si $u \in \mathcal{L}(E)$ est nilpotent, il en est alors de même de u^k pour tout entier $k \geq 1$ et en conséquence, $\text{Tr}(u^k) = 0$.
3. Pour la réciproque avec $\mathbb{K}$ de caractéristique nulle, on procède encore par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$, on a $u(x) = \lambda x$, $\text{Tr}(u) = \lambda$ et le résultat est trivial. Supposons le résultat acquis pour les espaces vectoriels de dimension au plus égale à $n - 1 \geq 1$ et soit $u \in \mathcal{L}(E)$ tel que $\text{Tr}(u^k) = 0$ pour tout k compris entre 1 et $n = \dim(E) \geq 2$. En désignant par $\chi_u(X) = \sum_{k=0}^n a_k X^k$ le polynôme caractéristique de u et en tenant compte de $\chi_u(u) = \sum_{k=0}^n a_k u^k = 0$ et $\text{Tr}(u^k) = 0$ pour $k = 1, \dots, n$, on déduit que $\text{Tr}(\chi_u(u)) = na_0 = 0$ et $a_0 = \det(u) = 0$ puisque $\mathbb{K}$ de caractéristique nulle. Donc 0 est valeur propre de u et il existe une base $\mathcal{B}$ de E , dans laquelle la matrice de u est de la forme $A = \begin{pmatrix} 0 & \alpha \\ 0 & B \end{pmatrix}$ où $\alpha \in \mathcal{M}_{1,n-1}(\mathbb{K})$ et $B \in \mathcal{M}_{n-1}(\mathbb{K})$. Avec $A^k = \begin{pmatrix} 0 & \alpha B^{k-1} \\ 0 & B^k \end{pmatrix}$, on déduit que $\text{Tr}(B^k) = \text{Tr}(A^k) = \text{Tr}(u^k) = 0$ pour tout $k = 1, \dots, n$ et l'hypothèse de récurrence nous dit que B est nilpotente.

Enfin, en notant p l'indice de nilpotence de B , avec $A^{p+1} = \begin{pmatrix} 0 & \alpha B^p \\ 0 & B^{p+1} \end{pmatrix} = 0$, on déduit que A est nilpotente et il en est de même de u .

On peut aussi procéder comme suit en écrivant le polynôme minimal de u sous la forme $\pi_u(X) = X^r Q(X)$ avec $Q(0) \neq 0$. Le théorème de décomposition des noyaux (théorème 1.11) nous dit que $E = F \oplus G$, où les espaces $F = \ker(u^r)$ et $G = \ker(Q(u))$ sont stables par u (commutativité de $\mathbb{K}[u]$). Si Q est non

constant, il s'écrit alors $Q(X) = X^{p-r} + \sum_{k=0}^{p-r-1} a_k X^k$ avec $0 \leq r \leq p-1$ et on a

$$0 = Q(u)|_G = Q(u|_G), \text{ donc } \text{Tr}(Q(u|_G)) = \text{Tr}(u|_G^{p-r}) + \sum_{k=0}^{p-r-1} a_k \text{Tr}(u|_G^k) = 0$$

et $\text{Tr}(u|_G^{p-r}) + \sum_{k=1}^{p-r-1} a_k \text{Tr}(u|_G^k) = -a_0 \dim(G) \neq 0$ (on est en caractéristique nulle et $a_0 = Q(0)$). Il existe donc un entier k compris entre 1 et $p-r \leq n$ tel que $\text{Tr}(u|_G^k) \neq 0$. Utilisant la matrice de u^k dans une base adaptée à la somme directe $E = F \oplus G$, F et G étant stables par u^k , on aboutit à $\text{Tr}(u^k) = \text{Tr}(u|_F^k) + \text{Tr}(u|_G^k) = \text{Tr}(u|_G^k) \neq 0$, ce qui n'est pas. Le polynôme Q est donc constant égal à 1, ce qui nous donne $\pi_u(X) = X^r$ et signifie que u est nilpotent d'ordre r .

□

1.2 Localisation des valeurs propres d'une matrice complexe

Pour ce paragraphe, $\mathbb{K}$ désigne le corps des nombres complexes et l'espace vectoriel $\mathbb{C}^n$ est muni de la norme $x \mapsto \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$.

Pour $\lambda \in \mathbb{C}$ et $R \in \mathbb{R}^+$, on note $\overline{D}(\lambda, R) = \{z \in \mathbb{C} \mid |z - \lambda| \leq R\}$ le disque fermé de centre λ et de rayon R dans le plan complexe.

Pour $n \geq 2$, $A \in \mathcal{M}_n(\mathbb{C})$ et i, j compris entre 1 et n , on note :

$$L_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad C_j = \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}|, \quad L = \max_{1 \leq i \leq n} (L_i + |a_{ii}|), \quad C = \max_{1 \leq j \leq n} (C_j + |a_{jj}|)$$

Théorème 1.5. Gerschgorin-Hadamard

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on a $\text{Sp}(A) \subset \bigcup_{i=1}^n \overline{D}(a_{ii}, L_i)$.

Preuve. Soient $\lambda \in \text{Sp}(A)$ et $x \in \mathbb{C}^n$ un vecteur propre associé tel que $\|x\|_\infty = 1$. Pour $i \in \{1, \dots, n\}$ tel que $|x_i| = \|x\|_\infty$, on a :

$$|\lambda - a_{ii}| = |(\lambda - a_{ii})x_i| = \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \right| \leq \left(\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right) \|x\|_\infty = L_i$$

soit $\lambda \in \overline{D}(a_{ii}, L_i)$. □

On a aussi l'inclusion $\text{Sp}(A) = \text{Sp}({}^t A) \subset \bigcup_{j=1}^n \overline{D}(a_{jj}, C_j)$.

Les disques fermés $\overline{D}(a_{ii}, L_i)$ sont les disques de Gerschgorin.

L'exercice 1.9 est une application du théorème de Gerschgorin-Hadamard au calcul des valeurs propres d'une matrice.

Corollaire 1.1 : Pour toute valeur propre $\lambda \in \mathbb{C}$ de $A \in \mathcal{M}_n(\mathbb{C})$ on a $|\lambda| \leq \min(L, C)$.

Preuve. Pour $\lambda \in \text{Sp}(A)$ et $i \in \{1, \dots, n\}$ tel que $|\lambda - a_{ii}| \leq L_i$, on a :

$$|\lambda| \leq |\lambda - a_{ii}| + |a_{ii}| \leq L_i + |a_{ii}| \leq L$$

Remplaçant A par sa transposées, on a aussi $|\lambda| \leq C$, donc $|\lambda| \leq \min(L, C)$. □

Définition 1.6. Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite à diagonale strictement dominante si :

$$\forall i \in \{1, \dots, n\}, |a_{ii}| > L_i$$

Les matrices à diagonale strictement dominante se rencontrent dans de nombreux problèmes, par exemple dans le problème de l'interpolation par des fonctions splines cubiques ou dans les problèmes de résolutions d'équations aux dérivées partielles par des méthodes de discrétisation par différences finies (voir [17]).

Corollaire 1.2 : Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ à diagonale strictement dominante est inversible.

Preuve. Soient $A \in \mathcal{M}_n(\mathbb{C})$, $\lambda \in \text{Sp}(A)$ et i compris entre 1 et n tel que $|\lambda - a_{ii}| \leq L_i$. Dans le cas où A est à diagonale strictement dominante, on ne peut avoir $\lambda = 0$. On a donc $\text{Sp}(A) \subset \mathbb{C}^*$, ce qui implique que A est inversible. □

Une généralisation du théorème de Gerschgorin et Hadamard est le théorème d'Ostrowski qui suit.

Lemme 1.3 Soit A dans $\mathcal{M}_n(\mathbb{C})$. S'il existe un réel $\alpha \in [0, 1]$ tel que :

$$\forall i \in \{1, \dots, n\}, |a_{ii}| > L_i^\alpha C_i^{1-\alpha}$$

la matrice A est alors inversible.

Preuve. Pour $\alpha = 1$, il s'agit du corollaire 1.2 et pour $\alpha = 0$, c'est encore le corollaire 1.2 appliqué à ${}^t A$.

On suppose que $\alpha \in]0, 1[$ et que A est non inversible, ce qui revient à dire que 0 est valeur propre de A . Si $x \in \mathbb{C}^n \setminus \{0\}$ est un vecteur propre non nul associé, il est alors solution non nulle du système linéaire :

$$\sum_{j=1}^n a_{ij} x_j = 0 \quad (1 \leq i \leq n)$$

et on a :

$$|a_{ii}| |x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j| \quad (1 \leq i \leq n)$$

Tenant compte de $L_i^\alpha C_i^{1-\alpha} < |a_{ii}|$, pour tout $i \in \{1, \dots, n\}$, on en déduit que :

$$L_i^\alpha C_i^{1-\alpha} |x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j| \quad (1 \leq i \leq n)$$

l'inégalité étant stricte pour tous les indices i tels que $x_i \neq 0$. Utilisant l'inégalité de Hölder, on obtient pour $1 \leq i \leq n$:

$$\begin{aligned} L_i^\alpha C_i^{1-\alpha} |x_i| &\leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|^\alpha \left(|a_{ij}|^{1-\alpha} |x_j| \right) \\ &\leq \left(\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right)^\alpha \left(\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} \right)^{1-\alpha} = L_i^\alpha \left(\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} \right)^{1-\alpha} \end{aligned}$$

Pour $x_i \neq 0$, l'inégalité est stricte et nécessairement $L_i > 0$. On en déduit donc que :

$$C_i |x_i|^{\frac{1}{1-\alpha}} \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} \quad (1 \leq i \leq n)$$

l'inégalité étant stricte pour $x_i \neq 0$ et évidente pour $x_i = 0$. En additionnant ces inégalités, on aboutit à :

$$S = \sum_{i=1}^n C_i |x_i|^{\frac{1}{1-\alpha}} < \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |x_j|^{\frac{1}{1-\alpha}} = \sum_{j=1}^n C_j |x_j|^{\frac{1}{1-\alpha}} = S$$

ce qui est impossible. La matrice A est donc nécessairement inversible. $\square$

Théorème 1.6. Ostrowski

Pour tout réel $\alpha \in [0, 1]$ et toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on a :

$$\mathrm{Sp}(A) \subset \bigcup_{i=1}^n \overline{D}(a_{ii}, L_i^\alpha C_i^{1-\alpha})$$

Preuve. Pour $\lambda \in \mathrm{Sp}(A)$, la matrice $A - \lambda I_n$ est non inversible et le lemme précédent nous dit que pour tout réel $\alpha \in [0, 1]$, il existe un indice $i \in \{1, \dots, n\}$ tel que $|a_{ii} - \lambda| \leq L_i^\alpha C_i^{1-\alpha}$. $\square$

Corollaire 1.3 : Pour $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda \in \mathrm{Sp}(A)$, il existe $i \in \{1, \dots, n\}$ tel que $|\lambda|^2 \leq (L_i + |a_{ii}|)(C_i + |a_{ii}|)$.

Preuve. Prenant $\alpha = \frac{1}{2}$ dans le théorème d'Ostrowski, on peut trouver i compris entre 1 et n tel que $|a_{ii} - \lambda| \leq \sqrt{L_i C_i}$, ce qui implique que :

$$|\lambda| \leq |a_{ii}| + \sqrt{L_i C_i} \leq \sqrt{(|a_{ii}| + L_i)(|a_{ii}| + C_i)}$$

la dernière inégalité résultant de $2\sqrt{L_i C_i} \leq C_i + L_i$. $\square$

1.3 Matrice compagnon d'un polynôme

$\mathbb{K}$ est à nouveau un corps commutatif.

À tout polynôme unitaire $P(X) = X^n - \sum_{k=0}^{n-1} a_k X^k$ de degré $n \geq 1$ dans $\mathbb{K}[X]$, on associe sa matrice compagnon $C_P \in \mathcal{M}_n(\mathbb{K})$ définie par :

$$C_P = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & a_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & a_{n-1} \end{pmatrix}$$

Pour $n = 1$, on a $P(X) = X - a_0$ et $C_P = (a_0)$.

Une telle matrice est aussi appelée matrice de Frobenius.

Pour ce paragraphe, on se fixe $P \in \mathbb{K}[X]$ unitaire de degré $n \geq 1$ et u_P est l'endomorphisme de $\mathbb{K}^n$ de matrice C_P dans la base canonique $\mathcal{B} = (e_k)_{1 \leq k \leq n}$.

On a donc $u_P(e_{k-1}) = e_k$ pour k compris entre 2 et n et $u_P(e_n) = \sum_{k=0}^{n-1} a_k e_{k+1}$. Il

en résulte que $e_k = u_P^{k-1}(e_1)$ pour k compris entre 1 et n .

Lemme 1.4 Pour tout polynôme $Q \in \mathbb{K}[X]$, on a $Q(u_P) = 0$ si, et seulement si, $Q(u_P)(e_1) = 0$.

Preuve. Si $Q(u_P) = 0$, on a alors en particulier $Q(u)(e_1) = 0$. Réciproquement, si $Q(u_P)(e_1) = 0$, on a alors pour tout k compris entre 2 et n :

$$Q(u_P)(e_k) = Q(u_P)(u_P^{k-1}(e_1)) = u_P^{k-1}(Q(u_P)(e_1)) = 0$$

donc $Q(u) = 0$. □

Théorème 1.7.

Un polynôme unitaire P est polynôme minimal et polynôme caractéristique de sa matrice compagnon C_P .

Preuve. L'égalité $u_P(e_n) = \sum_{k=0}^{n-1} a_k e_{k+1}$ se traduit par $u_P^n(e_1) = \sum_{k=0}^{n-1} a_k u_P^k(e_1)$, soit $P(u_P)(e_1) = 0$, ce qui équivaut à $P(u_P) = 0$ d'après le lemme précédent. Le polynôme minimal π_{u_P} divise donc P et $\deg(\pi_{u_P}) \leq n$. La famille de vecteurs $(e_k)_{1 \leq k \leq n} = (u_P^{k-1}(e_1))_{1 \leq k \leq n}$ étant libre, il n'existe pas de polynôme $Q \in \mathbb{K}_{n-1}[X] \setminus \{0\}$ tel que $Q(u_P) = 0$, donc $\deg(\pi_{u_P}) \geq n$. Le polynôme minimal π_{u_P} est donc de degré n et égal à P puisque unitaire divisant P . On a donc $\pi_{C_P} = \pi_{u_P} = P$.

Le calcul du polynôme caractéristique χ_{C_P} peut se faire en effectuant l'opération élémentaire $L_1 \leftarrow L_1 + XL_2 + X^2L_3 + \dots + X^{n-1}L_n$ dans $\mathcal{M}_n(\mathbb{K}(X))$ qui donne :

$$\begin{aligned} \chi_{C_P}(X) &= \begin{vmatrix} X & \cdots & 0 & -a_0 \\ -1 & \ddots & \vdots & -a_1 \\ \vdots & \ddots & X & \vdots \\ 0 & \cdots & -1 & X - a_{n-1} \end{vmatrix} = \begin{vmatrix} 0 & \cdots & 0 & P(X) \\ -1 & \ddots & \vdots & -a_1 \\ \vdots & \ddots & X & \vdots \\ 0 & \cdots & -1 & X - a_{n-1} \end{vmatrix} \\ &= (-1)^{n+1} P(X) (-1)^{n-1} = P(X) \end{aligned}$$

On peut aussi procéder comme suit. En notant $\chi_{(a_0, \dots, a_{n-1})}(X)$ le polynôme caractéristique de C_P et en le développant par rapport à la première ligne, on a :

$$\chi_{(a_0, \dots, a_{n-1})}(X) = \begin{vmatrix} X & \cdots & 0 & -a_0 \\ -1 & \ddots & \vdots & -a_1 \\ \vdots & \ddots & X & \vdots \\ 0 & \cdots & -1 & X - a_{n-1} \end{vmatrix} = X \cdot \chi_{(a_1, \dots, a_{n-1})}(X) - a_0$$

puis par récurrence, on obtient $\chi_{(a_0, \dots, a_{n-1})}(X) = X^n - \sum_{k=0}^{n-1} a_k X^k$. □

Corollaire 1.4 : C_P est inversible si, et seulement si, $P(0) \neq 0$. Dans ce cas, on a :

$$C_P^{-1} = \frac{1}{a_0} \left(C_P^{n-1} - \sum_{k=1}^{n-1} a_k C_P^{k-1} \right) = \begin{pmatrix} -\frac{a_1}{a_0} & 1 & \cdots & 0 \\ -\frac{a_2}{a_0} & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 \\ \frac{1}{a_0} & 0 & \cdots & 0 \end{pmatrix}$$

$$\text{et } \pi_{C_P^{-1}}(X) = \chi_{C_P^{-1}}(X) = -\frac{1}{a_0} X^n P\left(\frac{1}{X}\right)$$

Preuve. On a $\det(C_P) = (-1)^n \chi_{C_P}(0) = (-1)^n P(0) = (-1)^{n+1} a_0$, donc C_P est inversible si, et seulement si, $a_0 = P(0) \neq 0$. Dans ce cas, de l'égalité

$$P(C_P) = \pi_{C_P}(C_P) = 0, \text{ on déduit que } C_P^{-1} = \frac{1}{a_0} \left(C_P^{n-1} - \sum_{k=1}^{n-1} a_k C_P^{k-1} \right), \text{ donc}$$

$$u_P^{-1}(e_1) = \frac{1}{a_0} \left(e_n - \sum_{k=1}^{n-1} a_k e_k \right) \text{ et } u_P^{-1}(e_k) = u_P^{-1}(u_P(e_{k-1})) = e_{k-1} \text{ pour } k \text{ compris entre 2 et } n, \text{ ce qui nous donne :}$$

$$C_P^{-1} = \begin{pmatrix} -\frac{a_1}{a_0} & 1 & \cdots & 0 \\ -\frac{a_2}{a_0} & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 \\ \frac{1}{a_0} & 0 & \cdots & 0 \end{pmatrix}$$

Dans la base $(f_1, \dots, f_n) = (e_n, e_{n-1}, \dots, e_2, e_1)$, la matrice de u_P^{-1} est :

$$C_Q = \begin{pmatrix} 0 & \cdots & 0 & \frac{1}{a_0} \\ 1 & \ddots & \vdots & -\frac{a_{n-1}}{a_0} \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & -\frac{a_1}{a_0} \end{pmatrix}$$

soit la matrice compagnon du polynôme $Q(X) = X^n - \sum_{k=0}^{n-1} \alpha_k X^k$, où $\alpha_0 = \frac{1}{a_0}$ et

$\alpha_k = -\frac{a_{n-k}}{a_0}$ pour $1 \leq k \leq n-1$. On a donc :

$$\pi_{C_P^{-1}}(X) = \chi_{C_P^{-1}}(X) = -\frac{1}{a_0} X^n \left(\frac{1}{X^n} - \sum_{j=0}^{n-1} a_j \frac{1}{X^j} \right) = -\frac{1}{a_0} X^n P\left(\frac{1}{X}\right)$$

□

1.4 Le théorème de Cayley-Hamilton

$\mathbb{K}$ est un corps commutatif et E un $\mathbb{K}$ -espace vectoriel de dimension $n \geq 1$.

Dans ce qui suit, $u \in \mathcal{L}(E)$ peut être remplacé par $A \in \mathcal{M}_n(\mathbb{K})$.

On se donne un endomorphisme $u \in \mathcal{L}(E) \setminus \{0\}$ et à tout vecteur $x \in E \setminus \{0\}$, on associe les ensembles :

$$I_{u,x} = \{P \in \mathbb{K}[X] \mid P(u)(x) = 0\}, \quad E_{u,x} = \text{Vect} \{u^k(x) \mid k \in \mathbb{N}\}$$

On dit que $E_{u,x}$ est le sous espace cyclique de E engendré par x .

On désigne par $p_{u,x}$ le plus grand entier naturel non nul tel que la famille de vecteurs $\mathcal{B}_{u,x} = (u^k(x))_{0 \leq k \leq p_{u,x}-1}$ soit libre. La famille de vecteurs $(u^k(x))_{0 \leq k \leq p_{u,x}}$

étant liée, il existe des scalaires $a_0, \dots, a_{p_{u,x}-1}$ tels que $u^{p_{u,x}}(x) = \sum_{k=0}^{p_{u,x}-1} a_k u^k(x)$

(caractère minimal de $p_{u,x}$). On note $\pi_{u,x}(X) = X^{p_{u,x}} - \sum_{k=0}^{p_{u,x}-1} a_k X^k$ dans $\mathbb{K}[X]$.

Lemme 1.5 Avec les notations qui précèdent, $\mathcal{B}_{u,x}$ est une base de $E_{u,x}$ et $\pi_{u,x}$ est le polynôme minimal et le polynôme caractéristique de la restriction de u à $E_{u,x}$.

Preuve. On a $\text{Vect}(\mathcal{B}_{u,x}) \subset E_x$ et par définition de l'entier $p_{u,x}$, $\mathcal{B}_{u,x}$ est une base de $\text{Vect}(\mathcal{B}_{u,x})$. On a $u^{p_{u,x}}(x) \in \text{Vect}(\mathcal{B}_{u,x})$ et on vérifie facilement par récurrence que $u^{p_{u,x}+k}(x) \in \text{Vect}(\mathcal{B}_{u,x})$ pour tout entier naturel k . On a donc l'égalité $\text{Vect}(\mathcal{B}_{u,x}) = E_{u,x}$.

En écrivant que la matrice de $u|_{E_{u,x}}$ dans la base $\mathcal{B}_{u,x}$ est :

$$A_{u,x} = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & a_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & a_{p_x-1} \end{pmatrix}$$

soit la matrice compagnon de $\pi_{u,x}$, on en déduit que $\pi_{u,x}$ est le polynôme minimal et le polynôme caractéristique de $u|_{E_{u,x}}$. $\square$

Théorème 1.8. Cayley-Hamilton

Pour tout $u \in \mathcal{L}(E)$ de polynôme caractéristique χ_u , on a $\chi_u(u) = 0$.

Preuve. Le résultat étant évident pour $u = 0$, on suppose u non nul. Pour x dans $E \setminus \{0\}$, le sous espace cyclique $E_{u,x}$ étant stable par u , le polynôme caractéristique $\pi_{u,x}$ de $u|_{E_{u,x}}$ divise celui de u . Il existe donc $Q \in \mathbb{K}[X]$ tel que $\chi_u = Q \pi_{u,x}$ et on a $\chi_u(u)(x) = Q(u) \circ \pi_{u,x}(u)(x) = 0$. On a donc ainsi montré que pour tout $x \in E$, on a $\chi_u(u)(x) = 0$ (c'est clair pour $x = 0$), ce qui équivaut à $\chi_u(u) = 0$. $\square$

Corollaire 1.5 : Pour tout $u \in \mathcal{L}(E)$, le polynôme minimal π_u divise le polynôme caractéristique χ_u .

Preuve. Le polynôme minimal divisant tout polynôme annulateur de u , on déduit du théorème de Cayley-Hamilton que π_u divise le polynôme caractéristique χ_u . $\square$

Du corollaire précédent, on déduit que $\deg(\pi_u) \leq n$. On retrouve aussi le fait que les racines de π_u sont valeurs propres de u .

Pour $\mathbb{K}$ algébriquement clos (par exemple, $\mathbb{K} = \mathbb{C}$) le polynôme caractéristique de l'endomorphisme u s'écrit $\chi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$ où les α_k sont dans $\mathbb{N}^*$ et les λ_k dans $\mathbb{K}$ deux à deux distincts. Le polynôme minimal π_u étant un diviseur de χ_u avec les mêmes racines, il s'écrit $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_k \leq \alpha_k$.

Dans le cas où l'endomorphisme u est inversible, le théorème de Cayley-Hamilton nous donne un moyen de calculer l'inverse de u , si on connaît son polynôme caractéristique P_u . En effet l'égalité $\chi_u(u) = 0$ avec $\chi_u(X) = \sum_{k=0}^n a_k X^k$ donne :

$$u^{-1} = -\frac{1}{a_0} \sum_{k=1}^n a_k u^{k-1} = -\frac{1}{\det(u)} \sum_{k=1}^n a_k u^{k-1}$$

On peut aussi remarquer que u^{-1} est un polynôme en u . Ce résultat étant aussi valable pour une matrice inversible, on en déduit que l'inverse d'une matrice triangulaire supérieure [resp. inférieure] et inversible est une matrice triangulaire supérieure [resp. inférieure].

Le théorème de Cayley-Hamilton permet également de calculer u^p pour tout entier p supérieur ou égal à n en fonction de $Id, u, \dots, u^{n-1}$. En effet pour $p = n$, de $\chi_u(u) = \sum_{k=0}^n a_k u^k = 0$ avec $a_n = 1$, on déduit que $u^n = -\sum_{k=0}^{n-1} a_k u^k$ et pour $p > n$ la division euclidienne de X^p par χ_A , $X^p = Q\chi_A + R$ avec $R = 0$ ou $R \neq 0$ et $\deg(R) < n$, donne $u^p = R(u)$.

1.5 Méthodes de calcul du polynôme caractéristique d'une matrice complexe

On suppose le corps $\mathbb{K}$ de caractéristique nulle et on se donne $A \in \mathcal{M}_n(\mathbb{K})$ de polynôme caractéristique $\chi_A(X) = \sum_{k=0}^n a_k X^k$.

Le théorème de Cayley-Hamilton nous dit que pour tout vecteur $u \in \mathbb{K}^n \setminus \{0\}$, on a $\sum_{k=0}^n a_k A^k u = 0$, ce qui fait apparaître, en posant $u_k = A^k u$, les coefficients a_k

comme solutions du système linéaire $\sum_{k=0}^{n-1} a_k u_k = -u_n$. En prenant u de manière

aléatoire, on a toutes les chances que le système ci-dessus soit non dégénéré de sorte que les coefficients du polynôme caractéristique s'obtiennent comme solution d'un système de Cramer. C'est la méthode de Krylov.

On peut ensuite utiliser, pour les matrices réelle ou complexe de taille raisonnable, une méthode numérique de résolution d'une équation polynomiale pour en déduire des valeurs approchées des valeurs propres.

Les formules de Newton qui suivent pour un polynôme scindé sur $\mathbb{K}$ nous fournissent deux autres méthodes de calcul du polynôme caractéristique.

Théorème 1.9. Formules de Newton

Soit $P(X) = \sum_{k=0}^n a_k X^k = \prod_{k=1}^n (X - \lambda_k) \in \mathbb{K}[X]$ un polynôme unitaire de degré $n \geq 2$ scindé sur $\mathbb{K}$. On lui associe les suites $(\sigma_{n,k})_{1 \leq k \leq n}$ et $(S_{n,k})_{1 \leq k \leq n}$ définies par :

$$\sigma_{n,k} = \sum_{1 \leq i_1 < \dots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_k}, \quad S_{n,k} = \sum_{i=1}^n \lambda_i^k \quad (1 \leq k \leq n)$$

Pour tout entier k compris entre 1 et n on a :

$$S_{n,k} + a_{n-1} S_{n,k-1} + \dots + a_{n-(k-1)} S_{n,1} + a_{n-k} k = 0$$

Preuve. Les $\sigma_{n,k}$ sont les fonctions symétriques élémentaires des racines et en développant le produit $\prod_{k=1}^n (X - \lambda_k)$, on vérifie facilement que $a_{n-k} = (-1)^k \sigma_{n,k}$ pour k compris entre 0 et n , en notant $\sigma_{n,0} = 1$.

En notant, pour j compris entre 1 et n , $P_j(X) = \prod_{\substack{k=1 \\ k \neq j}}^n (X - \lambda_k)$, on a :

$$(X - \lambda_j) P_j(X) = P(X) - P(\lambda_j) = \sum_{k=0}^{n-1} (-1)^k \sigma_{n,k} (X^{n-k} - \lambda_j^{n-k})$$

soit $P_j(X) = \sum_{k=0}^{n-1} (-1)^k \sigma_{n,k} \frac{X^{n-k} - \lambda_j^{n-k}}{X - \lambda_j}$ avec :

$$\frac{X^{n-k} - \lambda_j^{n-k}}{X - \lambda_j} = X^{n-k-1} + X^{n-k-2} \lambda_j + \dots + X \lambda_j^{n-k-2} + \lambda_j^{n-k-1}$$

pour k compris entre 0 et $n-1$. On déduit que pour tout entier k compris entre 0 et $n-1$ le coefficient de X^{n-1-k} dans P_j est donné par :

$$a_{n-1-k,j} = \sum_{i=0}^k (-1)^i \sigma_{n,i} \lambda_j^{k-i}$$

D'autre part, le polynôme dérivé de P est donné par :

$$P'(X) = \sum_{k=0}^{n-1} (-1)^k (n-k) \sigma_{n,k} X^{n-k-1} = \sum_{j=1}^n P_j$$

et en identifiant les coefficients de X^{n-k-1} , pour tout entier k compris entre 0 et $n-1$, on en déduit que :

$$\begin{aligned} (-1)^k (n-k) \sigma_{n,k} &= \sum_{j=1}^n a_{n-1-k,j} = \sum_{j=1}^n \sum_{i=0}^k (-1)^i \sigma_{n,i} \lambda_j^{k-i} \\ &= \sum_{i=0}^k (-1)^i \sigma_{n,i} \sum_{j=1}^n \lambda_j^{k-i} = \sum_{i=0}^k (-1)^i \sigma_{n,i} S_{n,k-i} \end{aligned}$$

en notant $S_{n,0} = n$, ce qui s'écrit aussi :

$$S_{n,k} - \sigma_{n,1} S_{n,k-1} + \cdots + (-1)^{k-1} \sigma_{n,k-1} S_{n,1} + (-1)^k \sigma_{n,k} k = 0$$

En utilisant les égalités $a_{n-k} = (-1)^k \sigma_{n,k}$, on déduit que, pour tout entier k compris entre 1 et $n-1$, on a :

$$S_{n,k} + a_{n-1} S_{n,k-1} + \cdots + a_{n-(k-1)} S_{n,1} + a_{n-k} k = 0$$

Pour $k = n$ l'identité s'obtient en écrivant que pour tout entier j compris entre 1 et n on a :

$$0 = P(\lambda_j) = \lambda_j^n + a_{n-1} \lambda_j^{n-1} + \cdots + a_1 \lambda_j + a_0$$

et en faisant la somme de toutes ces égalités. $\square$

Pour $A \in \mathcal{M}_n(\mathbb{C})$ de polynôme caractéristique $\chi_A(X) = \sum_{k=0}^n a_k X^k = \prod_{k=1}^n (X - \lambda_k)$ avec $a_n = 1$, où les λ_k sont les valeurs propres de A comptées avec leur multiplicité, les valeurs propres de A^k , pour $k \in \mathbb{N}$ sont les λ_i^k et on a $S_{n,k} = \sum_{i=1}^n \lambda_i^k = \text{Tr}(A^k)$, de sorte que les formules de Newton pour χ_A s'écrivent :

$$\text{Tr}(A^k) + a_{n-1} \text{Tr}(A^{k-1}) + \cdots + a_{n-(k-1)} \text{Tr}(A) + a_{n-k} k = 0$$

pour k compris entre 1 et n . Les coefficients a_{n-k} du polynôme caractéristique χ_A apparaissent donc comme solutions du système triangulaire inférieur :

$$\begin{cases} a_{n-1} = -\text{Tr}(A) \\ a_{n-2} = -\frac{1}{2} (\text{Tr}(A^2) + a_{n-1} \text{Tr}(A)) \\ \vdots \\ a_{n-k} = -\frac{1}{k} (\text{Tr}(A^k) + a_{n-1} \text{Tr}(A^{k-1}) + \cdots + a_{n-(k-1)} \text{Tr}(A)) \\ \vdots \\ a_0 = -\frac{1}{n} (\text{Tr}(A^n) + a_{n-1} \text{Tr}(A^{n-1}) + \cdots + a_1 \text{Tr}(A)) \end{cases}$$

C'est la méthode de Leverrier.

La méthode de Souriau nous donne un autre algorithme de χ_A et de l'inverse dans le cas où A est inversible.

On définit les suites de matrices $(A_k)_{1 \leq k \leq n}$, $(B_k)_{1 \leq k \leq n}$ et la suite de scalaires $(\alpha_k)_{1 \leq k \leq n}$ de la manière suivante :

$$\begin{cases} A_1 = A & \alpha_1 = -\operatorname{Tr}(A_1) & B_1 = A_1 + \alpha_1 I_n \\ A_k = B_{k-1}A & \alpha_k = -\frac{\operatorname{Tr}(A_k)}{k} & B_k = A_k + \alpha_k I_n \quad (2 \leq k \leq n) \end{cases}$$

Théorème 1.10.

Avec les notations qui précèdent, on a $\chi_A(X) = X^n + \sum_{k=1}^n \alpha_k X^{n-k}$. Si de plus A est inversible, son inverse est alors $A^{-1} = -\frac{1}{\alpha_n} B_{n-1}$.

Preuve. En posant $\alpha_0 = 1$, il s'agit de montrer que $\alpha_k = a_{n-k}$ pour k compris entre 0 et n . Pour ce faire, on procède par récurrence sur k compris entre 0 et n .

Pour $k = 0$, on a $\alpha_0 = 1 = a_n$ et pour $k = 1$, on a $\alpha_1 = -\operatorname{Tr}(A) = a_{n-1}$. Supposons que $\alpha_j = a_{n-j}$ pour $j = 0, \dots, k-1$, avec $k > 1$. Par définition, on a $A_k = B_{k-1}A = A_{k-1}A + \alpha_{k-1}A$ et par récurrence, il vient :

$$A_k = A^k + \alpha_1 A^{k-1} + \dots + \alpha_{k-1} A = A^k + a_{n-1} A^{k-1} + \dots + a_{n-(k-1)} A$$

ce qui nous donne en utilisant les formules de Newton :

$$\operatorname{Tr}(A_k) = \operatorname{Tr}(A^k) + a_{n-1} \operatorname{Tr}(A^{k-1}) + \dots + a_{n-(k-1)} \operatorname{Tr}(A) = -a_{n-k} k$$

et en conséquence, $\alpha_k = -\frac{\operatorname{Tr}(A_k)}{k} = a_{n-k}$.

Pour $k = n$, on a $A_n = A^n + a_{n-1} A^{n-1} + \dots + a_1 A$ et $B_n = A_n + \alpha_n I_n = A_n + a_0 I_n = \chi_A(A) = 0$ soit $A_n = B_{n-1}A = -\alpha_n I_n$. Dans le cas où A est inversible, on a $\alpha_n = a_0 = \det(A) \neq 0$ et $A^{-1} = -\frac{1}{\alpha_n} B_{n-1}$. $\square$

1.6 Sous espaces caractéristiques

Théorème 1.11. Décomposition des noyaux

Soient r un entier supérieur ou égal à 2 et $P = \prod_{k=1}^r P_k$ dans $\mathbb{K}[X]$, les polynômes P_k étant deux à deux premiers entre eux. Pour tout endomorphisme u de E on a $\ker(P(u)) = \bigoplus_{k=1}^r \ker(P_k(u))$.

Preuve. On procède par récurrence sur $r \geq 2$. Pour $r = 2$, les polynômes P_1 et P_2 sont premiers entre eux et le théorème de Bézout nous dit qu'il existe A, B dans $\mathbb{K}[X]$ tels que $AP_1 + BP_2 = 1$. On a alors $A(u) \circ P_1(u) + B(u) \circ P_2(u) = Id$ et pour tout $x \in \ker(P(u))$, on a :

$$x = (A(u) \circ P_1(u))(x) + (B(u) \circ P_2(u))(x) \quad (1.1)$$

où $x_2 = (A(u) \circ P_1(u))(x) \in \ker(P_2(u))$, $x_1 = (B(u) \circ P_2(u))(x) \in \ker(P_1(u))$, du fait de la commutativité de $\mathbb{K}[u]$. On a donc ainsi montré l'égalité :

$$\ker(P(u)) = \ker(P_1(u)) + \ker(P_2(u))$$

l'inclusion $\ker(P_1(u)) + \ker(P_2(u)) \subset \ker(P(u))$ étant évidente. De l'égalité (1.1) on déduit immédiatement que $\ker(P_1(u)) \cap \ker(P_2(u)) = \{0\}$. On a donc bien la somme directe $\ker(P(u)) = \ker(P_1(u)) \oplus \ker(P_2(u))$. On conclut par récurrence sur r en considérant que si P_1 est premier avec $P_2, \dots, P_r$, il est alors premier avec leur produit. $\square$

Dans le cas où $P(u) = 0$, on a $E = \bigoplus_{k=1}^r \ker(P_k(u))$.

Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique s'écrit $\chi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{\alpha_k}$ avec $\alpha_k \in \mathbb{N}^*$ et les $\lambda_k \in \mathbb{K}$ deux à deux distincts. Le polynôme minimal de u s'écrit alors $\pi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_k \leq \alpha_k$.

Définition 1.7. Avec les notations qui précèdent, on appelle sous-espaces caractéristiques de u les sous espaces vectoriels $N_k = \ker(u - \lambda_k Id)^{\alpha_k}$ pour tout $k \in \{1, 2, \dots, r\}$.

Théorème 1.12.

Avec les notations qui précèdent, on a $E = \bigoplus_{k=1}^r N_k$ et pour tout k compris entre 1 et r , on a :

1. $N_k = \ker(u - \lambda_k Id)^{\beta_k}$;
2. λ_k est la seule valeur propre de la restriction de u à N_k ;
3. $\dim(N_k) = \alpha_k$;
4. la restriction de $u - \lambda_k Id$ à N_k est nilpotente d'indice β_k .

Preuve. De $\chi_u(u) = 0$ et du théorème de décomposition des noyaux, on déduit que $E = \bigoplus_{k=1}^r N_k$.

1. On note, pour $k \in \{1, 2, \dots, r\}$, $M_k = \ker(u - \lambda_k Id)^{\beta_k}$ et on a :

$$M_k \subset N_k, \quad E = \bigoplus_{k=1}^r N_k = \bigoplus_{k=1}^r M_k$$

donc $1 \leq \dim(M_k) \leq \dim(N_k)$, $n = \sum_{k=1}^r \dim(N_k) = \sum_{k=1}^r \dim(M_k)$ et nécessairement $M_k = N_k$.

2. Soit λ une valeur propre de $u|_{N_k}$. C'est aussi une valeur propre de u , donc il existe $j \in \{1, 2, \dots, r\}$ tel que $\lambda = \lambda_j$ et un vecteur $x \in N_k \setminus \{0\}$ tel que $(u - \lambda_j Id)(x) = 0$. On a alors $x \in N_k \cap N_j \setminus \{0\}$ et nécessairement $j = k$ ($N_k \cap N_j = \{0\}$ si $j \neq k$).

3. Soit $d_k = \dim(N_k)$, pour k compris entre 1 et r . De ce qui précède on déduit que le polynôme caractéristique de $u|_{N_k} \in \mathcal{L}(N_k)$ (N_k est stable par u) est

$$P_k(X) = (\lambda_k - X)^{d_k}. \text{ De } E = \bigoplus_{k=1}^r N_k, \text{ les } N_k \text{ étant stables par } u, \text{ on déduit}$$

$$\text{que } \chi_u = \prod_{k=1}^r P_k \text{ et } d_k = \alpha_k.$$

4. Pour tout k compris entre 1 et r , on note $v_k = (u - \lambda_k Id)|_{N_k} \in \mathcal{L}(N_k)$ et on a $v_k^{\beta_k} = 0$ ($N_k = \ker(u - \lambda_k Id)^{\beta_k}$). Si $v_k^{\beta_k-1} = 0$ pour un $k \in \{1, 2, \dots, r\}$, le polynôme $(X - \lambda_k)^{\beta_k-1} \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j)^{\beta_j}$ annule alors u puisqu'il annule tous les

$u|_{N_j}$ et $E = \bigoplus_{k=1}^r N_k$, ce qui contredit le fait que $\prod_{k=1}^r (X - \lambda_k)^{\beta_k}$ est le polynôme minimal de u . On a donc $v_k^{\beta_k-1} \neq 0$ et v_k est nilpotent d'indice β_k .

□

Du théorème précédent on va déduire que tout endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$ se décompose de manière unique comme somme d'un endomorphisme diagonalisable (un endomorphisme est diagonalisable s'il existe une base de E dans laquelle sa matrice est diagonale) d et d'un endomorphisme nilpotent v avec d et v qui commutent.

Lemme 1.6 Avec les notations qui précèdent, pour tout k compris entre 1 et r , le projecteur π_k de E sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^r N_j$ est un polynôme en u .

Preuve. Les polynômes $P_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j)^{\beta_j}$, ($1 \leq k \leq r$) étant premiers entre eux dans leur ensemble, on peut trouver des polynômes $Q_1, Q_2, \dots, Q_r$ tels que $\sum_{k=1}^r Q_k P_k = 1$ (théorème de Bézout). On a alors $Id = \sum_{k=1}^r Q_k(u) \circ P_k(u)$ et

tout $x \in E$ s'écrit $x = \sum_{k=1}^r (Q_k(u) \circ P_k(u))(x)$ avec :

$$(Q_k(u) \circ P_k(u))(x) \in \text{Ker}(u - \lambda_k \text{Id})^{\beta_k} = N_k$$

C'est-à-dire que $(Q_k(u) \circ P_k(u))(x)$ est la projection de x sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^r N_j$. On a donc $\pi_k = Q_k(u) \circ P_k(u) \in \mathbb{K}[u]$. $\square$

Les projecteurs π_k de E sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^r N_j$ sont les projecteurs spectraux de u .

Théorème 1.13. Dunford

Soit u un endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Il existe un unique couple (d, v) d'endomorphismes de E tel que d soit diagonalisable, v soit nilpotent, d et v commutent et $u = d + v$.

Preuve. On a vu que sur chaque N_k , l'endomorphisme $v_k = (u - \lambda_k \text{Id})|_{N_k}$ est nilpotent et en notant $d_k = \lambda_k \text{Id}|_{N_k}$, on a $u|_{N_k} = d_k + v_k$ avec d_k diagonalisable, v_k nilpotent et $d_k v_k = v_k d_k$. On définit alors les endomorphismes d et v par $d = \sum_{k=1}^r \lambda_k \pi_k$ et $v = u - d$. L'endomorphisme d est diagonalisable ($d = d_k$ sur N_k), l'endomorphisme v est nilpotent ($v = v_k$ sur N_k), d et v commutent puisqu'ils sont dans l'algèbre commutative $\mathbb{K}[u]$ et $u = d + v$. Il reste à montrer l'unicité d'un tel couple (d, v) . Soit (d', v') un autre couple d'endomorphismes vérifiant les mêmes conditions que (d, v) . Comme $u = d' + v'$ et d' et v' commutent, ils commutent avec u donc avec d et v qui sont des polynômes en u . On a alors $d - d' = v' - v$, avec $d - d'$ diagonalisable comme somme de deux endomorphismes diagonalisables qui commutent (exercice 2.3) et $v' - v$ nilpotent comme somme de deux endomorphismes nilpotents qui commutent. On a donc nécessairement, $d - d' = v' - v = 0$. D'où l'unicité de la décomposition. $\square$

Pratiquement la décomposition de Dunford d'un endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$ passe par le calcul des projecteurs spectraux π_k . Pour ce faire il suffit de disposer d'un polynôme annulateur de u

de la forme $P(X) = \prod_{k=1}^r (X - \lambda_k)^{m_k}$, avec $m_k \geq \beta_k$ pour tout k compris entre 1

et p où les λ_k sont les valeurs propres de u et β_k est la multiplicité de λ_k comme racine du polynôme minimal (on peut prendre pour polynôme P le polynôme caractéristique, ou mieux le polynôme minimal, de u). À partir de la décomposition

en éléments simples $\frac{1}{P(X)} = \sum_{k=1}^p \sum_{i=1}^{m_k} \frac{\alpha_{ik}}{(X - \lambda_k)^i}$, en posant pour k compris entre

1 et r , $Q_k(X) = \sum_{i=1}^{m_k} \alpha_{ik} (X - \lambda_k)^{m_k-i}$, $P_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j)^{m_j}$, on obtient

$\frac{1}{P(X)} = \sum_{k=1}^r \frac{Q_k(X)}{(X - \lambda_k)^{m_k}}$ et la décomposition de Bézout $1 = \sum_{k=1}^r Q_k P_k$ qui permet d'obtenir les projecteurs spectraux $\pi_k = (Q_k P_k)(u)$. On a alors $u = d + v$ avec $d = \sum_{k=1}^r \lambda_k \pi_k$ et $v = u - d$.

La décomposition de Dunford d'un endomorphisme u de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$ permet le calcul de ses puissances successives. En effet comme d et v commutent, on peut utiliser la formule du binôme de Newton pour écrire :

$$\forall p \geq 1, u^p = (d + v)^p = \sum_{k=0}^p \binom{p}{k} d^k \circ v^{p-k}$$

Le calcul des puissances successives de l'endomorphisme d peut se faire dans une base de diagonalisation ou en utilisant les propriétés des projecteurs pour écrire que :

$$\forall p \geq 1, d^p = \left(\sum_{k=1}^r \lambda_k \pi_k \right)^p = \sum_{k=1}^r \lambda_k^p \pi_k$$

et le calcul des puissances successives de l'endomorphisme nilpotent v s'arrête à v^{q-1} où q est son indice de nilpotence.

On peut aussi calculer v^p avec :

$$\forall p \geq 1, v^p = \left(\sum_{k=1}^r (u - \lambda_k Id) \pi_k \right)^p = \sum_{k=1}^r (u - \lambda_k Id)^p \pi_k$$

Avec l'exercice 1.12 on a un exemple de tels calculs.

1.7 Exercices

On désigne par E un $\mathbb{K}$ -espace vectoriel de dimension $n \geq 1$.

Exercice 1.1. Soit P une matrice inversible de $\mathcal{M}_n(\mathbb{C})$. On définit les matrices réelles R et J par $R = \operatorname{Re}(P)$ et $J = \operatorname{Im}(P)$. Montrer qu'il existe un réel λ tel que la matrice $R + \lambda J$ soit inversible. En déduire que si A, B appartenant à $\mathcal{M}_n(\mathbb{R})$ sont semblables dans $\mathcal{M}_n(\mathbb{C})$, elles sont alors semblables dans $\mathcal{M}_n(\mathbb{R})$.

Solution. Si le polynôme $\varphi(X) = \det(R + XJ)$ s'annule pour toute valeur réelle il est alors identiquement nul et $\varphi(i) = \det(R + iJ) = \det(P) = 0$, ce qui contredit P inversible. Il existe donc des réels λ tels que $\varphi(\lambda) = \det(R + \lambda J) \neq 0$. Si $A, B \in \mathcal{M}_n(\mathbb{R})$ sont semblables dans $\mathcal{M}_n(\mathbb{C})$ il existe alors une matrice P inversible dans $\mathcal{M}_n(\mathbb{C})$ telle que $A = P^{-1}BP$. On a alors en notant R la partie réelle de P et J sa partie imaginaire, $(R + iJ)A = B(R + iJ)$ et en identifiant parties réelles et parties imaginaires, $RA = BR$, $JA = BJ$. Pour tout réel λ tel que $R + \lambda J$ soit inversible on a alors $(R + \lambda J)A = B(R + \lambda J)$, ce qui prouve que les matrices A et B sont donc semblables dans $\mathcal{M}_n(\mathbb{R})$.

Exercice 1.2. On suppose que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Montrer que pour tout endomorphisme u de E il existe un sous espace vectoriel de E de dimension 1 ou 2 stable par u .

Solution. Le polynôme minimal π_u se décompose dans l'anneau factoriel $\mathbb{K}[X]$ en produit de facteurs irréductibles, ces facteurs étant de degré 1 ou 2. Si π_u admet un facteur irréductible de degré 1, ce dernier est de la forme $X - \lambda$ avec λ dans $\mathbb{K}$ et λ est une valeur propre de u . Il est alors facile de vérifier que pour tout vecteur propre x dans $E \setminus \{0\}$ associé à cette valeur propre la droite vectorielle $D = \mathbb{K}x$ est stable par u . Si $\mathbb{K} = \mathbb{R}$ et tous les facteurs irréductibles de π_u sont de degré 2, ils sont alors de la forme $X^2 + bX + c$ et le polynôme π_u s'écrit $\pi_u(X) = (X^2 + bX + c)Q(X)$. De l'égalité $0 = \pi_u(u) = (u^2 + bu + cId) \circ Q(u)$ et du caractère minimal de π_u on déduit que l'endomorphisme $u^2 + bu + cId$ n'est pas injectif, c'est-à-dire que son noyau n'est pas réduit à $\{0\}$. Pour tout vecteur x non nul dans ce noyau on vérifie alors que $F = \text{Vect}\{x, u(x)\}$ est un sous espace vectoriel de dimension 2 stable par u . En effet on sait déjà que F est de dimension 1 ou 2 et si F est de dimension 1, x est alors vecteur propre de u et la valeur propre correspondante est racine de π_u ce qui contredit le fait que π_u n'a pas de facteur irréductible de degré 1. Donc F est de dimension 2. Avec $u^2(x) + bu(x) + cx = 0$, on déduit que $u^2(x)$ est dans F , ce qui entraîne que F est stable par u .

Exercice 1.3. Soient A et B dans $\mathcal{M}_n(\mathbb{K})$. En remarquant que :

$$\begin{pmatrix} XI_n - AB & 0 \\ B & XI_n \end{pmatrix} = \begin{pmatrix} -A & XI_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} B & XI_n \\ I_n & A \end{pmatrix}$$

dans $\mathcal{M}_n(\mathbb{K}(X))$, montrer que AB et BA ont même polynôme caractéristique.

Solution. On a :

$$\begin{aligned} X^n \det(XI_n - AB) &= \det \begin{pmatrix} XI_n - AB & 0 \\ B & XI_n \end{pmatrix} \\ &= \det \begin{pmatrix} -A & XI_n \\ I_n & 0 \end{pmatrix} \det \begin{pmatrix} B & XI_n \\ I_n & A \end{pmatrix} \\ &= \det \begin{pmatrix} B & XI_n \\ I_n & A \end{pmatrix} \det \begin{pmatrix} -A & XI_n \\ I_n & 0 \end{pmatrix} \\ &= \det \begin{pmatrix} XI_n - BA & XB \\ 0 & XI_n \end{pmatrix} = X^n \det(XI_n - BA) \end{aligned}$$

donc AB et BA ont même polynôme caractéristique.

Exercice 1.4. Montrer qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est de rang r si, et seulement si, elle est équivalente à $A_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. En déduire que

pour toutes matrices A et B dans $\mathcal{M}_n(\mathbb{K})$, AB et BA ont même polynôme caractéristique (considérer d'abord le cas où $A = A_r$).

Solution.

- Soient $u \in \mathcal{L}(\mathbb{K}^n)$ de matrice A , de rang r , dans la base canonique de $\mathbb{K}^n$, H un supplémentaire de $\ker(u)$ dans $\mathbb{K}^n$, donc de dimension r , $\mathcal{B}_1 = (e_k)_{1 \leq k \leq r}$ une base de H et $\mathcal{B}_2$ une base de $\ker(u)$. La famille $u(\mathcal{B}_1) = (u(e_k))_{1 \leq k \leq r}$ est alors libre dans $\mathbb{K}^n$ (si $\sum_{k=1}^r \lambda_k u(e_k) = 0$, on a alors $\sum_{k=1}^r \lambda_k e_k \in H \cap \ker(u) = \{0\}$ et tous les λ_k sont nuls du fait que $\mathcal{B}_1$ est libre) et elle se complète en une base $\mathcal{B} = (u(e_1), \dots, u(e_r), f_{r+1}, \dots, f_n)$ de $\mathbb{K}^n$. La matrice de u dans les bases $\mathcal{B}_1 \cup \mathcal{B}_2$ et $\mathcal{B}$ a alors la forme indiquée. La réciproque est évidente.
- Pour $A = A_r$, en écrivant $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ avec B_1 dans $M_r(\mathbb{K})$ et B_4 dans $\mathcal{M}_{n-r}(\mathbb{K})$ on a $AB = \begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix}$, $BA = \begin{pmatrix} B_1 & 0 \\ B_3 & 0 \end{pmatrix}$ et :

$$\det(XI_n - AB) = X^{n-r} \det(XI_r - B_1) = \det(XI_n - BA)$$

Dans le cas général, on a $A = PA_rQ$ (le cas $A = 0$ est trivial) et :

$$\chi_{AB} = \chi_{P(A_rQB)P^{-1}} = \chi_{A_r(QBP)} = \chi_{(QBP)A_r} = \chi_{Q(BPA_rQ)Q^{-1}} = \chi_{BA}$$

Exercice 1.5. On suppose que $\mathbb{K}$ est infini. Montrer que pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$ il existe une infinité de scalaires λ tels que $\lambda I_n - A$ soit inversible. En utilisant ce résultat, montrer que pour toutes matrices A et B dans $\mathcal{M}_n(\mathbb{K})$, AB et BA ont même polynôme caractéristique (considérer d'abord le cas où la matrice A est inversible).

Solution.

- L'application polynomiale $\lambda \mapsto \det(\lambda I_n - A)$ a au plus n racines dans $\mathbb{K}$, donc ce déterminant est non nul pour une infinité de valeurs de λ et pour ces valeurs la matrice $\lambda I_n - A$ est inversible.
- Si A est inversible, les matrices AB et $BA = A^{-1}(AB)A$ sont alors semblables et elles ont même polynôme caractéristique.
- Pour A dans $\mathcal{M}_n(\mathbb{K})$ il existe une infinité de scalaires λ tels que $\lambda I_n - A$ soit inversible et pour ces valeurs de λ on a, $\chi_{(A-\lambda I_n)B}(X) = \chi_{B(A-\lambda I_n)}(X)$. Pour tout x fixé dans $\mathbb{K}$ on a donc deux polynômes en λ qui coïncident pour une infinité de valeurs, ils sont donc égaux et $\lambda = 0$ donne $\chi_{AB}(X) = \chi_{BA}(X)$.

Exercice 1.6. On suppose que $\mathbb{K} = \mathbb{C}$ et on se donne u, v dans $\mathcal{L}(E)$. Montrer que :

1. si u et v commutent, ils ont alors un vecteur propre en commun ;
2. si $u \circ v - v \circ u = u$, u et v ont alors un vecteur propre en commun ;
3. si $u \circ v - v \circ u = \alpha u + \beta v$, u et v ont alors un vecteur propre en commun.

Solution.

1. Avec la condition $u \circ v = v \circ u$ on vérifie facilement, pour toute valeur propre $\lambda \in \mathbb{C}$ de u , l'espace propre $E(u, \lambda)$ est stable par v . La restriction de v à $E(u, \lambda)$ définit alors un endomorphisme de $E(u, \lambda)$ et cet endomorphisme admet des valeurs propres dans $\mathbb{C}$. Tout vecteur propre de la restriction de v à $E(u, \lambda)$ associé à l'une de ces valeurs propres est alors un vecteur propre commun à u et v .
2. Si $u \circ v - v \circ u = u$ on vérifie alors par récurrence que, pour tout $k \in \mathbb{N}^*$, on a $u^k \circ v - v \circ u^k = ku^k$. En effet le résultat est vrai pour $k = 1$ et en le supposant vrai pour $k \geq 1$ on a :

$$\begin{aligned} u^{k+1} \circ v - v \circ u^{k+1} &= u \circ (u^k \circ v - v \circ u^k) + (u \circ v - v \circ u) \circ u^k \\ &= u \circ (ku^k) + u \circ u^k = (k+1)u^k \end{aligned}$$

Ce résultat peut se traduire en disant que l'application linéaire définie sur $\mathcal{L}(E)$ par $\varphi(w) = w \circ v - v \circ w$ est telle que $\varphi(u^k) = ku^k$ pour tout $k \in \mathbb{N}^*$. Si u^k est non nul pour tout $k \in \mathbb{N}^*$, chacun de ces vecteurs de $\mathcal{L}(E)$ est alors un vecteur propre de φ et cet endomorphisme de $\mathcal{L}(E)$ a une infinité de valeurs propres, ce qui est impossible. Il existe donc un entier $k \geq 1$ tel que $u^k = 0$, ce qui entraîne que u n'est pas injectif et 0 est valeur propre de u . Avec la condition $u \circ v - v \circ u = u$ on vérifie alors facilement que le noyau de u est stable par v . Tout vecteur propre de la restriction de v au noyau de u est alors un vecteur propre commun à u et v .

On peut aussi montrer que la condition $u \circ v - v \circ u = u$ entraîne que u n'est pas inversible. En effet si u est inversible on a alors $v = u^{-1} \circ (v + Id) \circ u$, donc v et $v + Id$ ont même polynôme caractéristique, ce qui nous dit que :

$$\forall \lambda \in \mathbb{C}, \chi_v(\lambda) = \det(\lambda Id - v - Id) = \chi_v(\lambda - 1)$$

On en déduit alors par récurrence que $\chi_v(k) = \chi_v(0)$ pour tout entier naturel k , le polynôme χ_v étant non constant, ce qui est impossible. On a donc $\ker(u) \neq \{0\}$ et on conclut comme dans la démonstration qui précède.

3. D'après ce qui précède on peut supposer que α et β sont non nuls et la condition $u \circ v - v \circ u = \alpha u + \beta v$ peut s'écrire :

$$(\alpha u + \beta v) \circ \left(\frac{1}{\alpha} v \right) - \left(\frac{1}{\alpha} v \right) \circ (\alpha u + \beta v) = \alpha u + \beta v$$

Avec ce qui précède, on déduit qu'il existe un vecteur non nul x dans le noyau de $\alpha u + \beta v$ qui soit un vecteur propre de v , c'est-à-dire que $(\alpha u + \beta v)(x) = 0$ et $v(x) = \lambda x$. On a alors $u(x) = -\frac{\beta}{\alpha} \lambda x$ et x est un vecteur propre commun à u et v .

Exercice 1.7. Soient α, β dans $\mathbb{K}$ et $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre $n \geq 3$ définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, \dots, n\} \setminus \{i\} \end{cases}$$

1. Calculer $\Delta(\alpha, \beta) = \det(A(\alpha, \beta))$.
2. Calculer le polynôme caractéristique et les valeurs propres avec leur multiplicité de la matrice $A(\alpha, \beta)$.
3. Calculer le polynôme minimal de la matrice $A(\alpha, \beta)$.
4. Dans le cas où la matrice $A(\alpha, \beta)$ est inversible, calculer son inverse.
5. Calculer $A(\alpha, \beta)^k$ pour tout entier naturel k .

Solution.

1. La matrice $A(\alpha, \beta)$ est de la forme :

$$A(\alpha, \beta) = \begin{pmatrix} \beta & \alpha & \alpha & \cdots & \alpha \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{pmatrix}$$

En ajoutant les lignes 2 à n à la première ligne on a :

$$\Delta(\alpha, \beta) = \det(A(\alpha, \beta)) = \beta + (n-1)\alpha \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{vmatrix}$$

Puis en retranchant la première colonne aux colonnes 2 à n on obtient :

$$\Delta(\alpha, \beta) = (\beta + (n-1)\alpha)(\beta - \alpha)^{n-1}$$

2. Le polynôme caractéristique de $A(\alpha, \beta)$ est donné par :

$$\chi_{(\alpha, \beta)}(X) = \Delta(-\alpha, X - \beta) = (X - (\beta + (n-1)\alpha))(X - (\beta - \alpha))^{n-1}$$

Les valeurs propres de $A(\alpha, \beta)$ sont donc $\lambda_1 = \beta + (n-1)\alpha$ et $\lambda_2 = \beta - \alpha$. Pour $\alpha = 0$, la matrice $A(0, \beta)$ est celle d'une homothétie et $\lambda_1 = \lambda_2 = \beta$ est valeur propre d'ordre n de cette matrice. Pour $\alpha \neq 0$, λ_1 est valeur propre simple et λ_2 est valeur propre d'ordre $n-1$. On peut aussi remarquer que $A(\alpha, \beta) + (\alpha - \beta)I_n = \alpha A(1, 1)$ avec $A(1, 1)$ de rang 1. On déduit donc que 0 est valeur propre d'ordre $n-1$ de $A(1, 1)$ et que $\text{Tr}(A(1, 1)) = n-1$ est valeur propre simple. On retrouve alors facilement les valeurs propres de $A(\alpha, \beta)$.

3. Pour $\alpha = 0$, on a $A(\alpha, \beta) = \beta I_n$ et $\pi_{(\alpha, \beta)}(X) = X - \beta$. Pour α non nul, on a :

$$\begin{aligned} A(\alpha, \beta) - (\beta - \alpha) I_n &= \alpha A(1, 1) \\ A(\alpha, \beta) - (\beta + (n-1)\alpha) I_n &= \alpha A(1, 1-n) \end{aligned}$$

avec $A(1, 1)A(1, 1-n) = 0$. La matrice $A(\alpha, \beta)$ n'étant pas celle d'une homothétie, on en déduit que son polynôme minimal est :

$$\pi_{(\alpha, \beta)}(X) = (X - (\beta - \alpha))(X - (\beta + (n-1)\alpha))$$

4. Pour $\alpha = 0$, on $A(\alpha, \beta) = \beta I_n$ et cette matrice est inversible si, et seulement si, $\beta \neq 0$ avec $A(\alpha, \beta)^{-1} = \beta^{-1} I_n$. Pour α non nul, la matrice $A(\alpha, \beta)$ est inversible si, et seulement si, $\beta \neq \alpha$ et $\beta \neq -(n-1)\alpha$ et son polynôme minimal est $\pi_{(\alpha, \beta)}(X) = (X - \lambda_1)(X - \lambda_2)$. L'égalité $\pi_{(\alpha, \beta)}(A(\alpha, \beta)) = 0$ donne alors :

$$\begin{aligned} A(\alpha, \beta)^{-1} &= \frac{1}{\lambda_1 \lambda_2} ((\lambda_1 + \lambda_2) I_n - A(\alpha, \beta)) \\ &= \frac{1}{\lambda_1 \lambda_2} A(-\alpha, \lambda_1 + \lambda_2 - \beta) = \frac{1}{\lambda_1 \lambda_2} A(-\alpha, \beta + (n-1)\alpha) \end{aligned}$$

5. Pour $\alpha = 0$, on $A(\alpha, \beta) = \beta I_n$ et $A(\alpha, \beta)^k = \beta^k I_n$ pour tout $k \in \mathbb{N}$. On suppose que α est non nul. Pour $k = 0$ et $k = 1$, on a $A(\alpha, \beta)^k = a_k A(\alpha, \beta) + b_k I_n$ avec $(a_0, b_0) = (0, 1)$ et $(a_1, b_1) = (1, 0)$. Pour $k \geq 2$, la division euclidienne de X^k par $\pi_{(\alpha, \beta)}(X)$ s'écrit $X^k = Q(X) \pi_{(\alpha, \beta)}(X) + a_k X + b_k$, avec :

$$\begin{cases} a_k = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} = \frac{\lambda_1^k - \lambda_2^k}{n\alpha} \\ b_k = \frac{\lambda_1 \lambda_2^k - \lambda_2 \lambda_1^k}{\lambda_1 - \lambda_2} = \frac{\lambda_1 \lambda_2}{n\alpha} (\lambda_2^{k-1} - \lambda_1^{k-1}) \end{cases}$$

Et on a $A(\alpha, \beta)^k = a_k A(\alpha, \beta) + b_k I_n$ pour tout $k \in \mathbb{N}$. On peut aussi écrire :

$$A(\alpha, \beta)^k = \frac{1}{n\alpha} (\lambda_1^k (A(\alpha, \beta) - \lambda_2 I_n) + \lambda_2^k (\lambda_1 I_n - A(\alpha, \beta)))$$

$(\frac{1}{nb} (A(\alpha, \beta) - \lambda_2 I_n))$ et $\frac{1}{nb} (\lambda_1 I_n - A(\alpha, \beta))$ sont les projecteurs spectraux, ou encore $A(\alpha, \beta)^k = \frac{1}{n} (\lambda_2^k A(1, 1) - \lambda_1^k A(1, 1-n))$.

Exercice 1.8.

1. Pour toute matrice complexe $A = ((a_{ij}))_{1 \leq i, j \leq n}$ d'ordre $n \geq 2$ et pour tout nombre complexe t , on note $A_t = ((a_{ij} + t))_{1 \leq i, j \leq n}$. Montrer que :

$$\forall t \in \mathbb{C}, \det(A_t) = \det(A) + tS(A)$$

la constante réelle $S(A)$ ne dépendant que des coefficients de A .

2. Soient a, b, c des nombres complexes et $M(a, b, c) = ((m_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre $n \geq 2$ définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} m_{ii} = b \\ m_{ij} = c \text{ si } j \in \{i+1, \dots, n\} \\ m_{ij} = a \text{ si } j \in \{1, \dots, i-1\} \end{cases}$$

- (a) En prenant, pour $a \neq c$, $A = M(a, b, c)$, $t = -a$ et $t = -c$ dans la question précédente, calculer le déterminant de $M(a, b, c)$.
- (b) Calculer le polynôme caractéristique et les valeurs propres de $M(a, b, c)$.

Solution.

1. En notant C_j la colonne numéro j de A , T le vecteur de composantes toutes égales à t et en utilisant la n -linéarité du déterminant on déduit que :

$$\det(A_t) = \det(C_1 + T, \dots, C_n + T) = \det(A) + tS(A)$$

où on a noté $S(A) = \sum_{j=1}^n \det(C_1, \dots, C_{j-1}, E, C_j, \dots, C_n)$, E désignant le vecteur de composantes toutes égales à 1.

2. La matrice $M(a, b, c)$ s'écrit :

$$M(a, b, c) = \begin{pmatrix} b & c & c & \cdots & c \\ a & b & c & \cdots & c \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & \cdots & a & b & c \\ a & \cdots & a & a & b \end{pmatrix}$$

- (a) Pour $a = c$, la matrice $M(a, b, a)$ est symétrique et ce cas a été étudié avec l'exercice 1.7. Dans ce cas, on a :

$$\det(M(a, b, a)) = (b + (n-1)a)(b-a)^{n-1}$$

Pour $a \neq c$, prenant $A = M(a, b, c)$ dans la question précédente, on a :

$$\det(A_{-a}) = (b-a)^n = \det(A) - aS(A)$$

$$\det(A_{-c}) = (b-c)^n = \det(A) - cS(A)$$

On en déduit facilement que $\det(M(a, b, c)) = \frac{c(b-a)^n - a(b-c)^n}{c-a}$.

- (b) Pour $a = c$, le polynôme caractéristique et les valeurs propres de la matrice $M(a, b, a)$ ont été calculés à l'exercice 1.7. On suppose donc que $a \neq c$. Dans ce cas, le polynôme caractéristique de $M(a, b, c)$ est donné par :

$$\chi_{a,b,c}(X) = \frac{c(b-a-X)^n - a(b-c-X)^n}{c-a}$$

Pour $a = 0$ ou $c = 0$, la matrice $M(a, b, c)$ est triangulaire avec b pour valeur propre d'ordre n . Pour $a \neq 0$ et $c \neq 0$, on a

$$\chi_{a,b,c}(b-a) = a(a-c)^{n-1} \neq 0, \quad \chi_{a,b,c}(b-c) = c(c-a)^{n-1} \neq 0$$

On peut donc écrire que les valeurs propres de $M(a, b, c)$ sont définies par $\left(\frac{b-\lambda-a}{b-\lambda-c}\right)^n = \frac{a}{c}$. En notant $\omega_1, \dots, \omega_n$ les racines n -ième dans $\mathbb{C}$ de $\frac{a}{c}$ (qui est non nul), on déduit alors que les valeurs propres de $M(a, b, c)$ sont les $\lambda_k = b + \frac{c\omega_k - a}{1 - \omega_k}$ pour k compris entre 1 et n (comme $a \neq c$, on a $\omega_k \neq 1$).

Exercice 1.9. Soient a, b dans $\mathbb{R}$ et $A(a, b)$ la matrice réelle d'ordre $n \geq 2$ définie par :

$$A(a, b) = \begin{pmatrix} a & b & 0 & \cdots & 0 \\ b & a & b & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b & a & b \\ 0 & \cdots & 0 & b & a \end{pmatrix}$$

Calculer les valeurs propres et les vecteurs propres associés de $A(a, b)$.

Solution. On peut écrire $A(a, b) = aI_n + bA(0, 1)$. Il suffit donc de considérer le cas où $(a, b) = (0, 1)$. La matrice $A(0, 1)$ est symétrique réelle, donc toutes ses valeurs propres sont réelles. Avec le théorème de Gerschgorin-Hadamard, on déduit que pour toute valeur propre $\lambda \in \mathbb{R}$ on a $|\lambda| \leq 2$. Une telle valeur propre peut donc s'écrire $\lambda = 2 \cos(\alpha)$ avec $\alpha \in [0, \pi]$. Si x est un vecteur propre non nul associé ses composantes sont solutions de la récurrence :

$$x_{k-1} - \lambda x_k + x_{k+1} = 0 \quad (1 \leq k \leq n)$$

avec les conditions aux limites $x_0 = 0$ et $x_{n+1} = 0$. Le polynôme caractéristique de cette récurrence est $P(r) = r^2 - 2 \cos(\alpha)r + 1$, soit $P(r) = (r - e^{i\alpha})(r - e^{-i\alpha})$. Les racines sont donc $r_1 = e^{i\alpha}$ et $r_2 = e^{-i\alpha}$. Ce qui donne :

$$x_k = c_1 e^{ik\alpha} + c_2 e^{-ik\alpha} \quad (0 \leq k \leq n+1)$$

Avec $x_0 = 0$, on déduit que $c_2 = -c_1$. Avec $x_{n+1} = 0$ et $(c_1, c_2) \neq (0, 0)$, on déduit que $\sin((n+1)\alpha) = 0$ et $\alpha = \frac{j\pi}{n+1}$ avec $1 \leq j \leq n$. Les valeurs propres de $A(a, b)$ sont donc les :

$$\lambda_k(a, b) = a + 2b \cos\left(\frac{k\pi}{n+1}\right) \quad (1 \leq k \leq n)$$

L'espace propre associé à $\lambda_k(a, b)$ est la droite engendrée par le vecteur $v^{(k)}$ de composantes :

$$v_j^{(k)} = \sin\left(j \frac{k\pi}{n+1}\right) \quad (1 \leq k, j \leq n)$$

Exercice 1.10. Soit u un endomorphisme de E nilpotent d'ordre 2. Montrer que si $\mathcal{B}_1 = (u(e_k))_{1 \leq k \leq r}$ est une base de $\text{Im}(u)$, elle se complète alors en une base $\mathcal{B}_2 = (u(e_1), \dots, u(e_r), e'_1, \dots, e'_p)$ de $\ker(u)$ et la famille de vecteurs $\mathcal{B} = (u(e_1), \dots, u(e_r), e'_1, \dots, e'_p, e_1, \dots, e_r)$ est une base de E . Écrire la matrice de u dans cette base.

Solution. De $u^2 = 0$ on déduit que $\text{Im}(u)$ est contenu dans $\ker(u)$. On peut donc compléter $\mathcal{B}_1$ en une base $\mathcal{B}_2$ de $\ker(u)$. En utilisant le théorème du rang on a $\text{card}(\mathcal{B}_2) = n - r = r + p$, on déduit donc que $p = n - 2r$ et $\text{card}(\mathcal{B}) = n$. Il suffit donc de montrer que le système $\mathcal{B}$ est libre. Si

$$\sum_{k=1}^r \alpha_k e_k + \sum_{k=1}^r \beta_k u(e_k) + \sum_{k=1}^p \gamma_k e'_k = 0$$

en appliquant u on a $\sum_{k=1}^r \alpha_k u(e_k) = 0$ et tous les α_k sont nuls. Le système $\mathcal{B}_2$ étant libre on en déduit ensuite que tous les β_k et les γ_k sont nuls. En conclusion le système $\mathcal{B}$ est libre et c'est une base de E . La matrice de u dans cette base est alors $A = \begin{pmatrix} 0 & 0 & I_r \\ 0 & 0 & 0 \end{pmatrix}$.

Exercice 1.11. On suppose $\mathbb{K}$ de caractéristique nulle. En utilisant les formules de Newton, montrer que pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$ les matrices AB et BA ont même polynôme caractéristique.

Solution. Notant $\chi_{AB}(X) = \sum_{k=0}^n a_k X^k$, $\chi_{BA}(X) = \sum_{k=0}^n b_k X^k$, on a $a_n = b_n = 1$ et $a_{n-1} = -\text{Tr}(AB) = -\text{Tr}(BA) = b_{n-1}$. Supposons que $a_{n-j} = b_{n-j}$ pour tout entier j compris entre 0 et $k-1$ avec $1 \leq k \leq n$. On a alors, en utilisant les formules de Newton relatives à la matrice AB :

$$ka_{n-k} = -\left(\text{Tr}\left((AB)^k\right) + a_{n-1} \text{Tr}\left((AB)^{k-1}\right) + \dots + a_{n-(k-1)} \text{Tr}(AB)\right)$$

avec, pour tout entier j compris entre 1 et k :

$$\text{Tr}\left((AB)^j\right) = \text{Tr}\left(A\left(B(AB)^{j-1}\right)\right) = \text{Tr}\left(\left(B(AB)^{j-1}\right)A\right) = \text{Tr}\left((BA)^j\right)$$

ce qui nous donne $ka_{n-k} = kb_{n-k}$, soit $a_{n-k} = b_{n-k}$ puisque $\mathbb{K}$ est de caractéristique nulle. On a donc ainsi montré par récurrence que AB et BA ont même polynôme caractéristique.

Exercice 1.12. *Ecrire la décomposition de Dunford de la matrice :*

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{C})$$

En déduire un calcul de A^r pour tout entier r strictement positif.

Solution. Le polynôme caractéristique de A est $\chi_A(X) = X(X-1)^3$ et son polynôme minimal est $\pi_A(X) = X(X-1)^2$. On a la décomposition en éléments simples :

$$\frac{1}{\pi_A(X)} = \frac{1}{X} + \left(\frac{1}{(X-1)^2} - \frac{1}{X-1} \right)$$

qui donne la décomposition de Bézout $1 = (X-1)^2 + (2X - X^2)$ et les projecteurs spectraux $\pi_1 = (A - I_4)^2$, $\pi_2 = 2A - A^2$ (on a $\pi_1 + \pi_2 = I_4$). On obtient alors la décomposition $A = D + V$ avec :

$$D = \pi_2 = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad V = A - D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Pour $r > 0$, on a $A^r = D^r + rD^{r-1}V$ ($V^2 = 0$) avec $D^r = \pi_2^r = \pi_2 = D$. Soit :

$$\forall r \geq 2, \quad A^r = D(I_4 + rV) = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & r & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Chapitre 2

Réduction des endomorphismes et des matrices

On garde les notations du chapitre précédent.

2.1 Trigonalisation

Définition 2.1. On dit qu'un endomorphisme $u \in \mathcal{L}(E)$ est trigonalisable s'il existe une base de E dans laquelle la matrice de u est triangulaire.

Si $u \in \mathcal{L}(E)$ a pour matrice A dans une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ et si de plus u est trigonalisable, il existe alors une base $\mathcal{B}' = (e'_k)_{1 \leq k \leq n}$ dans laquelle la matrice T de u est triangulaire. On sait alors que les matrices A et T sont semblables, c'est-à-dire qu'il existe une matrice inversible P d'ordre n à coefficients dans $\mathbb{K}$ telle que $T = P^{-1}AP$. Ce qui nous conduit à la définition suivante.

Définition 2.2. On dit qu'une matrice A d'ordre n à coefficients dans $\mathbb{K}$ est trigonalisable, si elle est semblable à une matrice triangulaire.

Deux matrices semblables ayant même déterminant, elles ont même polynôme caractéristique. En conséquence, si la matrice A est semblable à une matrice triangulaire T , les termes diagonaux de T sont alors les valeurs propres de A .

On en déduit également qu'une matrice à coefficients réels n'ayant pas toutes ses valeurs propres réelles n'est pas trigonalisable dans $\mathcal{M}_n(\mathbb{R})$. Par exemple la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ de valeurs propres complexes i et $-i$ n'est pas trigonalisable dans $\mathcal{M}_2(\mathbb{R})$.

Sur un corps algébriquement clos, tout endomorphisme et toute matrice sont trigonalisables.

Lemme 2.1 *On suppose que $\mathbb{K}$ est algébriquement clos et que E est de dimension $n \geq 2$. Si u est un endomorphisme de E , il existe alors un hyperplan de E stable par u .*

Preuve. $\mathbb{K}$ étant algébriquement clos, u admet au moins une valeur propre $\lambda \in \mathbb{K}$. L'endomorphisme $u - \lambda Id$ est alors non injectif et son image $\text{Im}(u - \lambda Id)$ est un sous espace vectoriel de dimension inférieure ou égale à $n - 1$. Il existe donc un hyperplan H de E qui contient $\text{Im}(u - \lambda Id)$. On a alors $u(x) - \lambda x \in H$ pour tout $x \in E$ et :

$$\forall x \in H, u(x) = (u(x) - \lambda x) + \lambda x \in H$$

ce qui signifie que H est stable par u . □

Théorème 2.1.

Pour $\mathbb{K}$ algébriquement clos, tout endomorphisme de E est trigonalisable.

Preuve. On raisonne par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$ le résultat est évident. Supposons-le acquis pour les $\mathbb{K}$ -espaces vectoriels de dimension $n - 1$. Si u est un endomorphisme sur le $\mathbb{K}$ -espace vectoriel E de dimension $n \geq 2$, il existe alors un hyperplan H de E stable par u et la restriction v de u à H est trigonalisable, c'est-à-dire qu'il existe une base $\mathcal{B}_1 = (e_k)_{1 \leq k \leq n-1}$ de H dans laquelle la matrice de v est triangulaire supérieure. Pour tout vecteur $e_n \in E \setminus H$ la famille $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ est une base de E et la matrice de u dans cette base est triangulaire supérieure. □

Corollaire 2.1 : *Toute matrice à coefficients dans un corps algébriquement clos est trigonalisable.*

On a déjà vu que le résultat du théorème 2.1 est faux si le corps de base est $\mathbb{R}$. Ce qui est important dans la démonstration précédente c'est que u admet des valeurs propres ainsi que la restriction de u à tout sous espace vectoriel de E stable par u . Pour tout corps commutatif, on a le résultat suivant.

Théorème 2.2.

Un endomorphisme $u \in \mathcal{L}(E)$ est trigonalisable sur $\mathbb{K}$ si, et seulement si, son polynôme caractéristique est scindé sur $\mathbb{K}$.

Preuve. Il est clair que si u est trigonalisable son polynôme caractéristique est alors scindé sur $\mathbb{K}$ (raisonner avec la matrice de u dans une base de trigonalisation). Réciproquement si χ_u est scindé sur $\mathbb{K}$, u admet alors au moins une valeur propre $\lambda \in \mathbb{K}$ et comme dans le cas algébriquement clos, on montre qu'il existe un hyperplan H de E qui est stable par u . En utilisant le lemme 1.1, on déduit que le polynôme caractéristique de la restriction de u à H est également scindé sur $\mathbb{K}$. La démonstration se termine alors par récurrence comme dans le cas algébriquement clos. □

Corollaire 2.2 : Soit u un endomorphisme de E . Si u est trigonalisable et si F est un sous espace vectoriel de E stable par u , la restriction de u à F est alors aussi trigonalisable.

Corollaire 2.3 : Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé sur $\mathbb{K}$ est trigonalisable.

2.2 Diagonalisation

Définition 2.3. On dit qu'un endomorphisme $u \in \mathcal{L}(E)$ est diagonalisable s'il existe une base de E dans laquelle la matrice de u est diagonale.

Si $u \in \mathcal{L}(E)$ a pour matrice A dans une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ et si de plus u est diagonalisable il existe alors une base $\mathcal{B}' = (e'_k)_{1 \leq k \leq n}$ dans laquelle la matrice D de u est diagonale. On sait alors que les matrices A et D sont semblables, c'est-à-dire qu'il existe une matrice inversible P d'ordre n à coefficients dans $\mathbb{K}$ telle que $D = P^{-1}AP$. Ce qui nous conduit à la définition suivante.

Définition 2.4. On dit qu'une matrice A d'ordre n à coefficients dans $\mathbb{K}$ est diagonalisable si elle est semblable à une matrice diagonale.

Une condition suffisante de diagonalisation est donnée par le résultat suivant.

Théorème 2.3.

Si u est un endomorphisme de E ayant n valeurs propres distinctes dans $\mathbb{K}$, il est alors diagonalisable.

Preuve. Si u a n valeurs propres distinctes, ces valeurs propres sont alors toutes simples et chaque sous espace propre de u est de dimension 1. Il est facile de vérifier que ces sous espaces propres sont en somme directe. Il en résulte alors que u est diagonalisable. $\square$

Des conditions nécessaires et suffisantes pour qu'un endomorphisme soit diagonalisable sont données par le résultat suivant.

Théorème 2.4.

Soit u un endomorphisme de E . Les conditions suivantes sont équivalentes :

1. u est diagonalisable ;
2. si $\lambda_1, \dots, \lambda_p$ sont les valeurs propres de u deux à deux distinctes dans

$$\mathbb{K}, \text{ on a alors } E = \bigoplus_{k=1}^p \ker(u - \lambda_k \text{Id}) ;$$

3. si $\lambda_1, \dots, \lambda_p$ sont les valeurs propres de u deux à deux distinctes dans $\mathbb{K}$, on a alors $\sum_{k=1}^p \dim(\ker(u - \lambda_k Id)) = n$;
4. le polynôme caractéristique de u est scindé sur $\mathbb{K}$ de racines deux à deux distinctes $\lambda_1, \dots, \lambda_p$ dans $\mathbb{K}$, chaque λ_k ($1 \leq k \leq p$) étant de multiplicité $\alpha_k = \dim(\ker(u - \lambda_k Id))$;
5. il existe un polynôme annulateur de u scindé à racines simples dans $\mathbb{K}$;
6. le polynôme minimal π_u est scindé à racines simples dans $\mathbb{K}$.

Preuve. (1) $\Leftrightarrow$ (2). Si u est diagonalisable, il existe alors une base :

$$\mathcal{B} = (e_{11}, \dots, e_{1, \alpha_1}, \dots, e_{p1}, \dots, e_{p, \alpha_p})$$

de E formée de vecteurs propres avec :

$$u(e_{k,j}) = \lambda_k e_{k,j} \quad (1 \leq k \leq p, 1 \leq j \leq \alpha_k)$$

où les $\lambda_1, \dots, \lambda_p$ sont les valeurs propres distinctes deux à deux de u . On déduit alors facilement que $E = \bigoplus_{k=1}^p \ker(u - \lambda_k Id)$ avec $\dim(\ker(u - \lambda_k Id)) = \alpha_k$. La réciproque est évidente.

(2) $\Leftrightarrow$ (3) Cette équivalence est évidente (les espaces propres sont en somme directe).

$$(3) \Rightarrow (4) \text{ Si } \sum_{k=1}^p \dim(\ker(u - \lambda_k Id)) = n, \text{ on a alors } E = \bigoplus_{k=1}^p \ker(u - \lambda_k Id) \text{ et}$$

le polynôme caractéristique de u s'écrit $\chi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$.

(4) $\Rightarrow$ (5) Du fait que $\dim(\ker(u - \lambda_k Id))$ est égal à la multiplicité de λ_k , on déduit que $\sum_{k=1}^p \dim(\ker(u - \lambda_k Id)) = n$ et donc que $E = \bigoplus_{k=1}^p \ker(u - \lambda_k Id)$. On

a alors $P(u) = 0$ avec $P(X) = \prod_{k=1}^p (X - \lambda_k)$ scindé sur $\mathbb{K}$ à racines simples.

(5) $\Rightarrow$ (6) Le polynôme minimal étant un diviseur de tout polynôme annulateur, il est également scindé sur $\mathbb{K}$ à racines simples si (5) est vérifié.

(6) $\Rightarrow$ (1) Si le polynôme minimal de u s'écrit $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)$, les λ_k étant deux à deux distincts, le théorème de décomposition des noyaux nous donne $E = \bigoplus_{k=1}^p \ker(u - \lambda_k Id)$, ce qui signifie que u est diagonalisable. $\square$

2.3 Espaces vectoriels euclidiens

Pour ce paragraphe, E est un espace vectoriel sur $\mathbb{R}$ de dimension $n \geq 1$.

Définition 2.5. On appelle produit scalaire euclidien sur E toute application :

$$\begin{aligned}\varphi : E \times E &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \varphi(x, y)\end{aligned}$$

qui vérifie les propriétés suivantes :

- pour tout x dans E , l'application $y \mapsto \varphi(x, y)$ est linéaire et pour tout y dans E , l'application $x \mapsto \varphi(x, y)$ est linéaire (φ est une forme bilinéaire sur E) ;
- $\varphi(x, y) = \varphi(y, x)$ pour tous x, y dans E (φ est symétrique) ;
- $\varphi(x, x) \geq 0$ pour tout x dans E (φ est positive) ;
- pour x dans E , $\varphi(x, x) = 0$ équivaut à $x = 0$ (φ est définie).

On note en général $(x, y) \mapsto \langle x | y \rangle$ un tel produit scalaire.

Définition 2.6. Un espace euclidien est un $\mathbb{R}$ -espace vectoriel de dimension finie muni d'un produit scalaire euclidien.

Exemple 2.1 L'espace vectoriel $\mathbb{R}^n$ est muni de sa base canonique $(e_k)_{1 \leq k \leq n}$.

On se donne un vecteur ω dans $\mathbb{R}^n$ et pour $x = \sum_{k=1}^n x_k e_k$, $y = \sum_{k=1}^n y_k e_k$ dans $\mathbb{R}^n$

on note $\langle x | y \rangle = \sum_{k=1}^n \omega_k x_k y_k$. L'application ainsi définie est un produit scalaire sur $\mathbb{R}^n$ si, et seulement si, toutes les composantes de ω sont strictement positives. Si tous les coefficients ω_k sont égaux à 1, le produit scalaire obtenu est appelé produit scalaire canonique de $\mathbb{R}^n$.

Dans la suite de ce paragraphe E désigne un espace euclidien et on note pour tout x dans E , $\|x\| = \sqrt{\langle x | x \rangle}$.

Théorème 2.5. Inégalité de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Preuve. Si $x = 0$ ou $x \neq 0$ et $y = \lambda x$ avec $\lambda \in \mathbb{R}$, on a alors l'égalité pour tout $y \in E$. On suppose donc que x est non nul et y non lié à x . La fonction polynomiale P définie sur $\mathbb{R}$ par $P(t) = \|y + tx\|^2 = \|x\|^2 t^2 + 2\langle x | y \rangle t + \|y\|^2$ est alors à valeurs strictement positives, le coefficient de t^2 étant non nul, il en résulte que son discriminant est strictement négatif, soit $\langle x | y \rangle^2 - \|x\|^2 \|y\|^2 < 0$, ce qui équivaut à $|\langle x | y \rangle| < \|x\| \|y\|$. $\square$

Théorème 2.6. Inégalité de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = \lambda y$ avec $\lambda \geq 0$ (on dit que x et y sont positivement liés).

Preuve. Si $x = 0$, on a alors l'égalité pour tout $y \in E$. Si $x \neq 0$ et $y = \lambda x$ avec $\lambda \in \mathbb{R}$, on a $\|x + y\| = |1 + \lambda| \|x\| \leq (1 + |\lambda|) \|x\| = \|x\| + \|y\|$, l'égalité étant réalisée pour $\lambda \geq 0$. Pour $\lambda < 0$, l'inégalité est stricte puisque dans ce cas $|1 + \lambda| < 1 + |\lambda| = 1 - \lambda$. On suppose que x est non nul et y non lié à x . On a :

$$\|x + y\|^2 = \|x\|^2 + 2 \langle x | y \rangle + \|y\|^2$$

et avec l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 < \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

ce qui équivaut à $\|x + y\| < \|x\| + \|y\|$. □

Corollaire 2.4 : *L'application $x \mapsto \|x\| = \sqrt{\langle x | x \rangle}$ définit une norme sur E .*

Les deux égalités qui suivent sont utiles en pratique.

Théorème 2.7.

Pour tous x, y dans E on a :

$$\langle x | y \rangle = \frac{1}{2} \left(\|x + y\|^2 - \|x\|^2 - \|y\|^2 \right) = \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right)$$

$$\|x + y\|^2 + \|x - y\|^2 = 2 \left(\|x\|^2 + \|y\|^2 \right)$$

La deuxième identité est l'égalité du parallélogramme. Elle est caractéristique des normes réelles déduites d'un produit scalaire (exercice 2.4).

Pour x et y non nuls, on a $-1 \leq \frac{\langle x | y \rangle}{\|x\| \|y\|} \leq 1$, donc il existe un unique réel θ dans $[0, \pi]$ tel que $\langle x | y \rangle = \cos(\theta) \|x\| \|y\|$. On dit que θ est la mesure dans $[0, \pi]$ de l'angle que font les vecteurs x et y dans $E \setminus \{0\}$. Pour $\theta = \frac{\pi}{2}$, on a $\langle x | y \rangle = 0$ et on dit que les vecteurs x et y sont orthogonaux.

Définition 2.7. *On dit que deux vecteurs x et y dans E sont orthogonaux si $\langle x | y \rangle = 0$.*

Le résultat suivant se démontre facilement.

Théorème 2.8. Pythagore

Les vecteurs x et y sont orthogonaux dans E si, et seulement si, $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Définition 2.8. On appelle famille orthogonale dans E toute famille $(e_i)_{i \in I}$ de vecteurs de E telle que $\langle e_i | e_j \rangle = 0$ pour tous i, j dans I tels que $i \neq j$. Si de plus $\|e_i\| = 1$ pour tout $i \in I$, on dit alors que cette famille est orthonormée ou orthonormale.

Une famille orthogonale de vecteurs non nuls de E est libre. En effet si $(e_i)_{i \in I}$ est une telle famille et si $\sum_{j \in J} \lambda_j e_j = 0$, où J est une partie finie de I , on a alors

pour tout $k \in J$, $0 = \left\langle \sum_{j \in J} \lambda_j e_j \mid e_k \right\rangle = \lambda_k \|e_k\|^2$ avec $\|e_k\| \neq 0$ et nécessairement,

$\lambda_k = 0$.

Théorème 2.9. procédé d'orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_k)_{1 \leq k \leq p}$ dans E , il existe une unique famille orthonormée $(e_k)_{1 \leq k \leq p}$ dans E telle que :

$$\forall k \in \{1, 2, \dots, p\}, \begin{cases} \text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\} \\ \langle x_k \mid e_k \rangle > 0 \end{cases}$$

Preuve. On procède par récurrence sur $p \geq 1$. Pour $p = 1$, on a nécessairement $e_1 = \lambda_1 x_1$ avec $\lambda_1 \in \mathbb{R}^*$ et $1 = \|e_1\| = |\lambda_1| \|x_1\|$, donc $|\lambda_1| = \frac{1}{\|x_1\|}$ ce qui donne deux solutions pour λ_1 . La condition supplémentaire $\langle x_1 \mid e_1 \rangle > 0$ entraîne $\lambda_1 > 0$ et on a l'unique solution $e_1 = \frac{1}{\|x_1\|} x_1$. Supposons $p \geq 2$ et construite la famille orthonormée $(e_k)_{1 \leq k \leq p-1}$ vérifiant les conditions :

$$\forall k \in \{1, 2, \dots, p-1\}, \begin{cases} \text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\} \\ \langle x_k \mid e_k \rangle > 0 \end{cases}$$

Si $(e'_1, e'_2, \dots, e'_{p-1}, e_p)$ est une solution à notre problème, on a alors nécessairement $e'_k = e_k$ pour tout k compris entre 1 et $p-1$ (unicité pour le cas $p-1$). Les conditions $\text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\}$ pour tout k compris entre 1 et p , entraînent $e_p = \sum_{j=1}^{p-1} \lambda_j e_j + \lambda_p x_p$. Avec les conditions d'orthogonalité :

$$\forall j \in \{1, \dots, p-1\}, \langle e_p \mid e_j \rangle = 0$$

on déduit que :

$$\lambda_j + \lambda_p \langle x_p \mid e_j \rangle = 0 \quad (1 \leq j \leq p-1)$$

$$\text{et } e_p = \lambda_p \left(x_p - \sum_{j=1}^{p-1} \langle x_p | e_j \rangle e_j \right) = \lambda_p y_p.$$

Du fait que $x_p \notin \text{Vect}\{x_1, \dots, x_{p-1}\} = \text{Vect}\{e_1, \dots, e_{p-1}\}$, on déduit que $y_p \neq 0$ et la condition $\|e_p\| = 1$ donne $|\lambda_p| = \frac{1}{\|y_p\|}$. La condition supplémentaire :

$$0 < \langle x_p | e_p \rangle = \left\langle \frac{1}{\lambda_p} \left(e_p - \sum_{j=1}^{p-1} \lambda_j e_j \right) | e_p \right\rangle = \frac{1}{\lambda_p}$$

entraîne $\lambda_p > 0$, ce qui donne en définitive une unique solution pour e_p . $\square$

Si on désigne par $\mathcal{E} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$ muni de sa structure euclidienne canonique, tout système orthonormé $\mathcal{F} = (f_k)_{1 \leq k \leq n}$ est alors une base de $\mathbb{R}^n$ (ce système est libre formé de n vecteurs) et la matrice de passage Ω de la base canonique $\mathcal{E}$ à la base $\mathcal{F}$ est telle que ${}^t\Omega\Omega = I_n$. En effet les colonnes de Ω sont formées des composantes des vecteurs f_j dans la base $\mathcal{E}$ et le coefficient d'indice (i, j) de ${}^t\Omega\Omega$ est $\langle f_i | f_j \rangle = \delta_{ij}$.

Définition 2.9. On dit qu'une matrice carrée A à coefficients réels d'ordre n est orthogonale si ${}^tAA = I_n$. On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales.

On peut également dire qu'une matrice orthogonale est la matrice de passage de la base canonique de $\mathbb{R}^n$ à une base orthonormée où $\mathbb{R}^n$ est muni de sa structure euclidienne canonique.

$\mathcal{O}_n(\mathbb{R})$ est un sous groupe multiplicatif de $GL_n(\mathbb{R})$ et l'application $A \mapsto \det(A)$ réalise un morphisme de groupe de $\mathcal{O}_n(\mathbb{R})$ sur $\{-1, 1\}$. Le noyau de ce morphisme est un sous groupe distingué de $\mathcal{O}_n(\mathbb{R})$ appelé groupe des matrices de rotation. Il est noté $\mathcal{O}_n^+(\mathbb{R})$. On note $\mathcal{O}_n^-(\mathbb{R})$ le sous ensemble de $\mathcal{O}_n(\mathbb{R})$ formé des matrices telles que $\det(A) = -1$.

Corollaire 2.5 : Toute matrice $A \in GL_n(\mathbb{R})$ s'écrit de manière unique $A = \Omega T$, où Ω est une matrice orthogonale et T une matrice triangulaire supérieure à coefficients diagonaux strictement positifs.

Preuve. L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique, on désigne par $\mathcal{E} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$ et par $C_1, \dots, C_n$ les colonnes de la matrice A (ce sont des vecteurs de $\mathbb{R}^n$). Si $A \in GL_n(\mathbb{R})$, la famille de vecteurs $\mathcal{C} = (C_k)_{1 \leq k \leq n}$ est alors une base de $\mathbb{R}^n$ et on peut lui appliquer le procédé de Gram-Schmidt. Il existe donc une base orthonormée $\mathcal{F} = (f_k)_{1 \leq k \leq n}$ de $\mathbb{R}^n$ telle que $\text{Vect}\{f_1, \dots, f_k\} = \text{Vect}\{C_1, \dots, C_k\}$ pour tout $k \in \{1, \dots, n\}$, ce qui entraîne que la matrice de passage T_1 de la base $\mathcal{C}$ à la base $\mathcal{F}$ est triangulaire supérieure et la matrice de passage Ω de la base canonique $\mathcal{E}$ à la base $\mathcal{F}$ est orthogonale. En considérant que A est la matrice de passage de la base canonique $\mathcal{E}$ à la base $\mathcal{C}$, on a (relation de Chasles pour les matrices de passage) :

$$A = P_{\mathcal{E}, \mathcal{C}} = P_{\mathcal{E}, \mathcal{F}} P_{\mathcal{F}, \mathcal{C}} = \Omega T_1^{-1} = \Omega T$$

avec Ω orthogonale et T triangulaire supérieure. En considérant que les termes diagonaux de T sont les $t_{jj} = \langle C_j | f_j \rangle$ pour $1 \leq j \leq n$, on déduit que T est à termes diagonaux strictement positifs, si on a les conditions $\langle C_j | f_j \rangle > 0$ pour tout $j \in \{1, \dots, n\}$. S'il existe deux décompositions $A = \Omega T$ et $A = \Omega' T'$ avec Ω, Ω' orthogonales et T, T' triangulaires supérieures, la matrice $\Delta = T' T^{-1} = {}^t \Omega' \Omega$ est alors triangulaire supérieure orthogonale et $\Delta^{-1} = {}^t \Delta$ est à la fois triangulaire supérieure et inférieure, elle est donc diagonale et orthogonale. Les termes diagonaux de Δ sont donc égaux à ± 1 . Si on suppose de plus que T et T' sont à termes diagonaux strictement positifs, il en est alors de même de Δ et nécessairement $\Delta = I_n$, ce qui donne $T = T'$ et $\Omega = \Omega'$. D'où l'unicité de la décomposition. $\square$

Cette décomposition est parfois notée $A = QR$ avec Q orthogonale et R triangulaire supérieure de termes diagonaux strictement positifs et appelée décomposition QR de la matrice A . Cette décomposition conduit à un algorithme de calcul des valeurs propres d'une matrice symétrique réelle moyennant certaines conditions.

Théorème 2.10.

Dans un espace vectoriel E euclidien de dimension n , tout système orthonormé peut se prolonger en une base orthonormée.

Preuve. Soit $(e_k)_{1 \leq k \leq p}$ un système orthonormé dans E . Pour $p = n$, c'est une base de E . Pour $1 \leq p < n$, ce système est libre dans E , donc se prolonge en une base $(e_1, \dots, e_p, x_{p+1}, \dots, x_n)$ de E . Le procédé d'orthonormalisation de Gram-Schmidt fournit, vu l'unicité, le système orthonormé $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ qui est une base de E (c'est un système libre à n éléments). $\square$

Si $(e_k)_{1 \leq k \leq n}$ est une base orthonormée de E , tout vecteur $x \in E$ s'écrit alors $x = \sum_{k=1}^n \lambda_k e_k$ avec $\lambda_k = \langle x | e_k \rangle$ pour $1 \leq k \leq n$ et on a $\|x\|^2 = \sum_{k=1}^n |\lambda_k|^2$.

Corollaire 2.6 (Schur) : *Soit A une matrice réelle d'ordre $n > 0$ dont toutes les valeurs propres sont réelles. Il existe une matrice orthogonale Ω telle que ${}^t \Omega A \Omega$ soit triangulaire supérieure. C'est-à-dire que la matrice A se trigonalise dans une base orthonormée.*

Preuve. On procède par récurrence sur $n \geq 1$. Pour $n = 1$ le résultat est évident. Supposons le acquis pour $n - 1 \geq 1$ et soit A une matrice réelle d'ordre n dont toutes les valeurs propres sont réelles. On note u l'application linéaire ayant A pour matrice dans la base canonique de $\mathbb{R}^n$. Si λ_1 est une valeur propre de u et e_1 un vecteur propre associé unitaire, on complète (e_1) en une base orthonormée $\mathcal{B}_1 = (e_k)_{1 \leq k \leq n}$ de $\mathbb{R}^n$. La matrice de passage Ω_1 de la base canonique à cette base est orthogonale et la matrice de u dans $\mathcal{B}_1$ s'écrit $A_1 = {}^t \Omega_1 A \Omega_1 = \begin{pmatrix} \lambda_1 & a_1 \\ 0 & B_1 \end{pmatrix}$ avec $a_1 \in \mathcal{M}_{1, n-1}(\mathbb{K})$ (ensemble des matrices à une ligne et $n - 1$ colonnes) et $B_1 \in \mathcal{M}_{n-1}(\mathbb{K})$. Le polynôme caractéristique de A (égal à celui de A_1) est alors donné par $\chi_A(X) = (X - \lambda_1) \det(XI_{n-1} - B_1)$. On en déduit alors que la matrice B_1 a toutes ses valeurs propres réelles. On peut donc utiliser l'hypothèse de récurrence pour écrire qu'il existe une matrice orthogonale Ω'_2 d'ordre $n - 1$

telle que ${}^t\Omega'_2 B_1 \Omega'_2 = T_2$ soit triangulaire supérieure. La matrice $\Omega_2 = \begin{pmatrix} 1 & 0 \\ 0 & \Omega'_2 \end{pmatrix}$ est alors orthogonale d'ordre n et la matrice :

$${}^t\Omega_2 {}^t\Omega_1 A \Omega_1 \Omega_2 = \begin{pmatrix} \lambda_1 & a_1 \\ 0 & {}^t\Omega'_2 B_1 \Omega'_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & a_1 \\ 0 & T_2 \end{pmatrix}$$

est triangulaire supérieure. □

2.4 Réduction des matrices orthogonales

Pour ce paragraphe, on désigne par E l'espace vectoriel $\mathbb{R}^n$, par $(e_k)_{1 \leq k \leq n}$ la base canonique de E et le produit scalaire euclidien canonique de E est noté :

$$(x, y) = \left(\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i \right) \mapsto \langle x | y \rangle = \sum_{i=1}^n x_i y_i$$

Définition 2.10. Un endomorphisme $u \in \mathcal{L}(E)$ est dit orthogonal si :

$$\forall (x, y) \in E^2, \langle u(x) | u(y) \rangle = \langle x | y \rangle$$

La matrice d'un endomorphisme orthogonal dans une base orthonormée de E est une matrice orthogonale.

On note $\mathcal{O}(E)$ l'ensemble des endomorphismes orthogonaux de E .

Lemme 2.2 Soit $u \in \mathcal{L}(\mathbb{R}^n)$. Il existe un sous espace vectoriel P de $\mathbb{R}^n$ de dimension égale à 1 ou 2 et stable par u .

Preuve. Voir l'exercice 1.2. □

Lemme 2.3 Soient E un espace vectoriel euclidien et $u \in \mathcal{O}(E)$ un endomorphisme orthogonal. Il existe des sous espaces vectoriels $P_1, \dots, P_r$ de E , de dimension égale à 1 ou 2, deux à deux orthogonaux et stables par u tels que $E = \bigoplus_{j=1}^r P_j$.

Preuve. On procède par récurrence sur la dimension $n \geq 2$ de E . Pour $n = 2$, le résultat est évident. Supposons le acquis pour tout endomorphisme orthogonal sur un espace vectoriel euclidien de dimension p comprise entre 1 et $n - 1$, avec $n > 2$. Si P_1 est un sous espace vectoriel de E non réduit à $\{0\}$ de dimension au plus 2 stable par u , endomorphisme orthogonal, alors $P_1^\perp$ est stable par u . En effet $u(P_1) \subset P_1$ et $u \in GL(E)$ entraînent $u(P_1) = P_1$ (un isomorphisme conserve la dimension), donc tout $y \in P_1$ s'écrit $y = u(x)$ avec $x \in P_1$ et pour tout $z \in P_1^\perp$, on a $\langle u(z) | y \rangle = \langle u(z) | u(x) \rangle = \langle z | x \rangle = 0$, c'est-à-dire que $u(z) \in P_1^\perp$. Comme $0 < \dim(P_1^\perp) < \dim(E)$, on peut trouver des sous espaces vectoriels $P_2, \dots, P_r$ de E , de dimension au plus égale à 2, deux à deux orthogonaux et

stables par la restriction de u à $P_1^\perp$, donc par u , tels que $P_1^\perp = \bigoplus_{j=2}^r P_j$. On a alors

$$E = P_1 \oplus P_1^\perp = \bigoplus_{j=1}^r P_j. \quad \square$$

Théorème 2.11.

Soit A une matrice orthogonale d'ordre $n \geq 2$. Il existe une matrice orthogonale P telle que :

$$P^{-1}AP = \begin{pmatrix} I_p & 0 & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & R_1 & 0 & \ddots & 0 \\ 0 & \ddots & 0 & R_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 0 & R_r \end{pmatrix}$$

où, pour tout $k \in \{1, \dots, r\}$, on a noté $R_k = \begin{pmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{pmatrix}$ avec $\theta_k \in]0, 2\pi[\setminus \{\pi\}$.

Preuve. On procède par récurrence sur $n \geq 2$. Pour $n = 2$, dire que la matrice $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est orthogonale équivaut à :

$$\begin{cases} a^2 + b^2 = c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases} \quad \begin{cases} a^2 + c^2 = b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

On déduit alors que $b^2 = 1 - a^2 = c^2$, $a^2 = 1 - b^2 = d^2$, donc $c = \varepsilon b$, $d = \eta a$, avec ε, η dans $\{-1, 1\}$ tels que $\varepsilon\eta = -1$. Ce qui entraîne $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ ou

$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ avec $a^2 + b^2 = 1$. Réciproquement, les matrices de l'une de ces deux formes sont orthogonales. Dans le premier cas, on peut poser $a = \cos(\theta)$ avec $\theta \in [0, 2\pi[$ et $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. Dans le deuxième cas, la matrice A est symétrique de valeurs propres 1 et -1 , elle est donc orthogonalement semblable à $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Supposons le résultat acquis pour toute matrice orthogonale

d'ordre p compris entre 2 et $n-1$ et soit A une matrice orthogonale d'ordre $n > 2$. On désigne par u l'endomorphisme orthogonal ayant A pour matrice dans la base canonique de $E = \mathbb{R}^n$ muni de sa structure euclidienne canonique. Si u admet 1 ou -1 comme valeur propre, pour tout vecteur propre x associé à cette valeur propre, le sous espace vectoriel $(\mathbb{R}x)^\perp$ est stable par u . En effet, pour tout y dans $(\mathbb{R}x)^\perp$, on a $\langle u(y) | x \rangle = \pm \langle u(y) | u(x) \rangle = \pm \langle y | x \rangle = 0$ et il existe une base orthonormée $\mathcal{B}$ de $(\mathbb{R}x)^\perp$ dans laquelle la matrice de la restriction de u à $(\mathbb{R}x)^\perp$

est de la forme :

$$A' = \begin{pmatrix} I_p & 0 & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & R_1 & 0 & \ddots & 0 \\ 0 & \ddots & 0 & R_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 0 & R_r \end{pmatrix}$$

Dans la base orthonormée $\left\{ \frac{1}{\|x\|_2} x \right\} \cup \mathcal{B}$ la matrice de u est $A'' = \begin{pmatrix} \pm 1 & 0 \\ 0 & A' \end{pmatrix}$, qui est bien de la forme souhaitée. Si toutes les valeurs propres de u sont complexes non réelles, alors dans une décomposition $E = \bigoplus_{k=1}^r P_k$ du type lemme 2.3 tous les sous espaces vectoriels P_k sont de dimension 2 et la restriction de u à chacun des P_k a ses valeurs propres complexes non réelles. D'après l'étude du cas $n = 2$ (les P_k sont stables par u), on déduit alors qu'il existe une base orthonormée de P_k dans laquelle la matrice de u est de la forme $R_k = \begin{pmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{pmatrix}$ avec $\theta_k \in]0, 2\pi[\setminus \{\pi\}$. En réunissant toutes ces bases, on obtient une base orthonormée de E dans laquelle la matrice de u est :

$$A' = \begin{pmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R_r \end{pmatrix}$$

□

On a $p = \dim(\ker(A - I_n))$ et $q = \dim(\ker(A + I_n))$ avec $p + q + 2r = n$. De plus $A \in \mathcal{O}_n^+(\mathbb{R})$ [resp. $A \in \mathcal{O}_n^-(\mathbb{R})$] si, et seulement si, q est pair [resp. impair].

2.5 Réduction des matrices symétriques réelles

On garde les notations du paragraphe 2.4.

Définition 2.11. Une matrice réelle A est dite *symétrique*, si ${}^t A = A$. On dit qu'elle est *symétrique positive* si on a de plus $\langle Ax | x \rangle \geq 0$ pour tout $x \in E$. On dit qu'elle est *symétrique définie positive* si on a de plus $\langle Ax | x \rangle > 0$ pour tout $x \in E \setminus \{0\}$.

On peut associer une forme quadratique à une matrice symétrique A en posant $q(x) = \langle Ax | x \rangle$ pour tout $x \in E$. La forme polaire associée est alors définie par :

$$\forall (x, y) \in E^2, \varphi(x, y) = \frac{1}{2} (q(x + y) - q(x) - q(y)) = \langle Ax | y \rangle$$

Pour A positive, on l'inégalité de Cauchy-Schwarz :

$$\forall (x, y) \in E^2, |\varphi(x, y)| \leq \sqrt{q(x)}\sqrt{q(y)}$$

Le cône isotrope de q est définie par $q^{-1}\{0\} = \{x \in E \mid q(x) = 0\}$. Pour A positive, on a $q^{-1}\{0\} = \ker(A)$. En effet, si $q(x) = 0$, on a alors :

$$\|Ax\|_2^2 = \langle Ax \mid Ax \rangle = \langle A(Ax) \mid x \rangle = \varphi(Ax, x) \leq \sqrt{q(Ax)}\sqrt{q(x)} = 0$$

et $Ax = 0$. La réciproque est claire.

Théorème 2.12.

Soit A une matrice symétrique réelle. Les valeurs propres de A sont toutes réelles et A se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe une matrice orthogonale P et une matrice diagonale D telles que $A = PD^tP$, les coefficients diagonaux de D étant les valeurs propres dans $\mathbb{R}$ de la matrice A .

Preuve. Soit $q : x \mapsto q(x) = \langle Ax \mid x \rangle$ la forme quadratique associée à la matrice A . Elle définit une application continue de E dans $\mathbb{R}$ et elle atteint son maximum sur la sphère unité de E (qui est compacte puisque E est de dimension finie). On peut donc poser $\lambda = \sup_{x \in E, \|x\|=1} q(x)$ et on a $\lambda = q(x_0)$ avec $\|x_0\|_2 = 1$. On en déduit alors que la forme quadratique

$$q_1 : x \mapsto q_1(x) = \lambda \|x\|_2^2 - q(x) = \langle \lambda x - Ax \mid x \rangle$$

est positive. Le vecteur x_0 qui est dans le cône isotrope de q_1 est alors dans le noyau de $\lambda I_n - A$. On a donc ainsi montré que λ est une valeur propre réelle de A et x_0 un vecteur propre réel associé. Une récurrence sur la dimension $n \geq 1$ de E permet alors de montrer que A se diagonalise dans une base orthonormée. En effet, du fait que $\langle Ax \mid x_0 \rangle = \langle x \mid Ax_0 \rangle = \lambda \langle x \mid x_0 \rangle$, on déduit que l'hyperplan $H = (\mathbb{R}x_0)^\perp$ est stable par A et dans une base orthogonale adaptée à la décomposition $E = \mathbb{R}x_0 \oplus H$, la matrice de l'application linéaire associée à la matrice A dans la base canonique est de la forme $A' = \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}$ avec B symétrique réelle d'ordre $n-1$. On applique alors l'hypothèse de récurrence à la matrice B pour conclure. $\square$

La démonstration de ce théorème de diagonalisation des matrices symétriques réelles est une adaptation au cas de la dimension finie d'un résultat plus général de diagonalisation d'un opérateur auto-adjoint compact sur un espace de Hilbert. Précisément si u est un endomorphisme auto-adjoint compact sur un espace de Hilbert E , il a alors toutes ses valeurs propres réelles et E est somme hilbertienne des espaces propres de u (voir [7], p. 125).

Avec l'exercice 2.7 on propose une autre démonstration du théorème précédent.

Corollaire 2.7 : *Une matrice symétrique réelle est définie positive [resp. positive] si, et seulement si, toutes ses valeurs propres sont strictement positives [resp. positives].*

Une autre caractérisation des matrices réelles symétriques définies positives est donnée avec l'exercice 2.6.

Théorème 2.13.

Une matrice symétrique réelle $A = ((a_{ij}))_{1 \leq i, j \leq n}$ à diagonale strictement dominante est définie positive si, et seulement si, on a $a_{ii} > 0$ pour tout i compris entre 1 et n .

Preuve. Supposons que $a_{ii} > 0$, pour tout $i = 1, \dots, n$. On sait déjà que les valeurs propres de A sont réelles. En reprenant la démonstration du théorème 1.5, on a $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| < a_{ii}$ et nécessairement $\lambda > 0$. La matrice A a donc toutes ses valeurs propres strictement positives, ce qui équivaut à dire qu'elle est définie positive. La réciproque provient de l'égalité $a_{ii} = \langle Ae_i | e_i \rangle$ où $(e_k)_{1 \leq k \leq n}$ désigne la base canonique de $\mathbb{R}^n$. $\square$

2.6 Tridiagonalisation des matrices symétriques réelles. Méthode de Householder

On a vu que toute matrice symétrique réelle est orthogonalement semblable à une matrice diagonale, donc tridiagonale. L'avantage de la méthode de tridiagonalisation due à Householder, décrite dans ce paragraphe, est qu'elle fournit un procédé algorithmique simple de réduction à la forme tridiagonale d'une matrice symétrique réelle sans avoir à calculer ses valeurs propres. Cette réduction est utilisée pour calculer des valeurs approchées des valeurs propres d'une matrice symétrique réelle (méthode de Givens-Householder, paragraphe 6.4).

Pour $n \geq 2$, l'espace vectoriel $\mathbb{R}^n$ est muni du produit scalaire euclidien canonique $\langle \cdot | \cdot \rangle$ et de la norme euclidienne associée $\|\cdot\|$. On désigne par $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$.

Pour tout réel α , le signe de α est défini par $\text{signe}(\alpha) = 0$ pour $\alpha = 0$ et $\text{signe}(\alpha) = \frac{\alpha}{|\alpha|}$ pour $\alpha \neq 0$.

Définition 2.12. On appelle matrice de Householder toute matrice réelle de la forme $P_u = I_n - 2u {}^t u$, où u est un vecteur unitaire dans $\mathbb{R}^n$.

Si, pour tout vecteur unitaire $u \in \mathbb{R}^n$, on note P_u la matrice de Householder qui lui est associée et H_u l'hyperplan orthogonal à u dans $\mathbb{R}^n$, on a alors :

$$\forall x \in H_u, P_u x = x - 2(u {}^t u) x = x - 2\langle u | x \rangle u = x$$

et, avec $P_u u = -u$, on déduit que P_u est la symétrie orthogonale par rapport à l'hyperplan H_u . En particulier l'inverse de P_u est donné par $P_u^{-1} = {}^t P_u = P_u$.

Lemme 2.4 Soit M une matrice symétrique réelle d'ordre $n \geq 2$. En utilisant les notations qui précèdent, on a :

$$P_u^{-1} M P_u = M - 2(v {}^t u + u {}^t v)$$

où $v = Mu - \langle u | Mu \rangle u$.

Preuve. On a $P_u^{-1}MP_u = P_uMP_u = M - 2u {}^t u M - 2Mu {}^t u + 4u {}^t u Mu {}^t u$ avec $u {}^t u M - u ({}^t u Mu) {}^t u = u {}^t (Mu) - u \langle u | Mu \rangle {}^t u = u {}^t v$ et :

$$Mu {}^t u - u ({}^t u Mu) {}^t u = (Mu - \langle u | Mu \rangle u) {}^t u = v {}^t u$$

d'où le résultat. $\square$

Dans ce qui suit, on se donne une matrice symétrique réelle $A = ((a_{ij}))_{1 \leq i, j \leq n}$ d'ordre $n \geq 3$. On note $x = (a_{i,1})_{1 \leq i \leq n}$ la première colonne de A et :

$$\begin{cases} x^{(2)} = x - a_{11}e_1 = \begin{pmatrix} 0 \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} \\ y^{(2)} = x^{(2)} + \text{signe}(a_{21}) \|x^{(2)}\| e_2 \\ P_1 = \begin{cases} I_n & \text{si } y^{(2)} = 0 \\ P_{u^{(2)}} & \text{si } y^{(2)} \neq 0 \end{cases} \end{cases}$$

où $u^{(2)} = \frac{1}{\|y^{(2)}\|} y^{(2)}$ pour $y^{(2)} \neq 0$. Enfin on note $A_1 = P_1^{-1}AP_1 = ((a_{ij}^1))_{1 \leq i, j \leq n}$.

Lemme 2.5 Avec les notations qui précèdent, la matrice A_1 est symétrique avec $a_{i1}^1 = 0$ pour $i = 3, 4, \dots, n$.

Preuve. Si $y^{(2)} = 0$, on a alors $a_{i1} = 0$ pour tout $i = 3, 4, \dots, n$ et $A_1 = A$ vérifie bien la condition voulue. On suppose donc que $y^{(2)} \neq 0$. On a $A_1 = P_1AP_1$ et A_1 est symétrique puisque P_1 et A le sont. La première colonne de A_1 est donnée par $A_1e_1 = P_1AP_1e_1$. Avec $e_1 \in (\mathbb{R}y^{(2)})^\perp$, on a $P_1e_1 = e_1$ et $A_1e_1 = P_1Ae_1 = P_1x$. En notant $y^{(2)} = x^{(2)} + z^{(2)}$, avec $z^{(2)} = \text{signe}(a_{21}) \|x^{(2)}\| e_2$, on a :

$$\langle y^{(2)} | x^{(2)} - z^{(2)} \rangle = \langle x^{(2)} + z^{(2)} | x^{(2)} - z^{(2)} \rangle = \|x^{(2)}\|^2 - \|x^{(2)}\|^2 = 0$$

donc :

$$\begin{cases} P_1(x^{(2)} + z^{(2)}) = -x^{(2)} - z^{(2)} \\ P_1(x^{(2)} - z^{(2)}) = x^{(2)} - z^{(2)} \end{cases}$$

et $P_1x^{(2)} = -z^{(2)} = -\text{signe}(a_{21}) \|x^{(2)}\| e_2$. Ce qui donne en définitive :

$$P_1x = P_1(a_{11}e_1 + x^{(2)}) = a_{11}e_1 - \text{signe}(a_{21}) \|x^{(2)}\| e_2$$

C'est à dire que :

$$a_{i1}^1 = \begin{cases} a_{11} & \text{si } i = 1 \\ -\text{signe}(a_{21}) \|x^{(2)}\| & \text{si } i = 2 \\ 0 & \text{si } i = 3, \dots, n \end{cases}$$

$\square$

Théorème 2.14.

Soit A une matrice symétrique réelle d'ordre $n \geq 3$. Il existe $n-2$ matrices orthogonales $P_1, \dots, P_{n-2}$ telles que la matrice :

$$T = P_{n-2}^{-1} P_{n-1}^{-1} \cdots P_1^{-1} A P_1 P_2 \cdots P_{n-2}$$

soit tridiagonale.

Preuve. On raisonne par récurrence sur $n \geq 3$. Pour $n = 3$, la matrice A est déjà tridiagonale. On suppose donc le résultat vrai pour les matrices symétriques d'ordre $n-1 \geq 3$. Avec les notations du lemme 2.5, on a :

$$A_1 = \begin{pmatrix} a_{11} & a_{12}^1 & 0 & \cdots & 0 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 & \cdots & a_{2n}^1 \\ 0 & a_{32}^1 & a_{33}^1 & \cdots & a_{3n}^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^1 & a_{n3}^1 & \cdots & a_{nn}^1 \end{pmatrix}$$

cette matrice étant symétrique ainsi que la matrice d'ordre $n-1$:

$$B = \begin{pmatrix} a_{22}^1 & \cdots & a_{2n}^1 \\ \vdots & \ddots & \vdots \\ a_{n2}^1 & \cdots & a_{nn}^1 \end{pmatrix}$$

Par hypothèse de récurrence, on peut trouver $Q_2, \dots, Q_{n-2}$ orthogonales telles que $T_1 = Q_{n-2}^{-1} \cdots Q_2^{-1} B Q_2 \cdots Q_{n-2}$ soit tridiagonale. On pose alors :

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

pour $j = 2, 3, \dots, n-2$ et :

$$T = \begin{pmatrix} a_{11} & a_{12}^1 & 0 & \cdots & 0 \\ a_{21}^1 & & & & \\ 0 & & T_1 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}$$

On a alors $T = P_{n-2}^{-1} \cdots P_2^{-1} A_1 P_2 \cdots P_{n-2} = P_{n-2}^{-1} \cdots P_2^{-1} P_1^{-1} A P_1 P_2 \cdots P_{n-2}$. $\square$

2.7 Espaces vectoriels hermitiens

Pour ce paragraphe, E est un espace vectoriel sur $\mathbb{C}$ de dimension $n \geq 1$.

Définition 2.13. On appelle produit scalaire hermitien sur E toute application

$$\begin{aligned}\varphi : E \times E &\rightarrow \mathbb{C} \\ (x, y) &\mapsto \varphi(x, y)\end{aligned}$$

qui vérifie les propriétés suivantes :

- pour tout y dans E l'application $x \mapsto \varphi(x, y)$ est linéaire ;
- $\varphi(y, x) = \overline{\varphi(x, y)}$ pour tous x, y dans E ;
- $\varphi(x, x) \geq 0$ pour tout x dans E ;
- pour x dans E , $\varphi(x, x) = 0$ équivaut à $x = 0$.

On note en général $(x, y) \mapsto \langle x | y \rangle$ un tel produit scalaire.

Définition 2.14. Un espace hermitien est un $\mathbb{C}$ -espace vectoriel de dimension finie muni d'un produit scalaire hermitien.

Dans la suite de ce paragraphe E désigne un espace hermitien et on note pour tout x dans E , $\|x\| = \sqrt{\langle x | x \rangle}$.

Théorème 2.15. Inégalité de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Preuve. Pour x, y fixés dans E on pose $\langle x | y \rangle = \rho e^{i\theta}$ avec $\rho = |\langle x | y \rangle| \in \mathbb{R}^+$ et $\theta \in [0, 2\pi[$. On désigne par P la fonction polynomiale définie par :

$$\forall t \in \mathbb{R}, P(t) = \|x + te^{i\theta}y\|^2 = \|y\|^2 t^2 + 2t|\langle x | y \rangle| + \|x\|^2$$

On a $P(t) \geq 0$ pour tout réel t et nécessairement le discriminant de ce polynôme de degré 2 est toujours négatif ou nul, soit $|\langle x | y \rangle|^2 - \|x\|^2 \|y\|^2 \leq 0$, ce qui équivaut à $|\langle x | y \rangle| \leq \|x\| \|y\|$. Si x et y sont liés, il est alors clair que l'égalité est réalisée. Réciproquement si l'égalité est réalisée, le polynôme P admet alors une racine réelle λ et $x + \lambda e^{i\theta}y = 0$, c'est-à-dire que x et y sont liés. $\square$

Théorème 2.16. Inégalité de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = \lambda y$ avec $\lambda \geq 0$ (on dit que x et y sont positivement liés).

Preuve. On a $\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re} \langle x | y \rangle + \|y\|^2$ et avec l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

ce qui équivaut à $\|x + y\| \leq \|x\| + \|y\|$. L'égalité est réalisée si, et seulement si, $\operatorname{Re} \langle x | y \rangle = |\langle x | y \rangle| = \|x\| \|y\|$, ce qui entraîne $|\langle x | y \rangle| = \|x\| \|y\|$ et il existe un

nombre complexe λ tel que $x = \lambda y$. Pour $x = 0$, l'égalité est réalisée et pour $x \neq 0$, on a $\operatorname{Re} \langle x | y \rangle = \operatorname{Re} \left(\lambda \|y\|^2 \right) = |\lambda| \|y\|^2$ avec $\|y\| \neq 0$, donc $\operatorname{Re}(\lambda) = |\lambda|$ et λ est un réel positif. $\square$

Corollaire 2.8 : *L'application $x \mapsto \|x\| = \sqrt{\langle x | x \rangle}$ définit une norme sur E .*

Comme dans le cas euclidien, on dit que deux vecteurs x et y dans E sont orthogonaux si $\langle x | y \rangle = 0$.

Le théorème de Pythagore est encore vrai sous la forme : si les vecteurs x et y sont orthogonaux dans E , on a alors $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. La réciproque étant fautive dans le cas hermitien (l'égalité $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ entraîne seulement $\operatorname{Re} \langle x | y \rangle = 0$).

On définit, comme dans le cas euclidien, les notions de famille orthogonale et orthonormale. Une famille orthogonale de vecteurs non nuls de E est libre.

Théorème 2.17. procédé d'orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_k)_{1 \leq k \leq p}$ dans E , il existe une unique famille orthonormée $(e_k)_{1 \leq k \leq p}$ dans E telle que :

$$\forall k \in \{1, 2, \dots, p\}, \begin{cases} \operatorname{Vect} \{e_1, \dots, e_k\} = \operatorname{Vect} \{x_1, \dots, x_k\} \\ \langle x_k | e_k \rangle > 0 \end{cases}$$

Preuve. La démonstration est analogue à celle du théorème 2.9. $\square$

En notant $\mathcal{E} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ muni de sa structure hermitienne canonique avec le produit scalaire $(x, y) \mapsto \langle x | y \rangle = \sum_{k=1}^n x_k \overline{y_k}$, tout système orthonormé $\mathcal{F} = (f_k)_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$ (ce système est libre formé de n vecteurs) et la matrice de passage U de la base canonique $\mathcal{E}$ à la base $\mathcal{F}$ est telle que $U^* U = I_n$, où $U^* = ((u_{ij}^*))_{1 \leq i, j \leq n}$ désigne la matrice adjointe de la matrice complexe $U = ((u_{ij}))_{1 \leq i, j \leq n}$, définie par :

$$u_{ij}^* = \overline{u_{ji}} \quad (1 \leq i, j \leq n)$$

En effet les colonnes de U sont formées des composantes des vecteurs f_j dans la base $\mathcal{E}$ et le coefficient d'indice (i, j) de $U^* U$ est $\langle f_j | f_i \rangle = \delta_{ij}$.

Définition 2.15. *On dit qu'une matrice carrée A à coefficients complexes d'ordre n est unitaire si $U^* U = I_n$. On note $\mathcal{U}_n(\mathbb{C})$ l'ensemble des matrices unitaires.*

On peut également dire qu'une matrice unitaire est la matrice de passage de la base canonique de $\mathbb{C}^n$ à une base orthonormée où $\mathbb{C}^n$ est muni de sa structure hermitienne canonique.

$\mathcal{U}_n(\mathbb{C})$ est un sous groupe multiplicatif de $GL_n(\mathbb{C})$ et l'application $A \mapsto \det(A)$ réalise un morphisme de groupe de $\mathcal{U}_n(\mathbb{C})$ sur le groupe multiplicatif des nombres complexes de module 1.

Corollaire 2.9 : Toute matrice $A \in GL_n(\mathbb{C})$ peut s'écrire de manière unique $A = UT$, où U est une matrice unitaire et T une matrice triangulaire supérieure à coefficients diagonaux strictement positifs.

Preuve. La démonstration est analogue à celle du corollaire 2.5. $\square$

Théorème 2.18.

Dans un espace vectoriel E hermitien de dimension n tout système orthonormé peut se prolonger en une base orthonormée.

Preuve. La démonstration est analogue à celle du théorème 2.10. $\square$

Corollaire 2.10 (Schur) : Soit A une matrice complexe d'ordre $n > 0$. Il existe alors une matrice unitaire U telle que U^*AU soit triangulaire supérieure. C'est-à-dire que la matrice A se trigonalise dans une base orthonormée.

Preuve. La démonstration est analogue à celle du corollaire 2.6. $\square$

2.8 Réduction des matrices normales

On note $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ et le produit scalaire hermitien canonique de $\mathbb{C}^n$ est noté :

$$(x, y) = \left(\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i \right) \mapsto \langle x | y \rangle = \sum_{i=1}^n x_i \overline{y_i}$$

La norme associée est la norme hermitienne notée $x \mapsto \|x\|$.

Définition 2.16. Une matrice complexe A est dite hermitienne si $A^* = A$. On dit qu'elle est hermitienne positive, si on a de plus $\langle Ax | x \rangle \geq 0$ pour tout $x \in E$. On dit qu'elle est hermitienne définie positive, si on a de plus $\langle Ax | x \rangle > 0$ pour tout $x \in E \setminus \{0\}$.

Définition 2.17. Une matrice complexe A est dite normale si $A^*A = AA^*$.

Les matrices hermitiennes et unitaires sont des cas particuliers de matrices normales.

Lemme 2.6 Les valeurs propres d'une matrice hermitienne [resp. hermitienne positive] sont réelles [resp. réelles et positives]. Les valeurs propres d'une matrice unitaire sont de module égal à 1.

Preuve. Soient $A \in \mathcal{M}_n(\mathbb{C})$ hermitienne, $\lambda \in \mathbb{C}$ une valeur propre de A et $x \in \mathbb{C}^n$ un vecteur propre non nul associé. On a :

$$\lambda \|x\|^2 = \langle Ax | x \rangle = \langle x | A^*x \rangle = \langle x | Ax \rangle = \bar{\lambda} \|x\|^2$$

donc $\lambda = \bar{\lambda}$ et $\lambda \in \mathbb{R}$. Si de plus A est positive, on a alors $\lambda \|x\|^2 = \langle Ax | x \rangle \geq 0$ avec $x \neq 0$, ce qui entraîne $\lambda \geq 0$.

Soient $A \in \mathcal{M}_n(\mathbb{C})$ unitaire, $\lambda \in \mathbb{C}$ une valeur propre de A et $x \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé. On a :

$$|\lambda|^2 \|x\|^2 = \|Ax\|^2 = \langle Ax | Ax \rangle = \langle x | A^* Ax \rangle = \|x\|^2$$

et nécessairement $|\lambda| = 1$. □

Théorème 2.19.

Si A une matrice complexe normale, elle se diagonalise alors dans une base orthonormée, c'est-à-dire qu'il existe une matrice unitaire U et une matrice diagonale D telles que $A = UDU^*$, les coefficients diagonaux de D étant les valeurs propres dans $\mathbb{C}$ de la matrice A .

Preuve. On montre tout d'abord que pour toute valeur propre $\lambda \in \mathbb{C}$ de A , le sous espace propre associé $E_\lambda = \ker(A - \lambda I_n)$ est stable par A^* et que son orthogonal $E_\lambda^\perp$ est stable par A et par A^* , ce qui permet de raisonner par récurrence sur l'ordre n de la matrice A .

Pour tout $x \in E_\lambda$ on a, puisque A est normale, $AA^*x = A^*Ax = \lambda A^*x$ et $A^*x \in E_\lambda$, c'est-à-dire que E_λ est stable par A^* .

Pour $x \in E_\lambda^\perp$ et $y \in E_\lambda$, on a $A^*y \in E_\lambda$ et $\langle Ax | y \rangle = \langle x | A^*y \rangle = 0$, c'est-à-dire que $Ax \in E_\lambda^\perp$ et $E_\lambda^\perp$ est stable par A . Puis en écrivant que :

$$\langle A^*x | y \rangle = \langle x | Ay \rangle = \langle x | \lambda y \rangle = \bar{\lambda} \langle x | y \rangle = 0$$

on déduit que $A^*x \in E_\lambda^\perp$ et $E_\lambda^\perp$ est stable par A^* .

On raisonne ensuite par récurrence sur la dimension n de l'espace vectoriel hermitien E . Pour $n = 1$ le résultat est évident. Supposons le acquis pour toute matrice normale d'ordre $p \in \{1, \dots, n-1\}$. Soient λ une valeur propre de A et E_λ l'espace propre associé. Si $E_\lambda = E$, la matrice A est alors une homothétie et toute base orthonormée de E convient. Si $E_\lambda \neq E$, $E_\lambda^\perp$ est alors de dimension $p \in \{1, \dots, n-1\}$ et la restriction B de A à $E_\lambda^\perp$ est un endomorphisme normal de $E_\lambda^\perp$ (une matrice est identifiée à l'application linéaire qu'elle définit dans la base canonique de E). Il existe donc une base orthonormée de $E_\lambda^\perp$ formée de vecteurs propres de B , donc de A . En complétant cette base par une base orthonormée de E_λ on obtient une base orthonormée de E formée de vecteurs propres de A . □

Si λ est une valeur propre de A , la stabilité de $E_\lambda^\perp$ par A^* est en fait vraie pour toute matrice A .

Du fait qu'une matrice hermitienne ou unitaire est normale on déduit les résultats suivants.

Corollaire 2.11 : *Une matrice hermitienne [resp. unitaire] a ses valeurs propres réelles [resp. de module 1] et se diagonalise dans une base orthonormée.*

De ce résultat on déduit l'existence d'une racine carrée pour une matrice complexe hermitienne positive.

Corollaire 2.12 : *Si A est une matrice complexe hermitienne positive, il existe alors une unique matrice hermitienne positive B telle que $A = B^2$.*

Preuve. La matrice A étant hermitienne positive a toutes ses valeurs propres réelles positives et se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe $U \in U_n(\mathbb{C})$ et $(\lambda_k)_{1 \leq k \leq n} \in (\mathbb{R}^+)^n$ tels que $A = UDU^*$, avec :

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

En posant :

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

et $B = U\Delta U^*$, on a $B^2 = A$, la matrice B étant hermitienne positive (ses valeurs propres sont positives).

Si P est le polynôme d'interpolation de Lagrange défini par $P(\lambda_i) = \sqrt{\lambda_i}$ pour tout $i \in \{1, \dots, n\}$ (le degré de P est $p-1$ où p est le nombre de valeurs propres distinctes de A), on a alors $P(A) = UP(D)U^* = U\Delta U^* = B$, c'est-à-dire que B est polynomiale en A . Si C est une autre racine carrée de A hermitienne positive, on a alors $C^2 = A$ et C commute avec A , donc avec B . En définitive les matrices B et C commutent et sont diagonalisables, on sait alors qu'elles sont simultanément diagonalisables (exercice 2.3), c'est-à-dire qu'il existe une matrice P dans $GL_n(\mathbb{C})$ telle que $C = P\Gamma P^{-1}$ et $B = P\Lambda P^{-1}$ avec Γ et Λ diagonales à coefficients réels positifs. De $C^2 = A = B^2$, on déduit alors que $\Gamma^2 = \Lambda^2$ et $\Gamma = \Lambda$ du fait que ces matrices sont diagonales à coefficients réels positifs. Et en définitive $B = C$. $\square$

Une matrice réelle symétrique et positive admet également une unique racine carrée symétrique positive.

On montre de manière analogue que si A est une matrice complexe hermitienne [resp. réelle symétrique] positive, alors pour tout entier $k > 0$, il existe une unique matrice complexe hermitienne [resp. réelle symétrique] positive B telle que $A = B^k$. De plus, si A est définie positive, il en est de même de B .

De ce résultat on peut déduire l'existence de la décomposition polaire d'une matrice réelle ou complexe inversible. Précisément on a le résultat suivant.

Corollaire 2.13 : *Toute matrice complexe [resp. réelle] inversible A peut s'écrire de manière unique $A = UH$ [resp. $A = \Omega S$] où U [resp. Ω] est une matrice unitaire [resp. orthogonale] et H [resp. S] une matrice hermitienne [resp. symétrique] définie positive.*

Preuve. On démontre le résultat dans le cas complexe. Le cas des matrices réelles se traitant de manière analogue. Si $A = UH$, on a alors $A^*A = HU^*UH = H^2$ et H

est la racine carrée de la matrice hermitienne définie positive A^*A ($\langle A^*Ax \mid x \rangle = \|Ax\|_2^2 > 0$ pour x non nul). La matrice U est alors donnée par $U = AH^{-1}$ (A inversible entraîne H inversible). On a donc, en cas d'existence, l'unicité des matrices U et H . Pour $A \in GL_n(\mathbb{C})$, la matrice A^*A est hermitienne définie positive et elle admet une unique racine carrée hermitienne définie positive H . En posant $U = AH^{-1}$, on a $A = UH$ et :

$$U^*U = (H^{-1})^* (A^*A) H^{-1} = (H^*)^{-1} H^2 H^{-1} = H^{-1}H = I_n$$

c'est-à-dire que U est unitaire. $\square$

2.9 Forme réduite de Jordan

Dans le cas d'un corps algébriquement clos, on peut affiner le théorème 2.1 de trigonalisation comme on va le voir dans ce paragraphe.

Pour ce paragraphe, E est un espace vectoriel de dimension $n \geq 2$ sur $\mathbb{K}$ et on désigne par E^* le dual algébrique de E , c'est-à-dire l'espace vectoriel des formes linéaires sur E .

L'orthogonal dans E d'une partie non vide Y de E^* est le sous espace vectoriel Y° de E défini par $Y^\circ = \{x \in E \mid \forall \varphi \in Y, \varphi(x) = 0\}$.

La base duale d'une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E est la base $\mathcal{B}^* = (e_k^*)_{1 \leq k \leq n}$ de E^* définie par :

$$\forall j \in \{1, \dots, n\}, \forall x = \sum_{i=1}^n x_i e_i, e_j^*(x) = x_j$$

Pour tout endomorphisme $u \in \mathcal{L}(E)$, on note ${}^t u \in \mathcal{L}(E^*)$ le transposé de u défini par :

$$\forall \varphi \in E^*, {}^t u(\varphi) = \varphi \circ u$$

Si $u \in \mathcal{L}(E)$ a pour matrice A dans une base $\mathcal{B}$ de E , la matrice de ${}^t u$ dans la base duale $\mathcal{B}^*$ est alors ${}^t A$.

Lemme 2.7 *Si $u \in \mathcal{L}(E)$ est nilpotent d'ordre $r \geq 1$, il en est alors de même de ${}^t u$ dans $\mathcal{L}(E^*)$.*

Preuve. Pour u, v dans $\mathcal{L}(E)$, on a ${}^t(u \circ v) = {}^t v \circ {}^t u$, donc $({}^t u)^r = {}^t(u^r) = 0$ et $({}^t u)^{r-1} = {}^t(u^{r-1}) \neq 0$, c'est-à-dire que ${}^t u$ est nilpotent d'ordre r . $\square$

Lemme 2.8 *Soit $u \in \mathcal{L}(E)$ nilpotent d'ordre $r \geq 1$. Pour tout vecteur $x \in E$ tel que $u^{r-1}(x) \neq 0$, la famille $\mathcal{B}_x = (u^k(x))_{0 \leq k \leq r-1}$ est libre et l'espace vectoriel $F_x = \text{Vect}(\mathcal{B}_x)$ est stable par u .*

Preuve. Comme $u^{r-1} \neq 0$, il existe des vecteurs $x \in E \setminus \{0\}$ tels que $u^{r-1}(x) \neq 0$. Si, pour un tel vecteur $x \in E \setminus \{0\}$, la famille $\mathcal{B}_x$ est liée, il existe alors des scalaires

$\lambda_0, \dots, \lambda_{r-1}$ non tous nuls tels que $\sum_{k=0}^{r-1} \lambda_k u^k(x) = 0$. En désignant par p le plus

petit entier compris entre 0 et $r - 1$ tel que $\lambda_p \neq 0$, on a nécessairement $p \leq r - 2$ (puisque $u^{r-1}(x) \neq 0$) et :

$$\lambda_p u^p(x) = - \sum_{k=p+1}^{r-1} \lambda_k u^k(x) = - \sum_{j=1}^{r-1-p} \lambda_{p+j} u^{p+j}(x)$$

donc $\lambda_p u^{r-1}(x) = - \sum_{j=1}^{r-1-p} \lambda_{p+j} u^{r+j-1}(x) = 0$ (on a $u^{r+k} = 0$ pour $k \geq 0$), ce

qui n'est pas. La famille $\mathcal{B}_x$ est donc libre et $F_x = \text{Vect}(\mathcal{B}_x)$ est un sous-espace vectoriel de dimension r de E . De l'égalité $u^r = 0$, on déduit que cet espace est stable par u . $\square$

Lemme 2.9 Soit $u \in \mathcal{L}(E)$ nilpotent d'ordre $r \geq 1$. Il existe une forme linéaire $\varphi \in E^*$ et un vecteur $x \in E$ tels que $F_x = \text{Vect}\{x, u(x), \dots, u^{r-1}(x)\}$ et l'orthogonal G_φ dans E de $H_\varphi = \text{Vect}\{\varphi, {}^t u(\varphi), \dots, ({}^t u)^{r-1}(\varphi)\}$ sont stables par u avec $E = F_x \oplus G_\varphi$.

Preuve. Comme ${}^t u \in \mathcal{L}(E^*)$ est nilpotent d'ordre r , pour toute forme linéaire $\varphi \in E^*$ telle que $({}^t u)^{r-1}(\varphi) \neq 0$, l'espace $H_\varphi = \text{Vect}\{\varphi, {}^t u(\varphi), \dots, ({}^t u)^{r-1}(\varphi)\}$ est de dimension r stable par ${}^t u$ et son orthogonal dans E , à savoir l'espace $G_\varphi = H_\varphi^\circ = \{x \in E \mid \forall \psi \in H_\varphi, \psi(x) = 0\}$, est de dimension $n - r$ stable par u (pour $x \in G_\varphi$ et $\psi \in H_\varphi$, on a ${}^t u(\psi) \in H_\varphi$, donc $\psi(u(x)) = {}^t u(\psi)(x) = 0$ et $u(x) \in G_\varphi$). Comme $({}^t u)^{r-1}(\varphi) = \varphi \circ u^{r-1} \neq 0$, pour tout $x \in E$ tel que $\varphi(u^{r-1}(x)) \neq 0$ l'espace $F_x = \text{Vect}\{x, u(x), \dots, u^{r-1}(x)\}$ est de dimension r stable par u . Enfin, pour vérifier que $E = F_x \oplus G_\varphi$, il nous suffit de vérifier que

$F_x \cap G_\varphi = \{0\}$ puisque $\dim(F_x) + \dim(G_\varphi) = n$. Si $y = \sum_{k=0}^{r-1} \lambda_k u^k(x) \in F_x \cap G_\varphi$, on a alors $u^{r-1}(y) \in G_\varphi$ (G_φ est stable par u), donc $0 = \varphi(u^{r-1}(y)) = \lambda_0 \varphi(u^{r-1}(x))$ et $\lambda_0 = 0$. Supposant que $\lambda_0 = \dots = \lambda_j = 0$ pour $0 \leq j \leq r-2$ (si $r = 1$ c'est fini), on a $y = \sum_{k=j+1}^{r-1} \lambda_k u^k(x)$ et $0 = \varphi(u^{r-j-2}(y)) = \lambda_{j+1} \varphi(u^{r-1}(x))$, donc $\lambda_{j+1} = 0$. $\square$

De ces lemmes on déduit les résultats suivants.

Théorème 2.20.

Si $u \in \mathcal{L}(E)$ est nilpotent d'ordre $r \geq 1$, il existe alors une base $\mathcal{B} = \bigcup_{i=1}^p \mathcal{B}_i$ de E telle que chaque sous-espace vectoriel $F_i = \text{Vect}(\mathcal{B}_i)$ soit stable par u

et la matrice dans la base $\mathcal{B}_i$ de la restriction u_i de u à F_i est :

$$J_i = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in M_{n_i}(\mathbb{K})$$

où $n_i = \dim(E_i)$ ($1 \leq i \leq p$), la suite $(n_i)_{1 \leq i \leq p}$ étant décroissante avec $n_1 = r$. La matrice de u dans la base $\mathcal{B}$ est donc :

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix}$$

Preuve. On procède par récurrence sur $n = \dim(E)$. Pour $n = 1$, on a $u = 0$ et le résultat est trivial. Supposons le acquis pour les espaces vectoriels de dimension strictement inférieure à n . Avec les notations du lemme précédent, la matrice de $u|_{F_x}$ (F_x est stable par u) dans la base $\mathcal{B}_1 = (u^k(x))_{0 \leq k \leq r-1}$ est :

$$J_r = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in M_r(\mathbb{K})$$

Si $r = n$ c'est fini, sinon en complétant cette base par une base $\mathcal{B}_\varphi$ de G_φ (notation du lemme précédent), la matrice de u dans la base $\mathcal{B}_1 \cup \mathcal{B}_\varphi$ est $A = \begin{pmatrix} J_r & 0 \\ 0 & A_{n-r} \end{pmatrix}$ (G_φ est stable par u) où $A_{n-r} \in \mathcal{M}_{n-r}(\mathbb{K})$ est la matrice de $u|_{G_\varphi}$ dans $\mathcal{B}_\varphi$, cette matrice étant nilpotente d'indice au plus égal à r avec G_φ de dimension strictement inférieure à n . L'hypothèse de récurrence permet alors de conclure. $\square$

Théorème 2.21. Jordan

Soit $u \in \mathcal{L}(E) \setminus \{0\}$ tel que son polynôme caractéristique s'écrive

$$\chi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k} \text{ avec } \alpha_k > 0 \text{ et les } \lambda_k \text{ distincts deux à deux.}$$

Il existe une base $\mathcal{B}$ de E dans laquelle la matrice de u est de la forme :

$$A = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix} \quad (2.1)$$

avec :

$$\forall k \in \{1, 2, \dots, p\}, J_k = \begin{pmatrix} \lambda_k & 0 & 0 & \cdots & 0 \\ \varepsilon_{k,2} & \lambda_k & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \varepsilon_{k,\alpha_k-1} & \lambda_k & 0 \\ 0 & \cdots & 0 & \varepsilon_{k,\alpha_k} & \lambda_k \end{pmatrix} \in M_{\alpha_k}(\mathbb{C})$$

où $\varepsilon_{k,i} \in \{0, 1\}$ (forme réduite de Jordan).

Preuve. En désignant pour tout k compris entre 1 et p , par $N_k = \ker(u - \lambda_k Id)^{\beta_k}$ les sous-espaces caractéristiques de u où β_k est la multiplicité de λ_k comme racine du polynôme minimal (théorème 1.12), on a $E = \bigoplus_{k=1}^p N_k$. Chaque sous espace vectoriel N_k est de dimension α_k stable par u , λ_k est la seule valeur propre de la restriction de u à N_k et la restriction de $u - \lambda_k Id$ à N_k est nilpotente d'indice β_k . On déduit alors qu'il existe une base $\mathcal{B}_k$ de N_k dans laquelle la matrice de la restriction de $u - \lambda_k Id$ à N_k est de la forme :

$$J_k = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \varepsilon_{k,2} & 0 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \varepsilon_{k,\alpha_k-1} & 0 & 0 \\ 0 & \cdots & 0 & \varepsilon_{k,\alpha_k} & 0 \end{pmatrix} \in M_{\alpha_k}(\mathbb{C})$$

Dans la réunion de ces bases, la matrice de u a la forme indiquée. $\square$

Corollaire 2.14 : Toute matrice non nulle A d'ordre n à coefficients complexes est semblable à une matrice triangulaire de la forme (2.1).

Avec les exercices 2.15 et 2.7, on propose deux autres démonstration du théorème 2.20 sans utiliser la dualité.

2.10 Exercices

Exercice 2.1. Soit $u \in \mathcal{L}(E)$ diagonalisable. Montrer que si F est un sous espace vectoriel de E stable par u , la restriction de u à F est alors aussi diagonalisable.

Solution. Dire que u est diagonalisable équivaut à dire que son polynôme minimal π_u est scindé sur $\mathbb{K}$ à racines simples. Si F est stable par u , la restriction v de u à F est alors un endomorphisme de F annulé par le polynôme π_u qui est scindé sur $\mathbb{K}$ à racines simples, il en résulte que v est diagonalisable.

Exercice 2.2. Soit $u \in \mathcal{L}(E)$ diagonalisable de valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_p$ ($1 \leq p \leq n$). Montrer que pour $1 \leq k \leq p$ la projection de E sur le sous espace propre $\ker(u - \lambda_k Id)$ est donnée par

$$p_k = \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) \text{ où } \alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)} \text{ (utiliser la décomposition en éléments simples de } \frac{1}{\pi_u}).$$

Solution. Si $u \in \mathcal{L}(E)$ est diagonalisable, son polynôme minimal s'écrit alors $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)$, où $\lambda_1, \dots, \lambda_p$ ($1 \leq p \leq n$) sont les valeurs propres deux à

deux distinctes de u . La décomposition en éléments simples $\frac{1}{\pi_u(X)} = \sum_{k=1}^p \frac{\alpha_k}{X - \lambda_k}$,

où $\alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)}$, nous donne l'égalité de Bézout $\sum_{k=1}^p \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (X - \lambda_j) = 1$

qui nous dit que $x = \sum_{k=1}^p \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) x$ pour tout $x \in E$. Considérant que

pour $k \in \{1, \dots, p\}$ et $x \in E$, on a $(u - \lambda_k Id) \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) (x) = \pi_u(u)(x) = 0$,

on déduit que $\alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) x \in \ker(u - \lambda_k Id)$, c'est à dire que les projec-

teurs p_k s'écrivent $p_k = \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id)$. En particulier ces projecteurs sont des polynômes en u .

Exercice 2.3. Soit $(u_i)_{i \in I}$ une famille d'endomorphismes de E (I étant un ensemble d'indice non nécessairement fini). On suppose que les u_i commutent deux à deux et sont diagonalisables. Montrer qu'il existe une base commune de diagonalisation.

Solution. On procède par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$, le résultat est évident. On suppose que E est de dimension $n + 1$ et que le résultat est acquis pour les espaces vectoriels de dimension inférieure ou égale à n . Si tous les u_i sont des homothéties, le résultat est alors clair. Sinon on se donne $j \in I$ tel que u_j ne soit pas une homothétie. On a la décomposition en sous espaces propres

$$E = \bigoplus_{k=1}^p \ker(u_j - \lambda_k Id).$$

L'endomorphisme u_j n'étant pas une homothétie, chaque sous espace propre est de dimension inférieure ou égale à n . Comme tous les u_i commutent à u_j , chaque sous espace propre est stable par u_i pour tout i dans I et la restriction de chaque u_i à chaque $\ker(u_j - \lambda_k Id)$ est diagonalisable. On peut donc appliquer, pour tout entier k compris entre 1 et p l'hypothèse de récurrence à la famille des restrictions des u_i à $\ker(u_j - \lambda_k Id)$, ce qui permet de construire une base de diagonalisation de $\ker(u_j - \lambda_k Id)$ commune à toutes les restrictions de u_i à cet espace. La réunion de ces bases donne une base de diagonalisation commune à tous les u_i .

Exercice 2.4. Soit $(E, \|\cdot\|)$ un espace vectoriel normé réel. On note :

$$\mu(E) = \sup_{(x,y) \neq (0,0)} \frac{\|x+y\|^2 + \|x-y\|^2}{2(\|x\|^2 + \|y\|^2)}$$

Montrer que les assertions suivantes sont équivalentes :

1. $\forall (x, y) \in E^2, \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (identité du parallélogramme) ;
2. la norme $\|\cdot\|$ dérive d'un produit scalaire sur E ;
3. $\mu(E) = 1$;
4. pour x, y fixés dans E , $\|x+ty\|^2$ est un trinôme en t .

Solution.

(1) $\Rightarrow$ (2) On définit la fonction $\langle \cdot | \cdot \rangle$ sur $E \times E$ par :

$$(x, y) \mapsto \langle x | y \rangle = \frac{1}{4} (\|x+y\|^2 - \|x-y\|^2)$$

Il est facile de vérifier que cette application est symétrique ($\langle x | y \rangle = \langle y | x \rangle$) et définie positive ($\langle x | x \rangle = \|x\|^2 \geq 0$ avec égalité si, et seulement si, $x = 0$). En utilisant l'identité du parallélogramme on va montrer que cette application est bilinéaire. C'est donc un produit scalaire et la norme $\|\cdot\|$ en dérive. Avec

l'identité du parallélogramme on a, pour x, y, z dans E :

$$\begin{aligned}
 & \langle x + z \mid y \rangle + \langle x - z \mid y \rangle \\
 &= \frac{\|x + z + y\|^2 - \|x + z - y\|^2 + \|x - z + y\|^2 - \|x - z - y\|^2}{4} \\
 &= \frac{\left(\|x + y + z\|^2 + \|x + y - z\|^2\right) - \left(\|x - y + z\|^2 + \|x - y - z\|^2\right)}{4} \\
 &= \frac{\left(\|x + y\|^2 + \|z\|^2\right) - \left(\|x - y\|^2 + \|z\|^2\right)}{2} = 2 \langle x \mid y \rangle
 \end{aligned}$$

ce qui donne pour $z = x$:

$$\forall (x, y) \in E^2, \langle 2x \mid y \rangle = 2 \langle x \mid y \rangle$$

En posant $u = \frac{1}{2}(x + y)$, $v = \frac{1}{2}(x - y)$, on a alors :

$$\begin{aligned}
 \langle x + y \mid z \rangle &= \langle 2u \mid z \rangle = 2 \langle u \mid z \rangle = \langle u + v \mid z \rangle + \langle u - v \mid z \rangle \\
 &= \langle x \mid z \rangle + \langle y \mid z \rangle
 \end{aligned}$$

On a donc $\langle x + y \mid z \rangle = \langle x \mid z \rangle + \langle y \mid z \rangle$, pour tout (x, y, z) dans E^3 . Par récurrence, on déduit que :

$$\forall (x, y) \in E^2, \forall n \in \mathbb{N}^*, \langle nx \mid y \rangle = n \langle x \mid y \rangle \quad (2.2)$$

Avec $\langle 0 \mid y \rangle = 0$ et $\langle -x \mid y \rangle = -\langle x \mid y \rangle$, on déduit que (2.2) est valable pour tout entier n dans $\mathbb{Z}$. Puis avec $\langle x \mid y \rangle = n \left\langle \frac{1}{n}x \mid y \right\rangle$ pour tout entier $n \geq 1$, on déduit que (2.2) est valable pour tout $r \in \mathbb{Q}$. Enfin avec la densité de $\mathbb{Q}$ dans $\mathbb{R}$ et la continuité de l'application :

$$(x, y) \mapsto \langle x \mid y \rangle = \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right)$$

on déduit que (2.2) est valable pour tout $r \in \mathbb{R}$. Ce qui achève de prouver que l'application $\langle \cdot \mid \cdot \rangle$ est bilinéaire.

- propriété 2) $\Rightarrow$ (3) Si la norme dérive d'un produit scalaire, on vérifie alors facilement qu'on a l'identité du parallélogramme et l'égalité $\mu(E) = 1$.
- (3) $\Rightarrow$ (1) Si $\mu(E) = 1$, on a alors :

$$\forall (x, y) \in E^2, \|x + y\|^2 + \|x - y\|^2 \leq 2 \left(\|x\|^2 + \|y\|^2 \right)$$

En posant $u = x + y$, $v = x - y$, on a :

$$\begin{aligned}
 \|2x\|^2 + \|2y\|^2 &= \|u + v\|^2 + \|u - v\|^2 \\
 &\leq 2 \left(\|u\|^2 + \|v\|^2 \right) = 2 \left(\|x + y\|^2 + \|x - y\|^2 \right)
 \end{aligned}$$

soit $2 \left(\|x\|^2 + \|y\|^2 \right) \leq \|x + y\|^2 + \|x - y\|^2$ et l'égalité.

On a donc les équivalences (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) .

(2) $\Rightarrow$ (4) C'est évident.

(4) $\Rightarrow$ (1) Pour x, y fixés dans E on peut écrire $P(t) = \|x + ty\|^2 = at^2 + 2bt + c$, pour tout réel t . Le coefficient c est donné en faisant $t = 0$, soit $c = \|x\|^2$ et le coefficient a est donné par :

$$a = \lim_{t \rightarrow +\infty} \frac{P(t)}{t^2} = \lim_{t \rightarrow +\infty} \left\| \frac{1}{t}x + y \right\|^2 = \|y\|^2$$

On en déduit alors que :

$$\|x + y\|^2 + \|x - y\|^2 = P(1) + P(-1) = 2(a + c) = 2(\|x\|^2 + \|y\|^2)$$

c'est-à-dire l'identité du parallélogramme.

Exercice 2.5.

1. Montrer que l'application $(A, B) \mapsto \langle A | B \rangle = \text{Tr}(A^t B)$ définit un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$.
2. Montrer que pour toute forme linéaire $\varphi \in \mathcal{M}_n(\mathbb{R})^*$ il existe une unique matrice $C \in \mathcal{M}_n(\mathbb{R})$ telle que :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \varphi(A) = \text{Tr}(A^t C)$$

3. Montrer que l'application déterminant, $\det : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ est différentiable et que sa différentielle u est donnée par :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \forall H \in \mathcal{M}_n(\mathbb{R}), u(A)(H) = \text{Tr}(H^t C)$$

où C est la comatrice de la matrice A .

4. Soit Y une fonction définie sur un intervalle réel I à valeurs dans $\mathcal{M}_n(\mathbb{R})$ et dérivable. Montrer que la fonction $f : t \mapsto \det(Y(t))$ est dérivable et calculer sa dérivée.

Solution.

1. Pour toutes matrice A, B dans $\mathcal{M}_n(\mathbb{R})$, on a $\langle A | B \rangle = \sum_{1 \leq i, j \leq n} a_{ij} b_{ij}$, c'est-à-dire que $\langle \cdot | \cdot \rangle$ est le produit scalaire euclidien de $\mathcal{M}_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$.
2. La base canonique $(E_{i,j})_{1 \leq i, j \leq n}$ de $\mathcal{M}_n(\mathbb{R})$ est orthonormée pour ce produit scalaire. Pour toute forme linéaire $\varphi \in \mathcal{M}_n(\mathbb{R})^*$ et toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, on a :

$$\varphi(A) = \varphi \left(\sum_{1 \leq i, j \leq n} a_{ij} E_{i,j} \right) = \sum_{1 \leq i, j \leq n} a_{ij} \varphi(E_{i,j}) = \langle X | C \rangle$$

où $C \in \mathcal{M}_n(\mathbb{R})$ est défini par $C = \sum_{1 \leq i, j \leq n} \varphi(E_{i,j}) E_{i,j}$. On a donc ainsi montré que l'application $C \mapsto (\varphi_C : A \mapsto \langle A | C \rangle)$ est surjective de $\mathcal{M}_n(\mathbb{R})$ sur

$\mathcal{M}_n(\mathbb{R})^*$. Cette application étant linéaire et les espaces vectoriels $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{M}_n(\mathbb{R})^*$ étant de même dimension, on en déduit que c'est un isomorphisme. C'est-à-dire que, pour toute forme linéaire $\varphi \in \mathcal{M}_n(\mathbb{R})^*$, il existe une unique matrice $C \in \mathcal{M}_n(\mathbb{R})$ telle que :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \varphi(A) = \text{Tr}(A {}^t C)$$

Ce dernier résultat est vrai sur tout corps commutatif $\mathbb{K}$.

3. La fonction $\det$ est continûment dérivable comme fonction polynomiale des coefficients, donc différentiable. Notons u sa différentielle. Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, $u(A)$ est une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$, donc il existe une unique matrice $C \in \mathcal{M}_n(\mathbb{R})$ telle que $u(A)(H) = \text{Tr}(H {}^t C)$ pour tout H dans $\mathcal{M}_n(\mathbb{R})$. Pour $H = E_{i,j}$ ($1 \leq i, j \leq n$), on a $\text{Tr}(H {}^t C) = c_{ij}$, ce qui nous donne $C = ((u(A)(E_{i,j})))_{1 \leq i, j \leq n}$. Il s'agit donc de calculer les $u(A)(E_{i,j})$, pour

$1 \leq i, j \leq n$. La différentielle de $\det$ en A s'écrit $u(A) = \sum_{1 \leq i, j \leq n} \frac{\partial \det}{\partial x_{ij}}(A) dx_{ij}$,

où on a noté $dx_{ij} = E_{i,j}^*$ (base duale de la base canonique de $\mathcal{M}_n(\mathbb{R})$). On a donc :

$$u(A)(E_{i,j}) = \frac{\partial \det}{\partial x_{ij}}(A) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{\det(A + tE_{i,j}) - \det(A)}{t}$$

En notant A_j la colonne j de la matrice A , $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{R}^n$ et en utilisant la n -linéarité du déterminant, on a :

$$\begin{aligned} \det(A + tE_{i,j}) &= \det(A_1, \dots, A_{j-1}, A_j + te_i, A_{j+1}, \dots, A_n) \\ &= \det(A) + t \det(A_1, \dots, A_{j-1}, e_i, A_{j+1}, \dots, A_n) \end{aligned}$$

et $c_{ij} = \frac{\partial \det}{\partial x_{ij}}(A) = \det(A_1, \dots, A_{j-1}, e_i, A_{j+1}, \dots, A_n)$ (cofacteur (i, j) de A). On a donc :

$$\forall H \in \mathcal{M}_n(\mathbb{R}), u(A)(H) = \text{Tr}(H {}^t C)$$

où C est la matrice des cofacteurs de la matrice A .

4. La fonction f est dérivable sur I comme composée de fonctions dérivables, avec :

$$\forall t \in I, f'(t) = u(Y)(Y'(t)) = \sum_{1 \leq i, j \leq n} \frac{\partial \det}{\partial x_{ij}}(Y)(t) y'_{ij}(t)$$

C'est-à-dire que :

$$\begin{aligned} f' &= \sum_{1 \leq i, j \leq n} \det(Y_1, \dots, Y_{j-1}, e_i, Y_{j+1}, \dots, Y_n) y'_{ij} \\ &= \sum_{j=1}^n \det \left(Y_1, \dots, Y_{j-1}, \sum_{i=1}^n y'_{ij} e_i, Y_{j+1}, \dots, Y_n \right) \\ &= \sum_{j=1}^n \det(Y_1, \dots, Y_{j-1}, Y'_j, \dots, Y_n) \end{aligned}$$

Exercice 2.6. Soit $A \in \mathcal{M}_n(\mathbb{R})$. On appelle sous-matrices principales de A les matrices :

$$A_k = ((a_{ij}))_{1 \leq i, j \leq k} \quad (k = 1, \dots, n)$$

et déterminants principaux les réels $\Delta_k = \det(A_k)$. Montrer qu'une matrice symétrique réelle est définie positive si, et seulement si, tous ses déterminants principaux sont strictement positifs.

Solution. Si la matrice A est symétrique définie positive, il en est alors de même de toutes ses sous-matrices principales. En effet, pour tout entier k compris entre 1 et n , si x est un vecteur non nul de $\mathbb{R}^k$, le vecteur $x' = \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{R}^n$ est non nul et $\langle A_k x \mid x \rangle = \langle A x' \mid x' \rangle > 0$. On déduit que toutes les valeurs propres de la matrice A_k sont strictement positives et $\det(A_k) > 0$ pour tout entier k compris entre 1 et n . Pour la réciproque on procède par récurrence sur $n \geq 1$. Pour $n = 1$ le résultat est évident. Supposons le acquis pour les matrices symétriques d'ordre $n - 1 \geq 1$ et soit A une matrice symétrique réelle d'ordre n ayant tous ses déterminants principaux strictement positifs. On désigne par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{R}^n$. Avec l'hypothèse de récurrence on déduit que la matrice symétrique A_{n-1} est définie positive et la restriction à l'espace vectoriel H engendré par $\{e_1, \dots, e_{n-1}\}$ de la forme quadratique $q : x \mapsto \langle Ax \mid x \rangle$ est définie positive. Il existe alors une base $(f_1, \dots, f_{n-1})$ de H orthonormée pour $q|_H$ et la matrice de q dans la base $(f_1, \dots, f_{n-1}, e_n)$ s'écrit :

$$A' = \begin{pmatrix} 1 & 0 & \cdots & \alpha_1 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & \alpha_{n-1} \\ \alpha_1 & \cdots & \alpha_{n-1} & \alpha_n \end{pmatrix}$$

En posant $f_n = e_n - \sum_{i=1}^{n-1} \alpha_i f_i$, le système $(f_1, \dots, f_{n-1}, f_n)$ est une base de $\mathbb{R}^n$ orthogonale pour q et la matrice de q dans cette base s'écrit :

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

Le déterminant de la matrice A'' est du même signe que celui de A , c'est-à-dire positif. En conséquence la forme quadratique q est définie positive et la matrice A symétrique définie positive.

Exercice 2.7. On se place dans $(\mathbb{R}^n, \langle \cdot | \cdot \rangle)$ euclidien, on note $\|\cdot\|$ la norme associée et on désigne par S_1 la sphère unité de $\mathbb{R}^n$ pour cette norme. Soit A une matrice réelle symétrique d'ordre n . On lui associe la forme quadratique q définie par :

$$\forall x \in \mathbb{R}^n, q(x) = \langle Ax | x \rangle$$

1. Montrer que l'application q est bornée sur S_1 et atteint ses bornes. On note $\lambda_0 = q(x_0) = \sup_{x \in S_1} q(x)$, où x_0 dans $\mathbb{R}^n$ est tel que $\|x_0\| = 1$.
2. On se donne un vecteur x_1 dans $(\mathbb{R}x_0)^\perp \cap S_1$ et pour tout réel $\theta \in [-\pi, \pi]$ on note $x_\theta = \cos(\theta)x_0 + \sin(\theta)x_1$. Utilisant les inégalités $q(x_\theta) \leq q(x_0)$ montrer que :

$$\forall \theta \in [-\pi, \pi], 2 \cos(\theta) \sin(\theta) \langle Ax_0 | x_1 \rangle \leq \sin^2(\theta) (q(x_0) - q(x_1))$$

puis en déduire que $\langle Ax_0 | x_1 \rangle = 0$.

3. Dédurre de ce qui précède que A se diagonalise dans une base orthonormée.

Solution.

1. L'application q est continue en tant qu'application polynomiale, elle est donc bornée et atteint ses bornes sur la sphère unité S_1 qui est compacte dans $\mathbb{R}^n$.
2. On a $x_\theta \in S_1$ pour tout $\theta \in [-\pi, \pi]$, donc :

$$q(x_\theta) = \cos^2(\theta) q(x_0) + 2 \cos(\theta) \sin(\theta) \langle Ax_0 | x_1 \rangle + \sin^2(\theta) q(x_1) \leq q(x_0)$$

ce qui entraîne que $2 \cos(\theta) \sin(\theta) \langle Ax_0 | x_1 \rangle \leq \sin^2(\theta) (q(x_0) - q(x_1))$. On en déduit que :

$$\begin{cases} \forall \theta \in]0, \pi[, 2 \cos(\theta) \langle Ax_0 | x_1 \rangle \leq \sin(\theta) (q(x_0) - q(x_1)) \\ \forall \theta \in]-\pi, 0[, 2 \cos(\theta) \langle Ax_0 | x_1 \rangle \geq \sin(\theta) (q(x_0) - q(x_1)) \end{cases}$$

En faisant tendre θ vers 0 par valeurs positives et négatives, on en déduit que $\langle Ax_0 | x_1 \rangle = 0$, c'est-à-dire que $x_1 \in \{Ax_0\}^\perp$.

3. On déduit de ce qui précède que pour tout x appartenant à $\{x_0\}^\perp \setminus \{0\}$, on a x qui appartient à $\{Ax_0\}^\perp$ (il suffit de prendre $x_1 = \frac{1}{\|x\|}x$), c'est-à-dire que $\{x_0\}^\perp \subset \{Ax_0\}^\perp$. Il en résulte que $\mathbb{R}(Ax_0) = \{Ax_0\}^{\perp\perp} \subset \{x_0\}^{\perp\perp} = \mathbb{R}x_0$ et en particulier, il existe $\lambda \in \mathbb{R}$ tel que $Ax_0 = \lambda x_0$, ce qui signifie que x_0 est un vecteur propre de A . La matrice A étant symétrique, l'orthogonal de x_0 dans $\mathbb{R}^n$ est stable par A et il suffit d'appliquer le raisonnement précédent à cet orthogonal pour conclure par récurrence.

Exercice 2.8. Soit $A = (C_1, \dots, C_n) \in \mathcal{M}_n(\mathbb{C})$ où C_k désigne la colonne k de A . Montrer que $|\det(A)| \leq \prod_{k=1}^n \|C_k\|$ (inégalité d'Hadamard), en notant $\|\cdot\|$ la norme hermitienne sur $\mathbb{C}^n$ (on peut utiliser le procédé d'orthogonalisation de Gram-Schmidt).

Solution. Pour $\det(A) = 0$, le résultat est évident. On suppose que $\det(A) \neq 0$. Dans ce cas le système $(C_1, \dots, C_n)$ est une base de $\mathbb{C}^n$ et le procédé de Gram-Schmidt nous permet alors de construire une base orthogonale $(E_1, \dots, E_n)$ de $\mathbb{C}^n$ telle que :

$$E_1 = C_1; \forall k \in \{2, \dots, n\}, E_k = C_k + \sum_{j=1}^{k-1} \lambda_{j,k} E_j$$

Un déterminant étant inchangé si on ajoute à une colonne une combinaison linéaire des autres, en notant $N = (E_1, \dots, E_n) \in \mathcal{M}_n(\mathbb{C})$, on a $\det(A) = \det(N)$. En notant $D = N^* N$, on a :

$$D = \begin{pmatrix} \|E_1\|^2 & 0 & \cdots & 0 \\ 0 & \|E_2\|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \|E_n\|^2 \end{pmatrix}$$

donc $|\det(A)|^2 = |\det(N)|^2 = \det(D) = \prod_{k=1}^n \|E_k\|^2$, soit $|\det(A)| = \prod_{k=1}^n \|E_k\|$.

Enfin avec :

$$\|C_k\|^2 = \|E_k\|^2 + \sum_{j=1}^{k-1} |\lambda_{j,k}|^2 \|E_j\|^2 \geq \|E_k\|^2$$

on déduit que $|\det(A)| \leq \prod_{k=1}^n \|C_k\|$.

Exercice 2.9. Montrer qu'une matrice A d'ordre $n > 0$ à coefficients complexes est normale si, et seulement si, il existe un polynôme P à coefficients complexes tel que $A^* = P(A)$.

Solution. Une matrice normale se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe une matrice unitaire U et une matrice diagonale D telles que $D = U^* A U$. En notant $\lambda_1, \lambda_2, \dots, \lambda_p$ les éléments diagonaux deux à deux distincts de D (ce sont les valeurs propres de A) avec $1 \leq p \leq n$, on désigne par P le polynôme d'interpolation de Lagrange défini par :

$$P \in \mathbb{C}_{p-1}[X], P(\lambda_k) = \overline{\lambda_k} \quad (1 \leq k \leq p)$$

on a $D^* = P(D) = U^* P(A) U$ et $P(A) = U D^* U^* = (U D U^*)^* = A^*$. La réciproque est évidente.

Exercice 2.10. Soit A une matrice d'ordre $n \geq 1$ à coefficients complexes.

1. Montrer que A est la matrice nulle si, et seulement si, $\text{Tr}(AA^*) = 0$.
2. Montrer que la matrice A est hermitienne si, et seulement si, $AA^* = A^2$.

Solution.

1. Pour $A = ((a_{ij}))_{1 \leq i, j \leq n}$, on a $\text{Tr}(AA^*) = \sum_{1 \leq i, j \leq n} |a_{ij}|^2$ (carré de la norme hermitienne de $\mathbb{C}^{n^2}$) et $\text{Tr}(AA^*) = 0$ équivaut bien à $A = 0$.
2. Soit $A \in \mathcal{M}_n(\mathbb{C})$ telle que $AA^* = A^2$. Notant $B = A - A^*$, on a $BB^* = A^*A - (A^*)^2$ et :

$$\begin{aligned}\text{Tr}(BB^*) &= \text{Tr}(A^*A) - \text{Tr}((A^*)^2) \\ &= \text{Tr}(AA^*) - \text{Tr}((A^*)^2) = \text{Tr}(AA^* - (A^*)^2)\end{aligned}$$

avec $AA^* = (AA^*)^* = (A^2)^* = (A^*)^2$. On déduit donc que $\text{Tr}(BB^*) = 0$ et $B = 0$, c'est-à-dire $A = A^*$. La réciproque est évidente.

Exercice 2.11. Soit $A \in GL_n(\mathbb{C})$, montrer qu'il existe deux matrices unitaires U, V et une matrice diagonale D à coefficients réels strictement positifs telles que $A = UDV^*$ (décomposition singulière de la matrice A).

Solution. La matrice A^*A est hermitienne définie positive (en effet, pour x non nul on a $\langle A^*Ax \mid x \rangle = \|Ax\|_2^2 > 0$), donc il existe une matrice unitaire V telle que $V^*(A^*A)V = \Delta$ où :

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

avec $\lambda_i \in \mathbb{R}_+^*$ pour tout $i \in \{1, \dots, n\}$ (la matrice A est inversible). En notant $W = AV$, on a alors $W^*W = D^2$, où :

$$D = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\lambda_n} \end{pmatrix}$$

et avec $U^*U = D^{-1}W^*WD^{-1} = D^{-1}D^2D^{-1} = I_n$, on déduit que $U = WD^{-1}$ est unitaire. Ce qui donne en définitive $U^*AV = D^{-1}W^*W = D^{-1}D^2 = D$, avec D diagonale à coefficients réels strictement positifs.

Exercice 2.12. Soient A et B deux matrices hermitiennes positives dans $\mathcal{M}_n(\mathbb{C})$. Montrer que :

$$0 \leq \operatorname{Tr}(AB) \leq \operatorname{Tr}(A) \operatorname{Tr}(B)$$

(on peut d'abord considérer le cas où la matrice A est diagonale).

Solution. Si $(e_i)_{1 \leq i \leq n}$ désigne la base canonique de $\mathbb{C}^n$, on a alors, pour toute matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{C})$, $a_{ij} = \langle Ae_j | e_i \rangle$ pour $1 \leq i, j \leq n$. Dans le cas particulier où A est hermitienne positive on a $a_{ii} = \langle Ae_i | e_i \rangle \geq 0$ pour $1 \leq i \leq n$. Si A est diagonale hermitienne positive, on a alors :

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

avec $\lambda_i \in \mathbb{R}^+$ pour tout i et pour toute matrice $B \in \mathcal{M}_n(\mathbb{C})$ hermitienne positive, on a $b_{ii} \geq 0$ pour tout i et :

$$0 \leq \operatorname{Tr}(AB) = \sum_{i=1}^n \lambda_i b_{ii} \leq \left(\sum_{i=1}^n \lambda_i \right) \left(\sum_{i=1}^n b_{ii} \right) = \operatorname{Tr}(A) \operatorname{Tr}(B)$$

Dans le cas général, on sait que, si A est hermitienne positive, toutes ses valeurs propres sont réelles positives et il existe une matrice unitaire U telle que :

$$A = U \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} U^* = U D U^*$$

avec $\lambda_i \in \mathbb{R}^+$ pour tout i . En remarquant que $AB = U D U^* B = U (D U^* B U) U^*$ est semblable à $D (U^* B U) = D C$, on a $\operatorname{Tr}(AB) = \operatorname{Tr}(DC)$. Mais D est diagonale hermitienne positive et C hermitienne positive unitairement semblable à B . Donc :

$$\operatorname{Tr}(AB) = \operatorname{Tr}(DC) \leq \operatorname{Tr}(D) \operatorname{Tr}(C) = \operatorname{Tr}(A) \operatorname{Tr}(B)$$

Exercice 2.13. Montrer que l'application $A \mapsto \|A\|_s = \sqrt{\operatorname{Tr}(A^* A)}$ est norme sur $\mathcal{M}_n(\mathbb{C})$ (norme de Schur) et qu'elle est sous-multiplicative, c'est-à-dire que $\|AB\|_s \leq \|A\|_s \|B\|_s$ pour toutes matrices A et B appartenant à $\mathcal{M}_n(\mathbb{C})$.

Solution. Pour tout $A \in \mathcal{M}_n(\mathbb{C})$, on a $\|A\|_s = \sqrt{\operatorname{Tr}(A^* A)} = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$, c'est-à-dire que $\|\cdot\|_s$ est la norme hermitienne de $\mathcal{M}_n(\mathbb{C})$ identifié à $\mathbb{C}^{n^2}$. Pour

toutes matrices A, B dans $\mathcal{M}_n(\mathbb{C})$, on a :

$$\|AB\|_s^2 = \text{Tr}(B^*(A^*AB)) = \text{Tr}((A^*AB)B^*) = \text{Tr}((A^*A)(BB^*))$$

et les matrices A^*A et BB^* étant hermitiennes positives, on peut utiliser le résultat de l'exercice 2.12 pour écrire que :

$$\|AB\|_s^2 \leq \text{Tr}(A^*A) \text{Tr}(BB^*) = \|A\|_s^2 \|B\|_s^2 = \|A\|_s^2 \|B\|_s^2$$

ce qui prouve que la norme de Schur est sous-multiplicative.

On peut aussi montrer ce résultat directement en écrivant que :

$$\|AB\|_s^2 = \sum_{i,j=1}^n \left| \sum_{k=1}^n a_{ik} b_{kj} \right|^2 \leq \sum_{i,j=1}^n \left(\sum_{k=1}^n |a_{ik}| |b_{kj}| \right)^2$$

et en utilisant l'inégalité de Cauchy-Schwarz :

$$\|AB\|_s^2 \leq \sum_{i,j=1}^n \left(\sum_{k=1}^n |a_{ik}|^2 \right) \left(\sum_{k=1}^n |b_{kj}|^2 \right) = \|A\|_s^2 \|B\|_s^2$$

Exercice 2.14. Montrer que toute matrice à coefficients réels ou complexes est semblable à sa transposée (utiliser la décomposition de Jordan sur $\mathbb{C}$).

Solution. On se place tout d'abord dans le cas où A est un bloc de Jordan complexe de la forme :

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

L'endomorphisme u associé à la matrice A est alors défini dans la base canonique $(e_i)_{1 \leq i \leq n}$ de $\mathbb{C}^n$ par :

$$\begin{cases} u(e_1) = \lambda e_1 \\ u(e_j) = e_{j-1} + \lambda e_j \quad (2 \leq j \leq n) \end{cases}$$

Dans la base $(e_n, \dots, e_1)$ l'endomorphisme u a pour matrice ${}^t A$. Si A est une matrice complexe quelconque, on sait qu'elle est semblable à une matrice de la forme :

$$A' = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix}$$

où les J_k sont des blocs de Jordan du type précédent. Chacun de ces J_k étant semblable à sa transposée on en déduit facilement que A est semblable à sa transposée. Ce résultat est encore valable pour les matrices réelles du fait que deux matrices réelles qui sont semblables sur $\mathbb{C}$ le sont sur $\mathbb{R}$ (exercice 1.1).

Exercice 2.15. Avec cet exercice, on propose une démonstration du théorème 2.20 sans utiliser la dualité. On se donne $u \in \mathcal{L}(E)$ nilpotente d'ordre r compris entre 2 et $n-1$ avec $n \geq 2$ (pour $r=1$, u est nulle et pour $r=n$, c'est fait avec le lemme 2.8).

1. Soit $(B, C) \in \mathcal{M}_{r, n-r}(\mathbb{K}) \times \mathcal{M}_{n-r}(\mathbb{K})$ tel que $A = \begin{pmatrix} J_r & B \\ 0_{n-r, r} & C \end{pmatrix}$ soit nilpotente d'ordre r . Montrer que B est nulle si, et seulement si, ses $r-1$ dernières lignes sont nulles.
2. En utilisant le lemme 2.8, justifier l'existence d'une base de E dans laquelle la matrice A de u est de la forme $\begin{pmatrix} J_r & B \\ 0_{n-r, r} & C \end{pmatrix}$.
3. En notant, pour $m = n-r$, toute matrice $X \in \mathcal{M}_{r, m}(\mathbb{K})$ et tout entier i compris entre 1 et r , $X_i = (x_{i1}, x_{i2}, \dots, x_{im})$ sa ligne numéro i , donner pour tout entier k compris entre 1 et $r-1$ une expression du produit $J_r^k X$ en fonction des X_i .
4. En notant, pour toute matrice $X \in \mathcal{M}_{r, n-r}(\mathbb{K})$, $T_X = \begin{pmatrix} I_r & X \\ 0 & I_{n-r} \end{pmatrix}$, montrer qu'il existe une matrice $X \in \mathcal{M}_{r, n-r}(\mathbb{K})$ telle que $T_X A (T_X)^{-1}$ soit de la forme $\begin{pmatrix} J_r & 0 \\ 0 & C \end{pmatrix}$.
5. Conclure.

Solution.

1. Soit $v \in \mathcal{L}(\mathbb{K}^n)$ de matrice A dans la base canonique $(e_k)_{1 \leq k \leq n}$. Dire que les $r-1$ dernières lignes de B sont nulles se traduit par $v(e_j) = b_{1j}e_1 + c_j$ pour $r+1 \leq j \leq n$, où $c_j \in G = \text{Vect}\{e_{r+1}, \dots, e_n\}$. Comme A est nilpotente d'ordre r , on a :

$$0 = v^r(e_j) = b_{1j}v^{r-1}(e_1) + v^{r-1}(c_j) = b_{1j}e_r + v^{r-1}(c_j) \quad (2.3)$$

avec $v(e_j) \in \text{Vect}\{e_1, e_{r+1}, \dots, e_n\}$ et $v^k(e_j) \in \text{Vect}\{e_1, \dots, e_k, e_{r+1}, \dots, e_n\}$ pour $1 \leq k \leq r-1$ (récurrence immédiate), donc $v^{r-1}(c_j)$ appartient à $\text{Vect}\{e_1, \dots, e_{r-1}, e_{r+1}, \dots, e_n\}$ et l'égalité (2.3) impose $b_{1j} = 0$ pour tout j , ce qui signifie que la première ligne de B est nulle et en conséquence, $B = 0$. La réciproque est évidente.

2. En complétant la famille libre $\mathcal{B}_x = (u^k(x))_{0 \leq k \leq r-1}$ en une base de E , la matrice de u dans cette base a la forme indiquée puisque $\text{Vect}(\mathcal{B}_x)$ est stable par u .

3. En notant $(e_j)_{1 \leq j \leq m}$ la base canonique de $\mathbb{K}^m$, on a pour tout entier j compris entre 1 et m , $J_r^k X e_j = J_r^k \left(\sum_{i=1}^r x_{ij} e_i \right) = \sum_{i=1}^r x_{ij} J_r^k e_i = \sum_{i=1}^{r-k} x_{ij} e_{i+k}$, ce qui

signifie que la colonne numéro j de $J_r^k X$ est $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_{1j} \\ \vdots \\ x_{r-k,j} \end{pmatrix}$, le coefficient x_{1j}

étant en ligne $k+1$, donc $J_r^k X = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ X_1 \\ \vdots \\ X_{r-k} \end{pmatrix}$, la ligne X_1 étant en position

$k+1$.

4. Pour tout $X \in \mathcal{M}_{r,n-r}(\mathbb{K})$, la matrice T_X est inversible d'inverse $(T_X)^{-1} = \begin{pmatrix} I_r & -X \\ 0 & I_{n-r} \end{pmatrix}$ et on a :

$$T_X A (T_X)^{-1} = \begin{pmatrix} I_r & X \\ 0 & I_{n-r} \end{pmatrix} \begin{pmatrix} J_r & B \\ 0 & C \end{pmatrix} \begin{pmatrix} I_r & -X \\ 0 & I_{n-r} \end{pmatrix} = \begin{pmatrix} J_r & B(X) \\ 0 & C \end{pmatrix}$$

où $B(X) = XC - J_r X + B$. Il s'agit alors de montrer que le système linéaire $B(X) = 0$ a une solution dans $\mathcal{M}_{r,n-r}(\mathbb{K})$. En notant M_i la ligne numéro i d'une matrice M , on a $B_1(X) = X_1 C + B_1$ et $B_i(X) = X_i C - X_{i-1} + B_i$ pour $2 \leq i \leq r$. En définissant la matrice X par :

$$X_r = 0, X_{r-1} = B_r, X_{r-2} = X_{r-1}C + B_{r-1}, \dots, X_1 = X_2C + B_2$$

la matrice $B(X)$ a ses dernières lignes nulles avec $\begin{pmatrix} J_r & B(X) \\ 0 & C \end{pmatrix}$ nilpotente d'ordre r (elle est semblable à la matrice A), ce qui revient à dire qu'elle est nulle.

5. Pour retrouver le théorème 2.20, il revient au même de prouver que toute matrice nilpotente $A \in \mathcal{M}_n(\mathbb{K}) \setminus \{0\}$ est semblable à une matrice de la forme

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix}. \text{ Pour ce faire, on procède par récurrence sur } n \geq 1.$$

Pour $n = 1$ c'est clair. Supposons le résultat acquis jusqu'au rang $n-1 \geq 1$ et soit $A \in \mathcal{M}_n(\mathbb{K}) \setminus \{0\}$ nilpotente d'ordre $r \geq 2$. Pour $r = n$, on déduit du lemme 2.8 que A est semblable à J_n . Pour $2 \leq r \leq n-1$, la question précédente nous dit que A est semblable à une matrice de la forme $A_1 = \begin{pmatrix} J_r & 0 \\ 0 & C \end{pmatrix}$. Avec

$0 = A_1^r = \begin{pmatrix} J_r^r & 0 \\ 0 & C^r \end{pmatrix}$, on déduit que C est nilpotente d'ordre $r_1 \leq r < n$ et l'hypothèse de récurrence nous permet de conclure.

Exercice 2.16. Avec cet exercice, on propose une démonstration du théorème 2.20 sans utiliser la dualité. Précisément, on se propose de démontrer que pour tout endomorphisme nilpotent $u \in \mathcal{L}(E)$ d'ordre $r \geq 1$, il existe un entier p compris entre 1 et n et des sous espaces $E_1, \dots, E_p$ de E tels que pour tout i compris entre 1 et p , l'espace E_i est de dimension $n_i \geq 1$, stable par u , la restriction u_i de u à E_i est nilpotente d'ordre n_i , la suite $(n_i)_{1 \leq i \leq p}$ est décroissante avec $n_1 = r$ et $E = \bigoplus_{i=1}^p E_i$ (décomposition de E en sous-espaces cycliques). Pour $n = 1$, le résultat est évident ($u = 0$ et $E_1 = E$). Le supposant acquis pour les espaces de dimension inférieure à $n - 1$, on se donne un espace E de dimension $n \geq 1$ et un endomorphisme $u \in \mathcal{L}(E)$ nilpotent d'ordre r compris entre 1 et n . Comme pour $r = 1$ ($u = 0$) ou $r = n$ ($p = 1$, $E_1 = E$ et $n_1 = n$) il n'y a rien à montrer, on suppose que r est compris entre 2 et $n - 1$.

1. Montrer que $F = \text{Im}(u)$ est un sous-espace vectoriel de E stable par u , de dimension comprise entre 1 et $n - 1$ et que la restriction de u à F est nilpotente d'ordre $r_1 \leq r - 1$.

Par hypothèse de récurrence appliquée à la restriction de u à $\text{Im}(u)$, on peut donc écrire $\text{Im}(u) = \bigoplus_{i=1}^p F_i$, où chaque F_i est un sous-espace de $\text{Im}(u)$ de dimension $n_i \geq 1$, stable par u tel que $u|_{F_i}$ est nilpotent d'ordre n_i , la suite $(n_i)_{1 \leq i \leq p}$ étant décroissante avec $n_1 = r_1 \leq r - 1$.

2. Montrer que, pour tout i compris entre 1 et p , il existe un vecteur $x_i \in E$ tel que la famille $\mathcal{B}_i = (u(x_i), u^2(x_i), \dots, u^{n_i}(x_i))$ soit une base de F_i et la famille $\{x_i\} \cup \mathcal{B}_i$ est libre.

3. En notant $G_i = \bigoplus_{k=0}^{n_i} \mathbb{K}u^k(x_i)$ pour tout entier i compris entre 1 et p et

$$G = \sum_{i=1}^p G_i, \text{ montrer que :}$$

$$(a) \quad G_i \cap \ker(u) = \mathbb{K}u^{n_i}(x_i);$$

$$(b) \quad G \cap \ker(u) = \bigoplus_{i=1}^p \mathbb{K}u^{n_i}(x_i);$$

$$(c) \quad G = \bigoplus_{i=1}^p G_i;$$

$$(d) \quad \dim(G) = \dim(\text{Im}(u)) + p.$$

4. En désignant par H un supplémentaire de $G \cap \ker(u)$ dans $\ker(u)$, montrer que $E = G \oplus H$ et conclure.

Solution.

1. Comme u est non nul et nilpotent, on a $\dim(\operatorname{Im}(u)) \geq 1$ et $\dim(\ker(u)) \geq 1$ puisque 0 est valeur propre de u , donc $\dim(\operatorname{Im}(u)) = n - \dim(\ker(u)) \leq n - 1$. Il est bien connu que $\operatorname{Im}(u)$ est stable par u .
2. L'endomorphisme $u_i = u|_{F_i} \in \mathcal{L}(F_i)$ étant nilpotent d'ordre $n_i = \dim(F_i)$, on a vu avec le lemme 2.8 qu'il existe une famille libre de la forme $\mathcal{B}_i = (u^k(y_i))_{0 \leq k \leq n_i-1}$ dans F_i et cette famille en est une base puisque formée de n_i vecteurs. Comme $y_i \in F_i \subset \operatorname{Im}(u)$, il existe $x_i \in E$ tel que $y_i = u(x_i)$, donc $\mathcal{B}_i = (u^k(x_i))_{1 \leq k \leq n_i}$. Si $x = \sum_{k=0}^{n_i} \lambda_k u^k(x_i) = 0$, on a alors $0 = u(x) = \sum_{k=0}^{n_i} \lambda_k u^{k+1}(x_i) = \sum_{k=0}^{n_i-1} \lambda_k u^k(y_i)$ (on a $u^{n_i}(y_i) = 0$ car $y_i \in F_i$ et $u_i^{n_i} = 0$), ce qui équivaut à la nullité de tous les λ_i car $(u^k(y_i))_{0 \leq k \leq n_i-1}$ est une base de F_i . La famille $\{x_i\} \cup \mathcal{B}_i$ est donc libre.

3.

- (a) Chaque $u^{n_i}(x_i)$ est dans G_i et $u(u^{n_i}(x_i)) = u^{n_i}(y_i) = 0$, donc $u^{n_i}(x_i)$ est dans $\ker(u)$. Un élément de $G_i \cap \ker(u)$ s'écrit $x = \sum_{k=0}^{n_i} \lambda_k u^k(x_i)$ et

$$\text{on a } 0 = u(x) = \sum_{k=0}^{n_i} \lambda_k u^{k+1}(x_i) = \sum_{k=0}^{n_i-1} \lambda_k u^{k+1}(x_i), \text{ donc } \lambda_k = 0 \text{ pour}$$

$0 \leq k \leq n_i - 1$ et $x = \lambda_{n_i} u^{n_i}(x_i)$ avec $u^{n_i}(x_i) = u^{n_i-1}(y_i) \neq 0$, ce qui signifie que $G_i \cap \ker(u)$ est la droite vectorielle dirigée par $u^{n_i}(x_i)$.

- (b) Un élément de $G \cap \ker(u)$ s'écrit $z = \sum_{i=1}^p z_i$ avec $z_i \in G_i$ et on a $0 = u(z) =$

$$\sum_{i=1}^p u(z_i) \text{ avec } u(z_i) \in u(G_i) \text{ et :}$$

$$u(G_i) = \operatorname{Vect}\{u(x_i), \dots, u^{n_i+1}(x_i)\} = \operatorname{Vect}\{u(x_i), \dots, u^{n_i}(x_i)\} = F_i$$

les F_i étant en somme directe, ce qui impose la nullité de tous les $u(z_i)$, donc $z_i \in G_i \cap \ker(u)$, soit $z_i = \lambda_i u^{n_i}(x_i)$. La famille $(u^{n_i}(x_i))_{1 \leq i \leq p}$ engendre donc $G \cap \ker(u)$ et elle est libre puisque chaque $u^{n_i}(x_i) = u^{n_i-1}(y_i)$ est non nul dans F_i , les F_i étant en somme directe.

- (c) Si $\sum_{i=1}^p z_i = 0$ avec $z_i \in G_i$ pour tout i , on a alors $\sum_{i=1}^p u(z_i) = 0$ avec $u(z_i)$

dans $u(G_i) = F_i$, les F_i étant en somme directe, donc $u(z_i) = 0$ pour tout i , c'est-à-dire que $z_i \in G_i \cap \ker(u) = \operatorname{Vect}\{u^{n_i}(x_i)\}$, soit $z_i = \lambda_i u^{n_i}(x_i)$.

On a donc $\sum_{i=1}^p \lambda_i u^{n_i}(x_i) = 0$, la famille $(u^{n_i}(x_i))_{1 \leq i \leq p}$ étant libre, ce qui impose $\lambda_i = 0$, soit $z_i = 0$ pour tout i .

(d) Il en résulte que :

$$\dim(G) = \sum_{i=1}^p \dim(G_i) = \sum_{i=1}^p (n_i + 1) = \sum_{i=1}^p \dim(F_i) + p = \dim(\operatorname{Im}(u)) + p$$

4. On a $\ker(u) = (G \cap \ker(u)) \oplus H$, donc :

$$\begin{aligned} \dim(H) &= \dim(\ker(u)) - \dim(G \cap \ker(u)) \\ &= n - (\dim(\operatorname{Im}(u)) + p) = n - \dim(G) \end{aligned}$$

Pour tout $x \in G \cap H$, on a $x \in H \subset \ker(u)$ et $x \in G$, donc $x \in (G \cap \ker(u)) \cap H$,

soit $x = 0$. En conclusion, on a $E = \bigoplus_{i=1}^p G_i \oplus H$, chaque espace G_i ayant une

base de la forme $\mathcal{B}'_i = (u^k(x_i))_{0 \leq k \leq n_i}$ et $H \subset \ker(u)$. En écrivant $H = \bigoplus_{i=r+1}^q H_i$

avec $H_i = \mathbb{K}h_i$, on a bien une décomposition de E de la forme annoncée et la matrice de u dans une base adaptée à cette somme directe est bien diagonale par blocs de Jordan.

Chapitre 3

L'espace vectoriel normé

$\mathcal{M}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)

Pour ce chapitre, $\mathbb{K}$ est le corps des réels ou des complexes.

3.1 Norme matricielle induite par une norme vectorielle

Pour ce paragraphe on se donne une norme $x \mapsto \|x\|$ sur $\mathbb{K}^n$. On note respectivement, $B = \{x \in \mathbb{K}^n \mid \|x\| \leq 1\}$ la boule unité et $S = \{x \in \mathbb{K}^n \mid \|x\| = 1\}$ la sphère unité de $(\mathbb{K}^n, \|\cdot\|)$.

On rappelle que sur un espace vectoriel normé une application linéaire est continue si, et seulement si, elle est bornée sur la sphère unité et qu'en dimension finie la sphère unité est compacte et tout endomorphisme est continu (voir [18]). On peut donc définir, pour toute matrice A , la quantité $\|A\| = \sup_{x \in S} \|Ax\|$.

Théorème 3.1.

Pour toute norme $x \mapsto \|x\|$ sur $\mathbb{K}^n$, l'application $A \mapsto \|A\| = \sup_{x \in S} \|Ax\|$ définit une norme sur $\mathcal{M}_n(\mathbb{K})$.

Preuve. L'égalité $\|A\| = 0$ équivaut à $\|Ax\| = 0$ pour tout x dans S . En remarquant que pour tout x dans $\mathbb{K}^n$ on a $\frac{1}{\|x\|}x \in S$ et en utilisant la linéarité de A , on déduit que $A = 0$. La vérification des autres propriétés d'une norme ne pose pas de problèmes particuliers. $\square$

Définition 3.1. *L'application $A \mapsto \|A\| = \sup_{x \in S} \|Ax\|$ est la norme matricielle induite (ou subordonnée à) la norme vectorielle $x \mapsto \|x\|$.*

L'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ étant de dimension finie, on sait que toutes les normes sur cet espace sont équivalentes, c'est-à-dire qu'on a une seule topologie sur cet espace compatible avec la structure d'espace vectoriel.

La norme matricielle subordonnée à la norme euclidienne $\|\cdot\|_{2,\mathbb{R}}$ sur $\mathbb{R}^n$ est définie par :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \|A\|_{2,\mathbb{R}} = \sup_{\|x\|_{2,\mathbb{R}}=1} \|Ax\|_{2,\mathbb{R}}$$

et *a priori* cette norme pourrait être différente de celle induite par la norme hermitienne $\|\cdot\|_{2,\mathbb{C}}$ de $\mathbb{C}^n$ définie par :

$$\|A\|_{2,\mathbb{C}} = \sup_{\|z\|_{2,\mathbb{C}}=1} \|Az\|_{2,\mathbb{C}}$$

où $\|z\|_{2,\mathbb{C}}^2 = \|x\|_{2,\mathbb{R}}^2 + \|y\|_{2,\mathbb{R}}^2$ pour tout $z = x + iy \in \mathbb{C}^n$. En fait, il n'en est rien. En effet pour tout $x \in \mathbb{R}^n$, on a $\|x\|_{2,\mathbb{C}} = \|x\|_{2,\mathbb{R}}$ et $\|Ax\|_{2,\mathbb{C}} = \|Ax\|_{2,\mathbb{R}}$, donc $\|A\|_{2,\mathbb{R}} \leq \|A\|_{2,\mathbb{C}}$. Et pour $z = x + iy \in \mathbb{C}^n$, on a :

$$\|Az\|_{2,\mathbb{C}}^2 = \|Ax\|_{2,\mathbb{R}}^2 + \|Ay\|_{2,\mathbb{R}}^2 \leq \|A\|_{2,\mathbb{R}}^2 (\|x\|_{2,\mathbb{R}}^2 + \|y\|_{2,\mathbb{R}}^2) = \|A\|_{2,\mathbb{R}}^2 \|z\|_{2,\mathbb{C}}^2$$

donc $\|A\|_{2,\mathbb{C}} \leq \|A\|_{2,\mathbb{R}}$.

Les deux exemples qui suivent nous permettent de calculer effectivement les normes matricielles induites par les normes vectorielles $\|\cdot\|_\infty$ et $\|\cdot\|_1$.

Lemme 3.1 *La norme matricielle induite par $\|\cdot\|_\infty$ est définie par :*

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

Preuve. Le résultat étant évident pour $A = 0$, on suppose A non nulle. On note $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$ et S_∞ la sphère unité de $(\mathbb{K}^n, \|\cdot\|_\infty)$.

Pour tout $x \in S_\infty$, on a $\|Ax\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \left(\max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \right)$, donc

$\|A\|_\infty \leq \alpha = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$. Soient k un entier compris entre 1 et n tel que

$\alpha = \sum_{j=1}^n |a_{kj}|$ et $x = \sum_{j=1}^n x_j e_j \in S_\infty$ défini par :

$$x_j = \begin{cases} \frac{\overline{a_{kj}}}{|a_{kj}|} & \text{si } a_{kj} \neq 0 \\ 0 & \text{si } a_{kj} = 0 \end{cases}$$

(A étant non nulle, on a $\alpha > 0$ et il existe un indice j tel que $a_{kj} \neq 0$). En notant $y = Ax$, on a :

$$\forall i \in \{1, \dots, n\}, |y_i| \leq \|y\|_\infty \leq \|A\|_\infty \leq \alpha$$

avec $|y_k| = \left| \sum_{j=1}^n a_{kj} x_j \right| = \sum_{j=1}^n |a_{kj}| = \alpha$, donc $\|Ax\|_\infty = \alpha$ et $\|A\|_\infty = \alpha$. □

Lemme 3.2 *La norme matricielle induite par $\|\cdot\|_1$ est définie par :*

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

Preuve. Posons $\beta = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$. Pour x dans S_1 (sphère unité de $(\mathbb{K}^n, \|\cdot\|_1)$), on a :

$$\|Ax\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{j=1}^n |x_j| \sum_{i=1}^n |a_{ij}| \leq \beta \sum_{j=1}^n |x_j| = \beta$$

donc $\|A\|_1 \leq \beta$. Pour montrer que $\beta \leq \|A\|_1$, il suffit de trouver un vecteur x dans S_1 tel que $\|Ax\|_1 = \beta$. Soit k compris entre 1 et n tel que $\beta = \sum_{i=1}^n |a_{ik}|$. Pour $x = e_k$ ($k^{\text{ème}}$ vecteur de la base canonique de $\mathbb{K}^n$), on a :

$$\|Ax\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| = \sum_{i=1}^n |a_{ik}| = \beta$$

On a donc bien $\|A\|_1 = \beta$. □

On a $\|A\|_1 = \|A^*\|_\infty$, où $A^* = {}^t\bar{A}$ désigne la matrice adjointe de A .

Lemme 3.3 *Pour toute matrice P dans $GL_n(\mathbb{K})$ l'application :*

$$x \mapsto \|x\|_P = \|P^{-1}x\|_\infty$$

définit une norme sur $\mathbb{K}^n$ et la norme matricielle induite par cette norme est définie par :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \|A\|_P = \|P^{-1}AP\|_\infty$$

Preuve. Il est facile de vérifier que $\|\cdot\|_P$ définit une norme sur $\mathbb{K}^n$. On peut aussi remarquer que P est la matrice de passage de la base canonique $\mathcal{B} = (e_k)_{1 \leq k \leq n}$

de $\mathbb{K}^n$ à une base $\mathcal{B}' = (e'_k)_{1 \leq k \leq n}$ et en notant $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$, les coordonnées du vecteur x dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$, on a :

$$\|x\|_P = \|P^{-1}X\|_\infty = \|X'\|_\infty$$

Pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$, on a :

$$\|A\|_P = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|_P}{\|x\|_P} = \sup_{X \in \mathbb{K}^n \setminus \{0\}} \frac{\|P^{-1}AX\|_\infty}{\|P^{-1}X\|_\infty}$$

soit en posant $X = PX'$:

$$\|A\|_P = \sup_{X' \in \mathbb{K}^n \setminus \{0\}} \frac{\|P^{-1}APX'\|_\infty}{\|X'\|_\infty} = \|P^{-1}AP\|_\infty$$

□

Le théorème qui suit nous donne quelques propriétés des normes matricielles induites par une norme vectorielle.

Théorème 3.2.

Soit $A \mapsto \|A\|$ une norme sur $\mathcal{M}_n(\mathbb{K})$ induite par une norme $x \mapsto \|x\|$ sur $\mathbb{K}^n$. On a $\|I_n\| = 1$ et :

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \|A\| = \sup_{x \in B} \|Ax\| = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} \quad (3.1)$$

$$\forall x \in \mathbb{K}^n, \forall A \in \mathcal{M}_n(\mathbb{K}), \|Ax\| \leq \|A\| \|x\| \quad (3.2)$$

$$\forall A \in \mathcal{M}_n(\mathbb{K}), \forall B \in \mathcal{M}_n(\mathbb{K}), \|AB\| \leq \|A\| \|B\| \quad (3.3)$$

$$\exists x \in S \text{ tel que } \|A\| = \|Ax\|$$

$$\|A\| = \inf \{ \alpha \in \mathbb{R}^+ \mid \forall x \in \mathbb{K}^n, \|Ax\| \leq \alpha \|x\| \} \quad (3.4)$$

Preuve. Pour tout $x \in S$, on a $\|I_n x\| = \|x\| = 1$, donc $\|I_n\| = 1$.

L'égalité (3.1) résulte de l'inclusion $S \subset B$ et de la linéarité de A .

L'inégalité (3.2) se déduit facilement de (3.1).

De (3.2) on déduit que pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$ on a :

$$\forall x \in \mathbb{K}^n, \|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|$$

D'où l'inégalité (3.3).

Du fait qu'en dimension finie la sphère unité est compacte et qu'une fonction continue sur un compact admet une borne supérieure qui est atteinte, on déduit qu'il existe x dans S tel que $\|A\| = \|Ax\|$.

On note $\mathcal{D} = \{ \alpha \in \mathbb{R}^+ \mid \forall x \in \mathbb{K}^n, \|Ax\| \leq \alpha \|x\| \}$. On a $\mathcal{D} \neq \emptyset$, car $\|A\| \in \mathcal{D}$, donc $\mathcal{D}$ admet une borne inférieure comme partie non vide et minorée de $\mathbb{R}^+$ et $\inf(\mathcal{D}) \leq \|A\|$. Soient $\alpha \in \mathcal{D}$ et x dans S tel que $\|A\| = \|Ax\|$. On alors :

$$\|A\| = \|Ax\| \leq \alpha \|x\| = \alpha$$

soit $\|A\| \leq \alpha$. On en déduit donc que $\|A\| \leq \inf(\mathcal{D})$ et $\|A\| = \inf(\mathcal{D})$. □

La propriété (3.3) se traduit en disant que toute norme matricielle induite par une norme vectorielle est sous-multiplicative.

En fait pour toute norme $A \mapsto N(A)$ sur $\mathcal{M}_n(\mathbb{K})$, il existe une constante réelle $\lambda > 0$ telle que :

$$\forall (A, B) \in \mathcal{M}_n(\mathbb{K}) \times \mathcal{M}_n(\mathbb{K}), N(AB) \leq \lambda N(A) N(B)$$

(exercice 3.1).

Il existe quand même des normes sur $\mathcal{M}_n(\mathbb{K})$ qui ne sont pas sous-multiplicatives (exercice 3.2).

Une norme sous-multiplicative n'est pas nécessairement induite par une norme vectorielle. Par exemple la norme de Schur est sous-multiplicative (exercice 2.13) et n'est pas induite par une norme vectorielle du fait que pour tout entier $n \geq 2$ on a $\|I_n\|_s = \sqrt{n} \neq 1$.

3.2 Le groupe topologique $GL_n(\mathbb{K})$

L'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ est muni d'une norme quelconque (elles sont toutes équivalentes).

Lemme 3.4 *L'application $\det : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ est continue.*

Preuve. L'application déterminant est continue comme fonction polynomiale des coefficients m_{ij} d'une matrice M . $\square$

Théorème 3.3.

L'ensemble $GL_n(\mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$.

Preuve. L'ensemble $GL_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$ en tant qu'image réciproque de l'ouvert $\mathbb{K}^*$ par l'application continue déterminant.

La fonction polynomiale $z \mapsto \det(zI_n - A)$ a au plus n racines dans $\mathbb{K}$, il existe donc un entier k_0 tel que :

$$\forall k > k_0, \det\left(\frac{1}{k}I_n - A\right) \neq 0$$

et on a $A = \lim_{\substack{k \rightarrow +\infty \\ k > k_0}} A_k$ avec les $A_k = \frac{1}{k}I_n - A$ inversibles pour tout $k > k_0$. $\square$

Ce théorème peut aussi se montrer en utilisant la caractérisation des matrices de rang r donné (exercice 3.3).

Théorème 3.4.

L'application $A \mapsto A^{-1}$ est continue de $GL_n(\mathbb{K})$ dans $GL_n(\mathbb{K})$.

Preuve. L'application $A \mapsto C(A)$ qui associe à une matrice A sa comatrice $C(A)$ est continue de $\mathcal{M}_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ car ses composantes sont des fonctions polynomiales des coefficients a_{ij} . On en déduit alors que l'application $A \mapsto A^{-1} = \frac{1}{\det(A)} {}^t C(A)$ est continue de $GL_n(\mathbb{K})$ dans $GL_n(\mathbb{K})$. $\square$

La continuité de $(U, V) \mapsto UV$ et $M \mapsto M^{-1}$ peut se traduire en disant que la topologie de $GL_n(\mathbb{K})$ est compatible avec sa structure de groupe. On dit aussi que $GL_n(\mathbb{K})$ est un groupe topologique.

Théorème 3.5.

Le centre de $\mathcal{M}_n(\mathbb{K})$ est formé des homothéties et le centre de $GL_n(\mathbb{K})$ est formé des homothéties non nulles.

Preuve. Soit $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans le centre de $\mathcal{M}_n(\mathbb{K})$, c'est-à-dire commutant avec toutes les matrices. On a $AE_{ij} = E_{ij}A$, pour tous i, j et en désignant par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$, on a :

$$AE_{ij}e_j = Ae_i = \sum_{k=1}^n a_{ki}e_k = E_{ij}Ae_j = E_{ij}\left(\sum_{k=1}^n a_{kj}e_k\right) = a_{jj}e_i$$

Donc $a_{ki} = 0$ pour $k \in \{1, \dots, n\} \setminus \{i\}$ et $a_{ii} = a_{jj}$. C'est-à-dire que $A = \lambda I_n$. On peut aussi remarquer que si A est dans le centre de $\mathcal{M}_n(\mathbb{K})$, elle commute alors à tout projecteur p_x sur la droite $\mathbb{K}x$. On en déduit que toutes les droites sont stables par A et A est une homothétie.

Avec la densité de $GL_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ et la continuité du produit matriciel, on déduit que toute matrice dans le centre de $GL_n(\mathbb{K})$ est aussi dans le centre de $\mathcal{M}_n(\mathbb{K})$, c'est donc une homothétie et son rapport est non nulle puisqu'elle est inversible. $\square$

Ces résultats sont en fait valables sur tout corps commutatif infini. Pour le cas de $\mathcal{M}_n(\mathbb{K})$, la démonstration est identique. Si A dans le centre de $GL_n(\mathbb{K})$. Pour tout $B \in \mathcal{M}_n(\mathbb{K})$, le polynôme $\det(\lambda I_n - B)$ a au plus n racines dans $\mathbb{K}$, donc pour $\mathbb{K}$ infini, il existe $\lambda \in \mathbb{K}$ tel que $\lambda I_n - B$ soit inversible. On a alors $A(\lambda I_n - B) = (\lambda I_n - B)A$, c'est-à-dire $\lambda A - AB = \lambda A - BA$ et $AB = BA$. Donc A est dans le centre de $\mathcal{M}_n(\mathbb{K})$.

Corollaire 3.1 : Les groupes multiplicatifs $GL_n(\mathbb{R})$ et $GL_n(\mathbb{C})$ ne sont pas isomorphes.

Preuve. Si $\varphi : GL_n(\mathbb{R}) \rightarrow GL_n(\mathbb{C})$ est un isomorphisme de groupes multiplicatifs, il induit alors un isomorphisme du centre de $GL_n(\mathbb{R})$ sur celui de $GL_n(\mathbb{C})$ (dire que A est dans le centre de $GL_n(\mathbb{R})$ équivaut à dire que $AM = MA$ pour tout $M \in GL_n(\mathbb{R})$ ce qui équivaut à :

$$\forall M \in GL_n(\mathbb{R}), \varphi(A) \varphi(M) = \varphi(M) \varphi(A)$$

encore équivalent à :

$$\forall M \in GL_n(\mathbb{C}), \varphi(A) M = M \varphi(A)$$

et donc à $\varphi(A)$ est dans le centre de $GL_n(\mathbb{C})$. On aurait alors un isomorphisme de groupes multiplicatifs de $\mathbb{R}^*$ sur $\mathbb{C}^*$, ce qui est impossible (i est d'ordre 4 dans $\mathbb{C}^*$ et il n'y a pas d'élément d'ordre 4 dans $\mathbb{R}^*$). $\square$

Corollaire 3.2 : Il existe une base de $\mathcal{M}_n(\mathbb{K})$ formée de matrices inversibles.

Preuve. $V = \text{Vect}(GL_n(\mathbb{K}))$ est un sous espace vectoriel fermé de $\mathcal{M}_n(\mathbb{K})$ (on est en dimension finie) qui contient $GL_n(\mathbb{K})$, il contient donc son adhérence c'est-à-dire $\mathcal{M}_n(\mathbb{K})$. Donc $V = \mathcal{M}_n(\mathbb{K})$ et du système générateur $GL_n(\mathbb{K})$ on peut extraire une base. $\square$

Ce résultat est encore valable sur un corps commutatif infini. En effet il existe une infinité de scalaires $\lambda \in \mathbb{K}$ tels que, pour tous i, j dans $\{1, \dots, n\}$, $E_{ij} + \lambda I_n$ est inversible. Pour un tel $\lambda \in \mathbb{K}$ la famille $\mathcal{B} = (E_{ij} + \lambda I_n)_{1 \leq i, j \leq n}$ est une base de $\mathcal{M}_n(\mathbb{K})$. En effet l'égalité $\sum_{1 \leq i, j \leq n} a_{ij} (E_{ij} + \lambda I_n) = 0$ entraîne $a_{ij} = 0$ pour

$$i \neq j \text{ et on a } \sum_{i=1}^n \left(a_{ii} + \lambda \sum_{j=1}^n a_{jj} \right) E_{ii} = 0 \quad (I_n = \sum_{i=1}^n E_{ii}), \text{ donc } a_{ii} + \lambda \sum_{j=1}^n a_{jj} = 0$$

pour tout $i \in \{1, \dots, n\}$ et en sommant on a $(1 + n\lambda) \sum_{j=1}^n a_{jj} = 0$, ce qui donne

$\sum_{j=1}^n a_{jj} = 0$ en choisissant $\lambda \in \mathbb{K}$ tel que $1 + n\lambda \neq 0$ et $a_{ii} = 0$ pour tout i . La

famille $\mathcal{B}$ est donc libre et c'est une base car formé de n^2 éléments.

Corollaire 3.3 : Pour $n \geq 2$ il n'existe pas de norme sur $\mathcal{M}_n(\mathbb{K})$ telle que $\|P^{-1}AP\| = \|A\|$ pour tout A dans $\mathcal{M}_n(\mathbb{K})$ et tout P dans $GL_n(\mathbb{K})$.

Preuve. Supposons que pour toutes matrices A dans $\mathcal{M}_n(\mathbb{K})$ et P dans $GL_n(\mathbb{K})$ on ait $\|P^{-1}AP\| = \|A\|$. En appliquant cette égalité à PA et P , on a $\|AP\| = \|PA\|$. Par densité de $GL_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ et continuité du produit matriciel, on déduit que $\|AP\| = \|PA\|$ pour tous A, P dans $\mathcal{M}_n(\mathbb{K})$. Mais ce dernier résultat est impossible pour $n \geq 2$. En effet $E_{12}E_{11} = 0$ et $E_{11}E_{12} \neq 0$. $\square$

De la densité de $GL_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$, on peut déduire une généralisation à $\mathcal{M}_n(\mathbb{K})$ du théorème de décomposition polaire des matrices inversibles (corollaire 2.13). Pour ce faire on utilise le lemme suivant.

Lemme 3.5 L'ensemble $\mathcal{U}_n(\mathbb{C})$ [resp. $\mathcal{O}_n(\mathbb{R})$] des matrices complexes unitaires [resp. réelles orthogonales] est compact dans $\mathcal{M}_n(\mathbb{C})$ [resp. $\mathcal{M}_n(\mathbb{R})$].

Preuve. On munit l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ de la norme matricielle $\|\cdot\|_2$ induite par la norme hermitienne ou euclidienne de $\mathbb{K}^n$. Du fait qu'une transformation unitaire ou orthogonale conserve la norme hermitienne ou euclidienne de $\mathbb{K}^n$, on déduit que pour toute matrice A unitaire ou orthogonale on a $\|A\|_2 = 1$. On déduit donc que $\mathcal{U}_n(\mathbb{C})$ [resp. $\mathcal{O}_n(\mathbb{R})$] est borné dans $(\mathcal{M}_n(\mathbb{K}), \|\cdot\|_2)$. De plus cet ensemble est fermé comme image réciproque du fermé $\{I_n\}$ de $\mathcal{M}_n(\mathbb{K})$ par l'application continue $A \mapsto A^*A$. En conclusion $\mathcal{U}_n(\mathbb{C})$ [resp. $\mathcal{O}_n(\mathbb{R})$] est compact dans $(\mathcal{M}_n(\mathbb{K}), \|\cdot\|_2)$ en tant que fermé borné (on est en dimension finie). $\square$

Corollaire 3.4 : Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ [resp. $A \in \mathcal{M}_n(\mathbb{R})$] peut s'écrire $A = UH$ [resp. $A = \Omega S$] où U [resp. Ω] est une matrice unitaire [resp. orthogonale] et H [resp. S] une matrice hermitienne [resp. symétrique] positive.

Preuve. Toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ peut s'écrire $A = \lim_{k \rightarrow +\infty} A_k$ où $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices inversibles. Avec le théorème de décomposition polaire (corollaire 2.13), on peut écrire pour tout entier k , $A_k = U_k H_k$ [resp. $A_k = \Omega_k S_k$] où $(U_k)_{k \in \mathbb{N}}$ [resp. $(\Omega_k)_{k \in \mathbb{N}}$] est une suite de matrices unitaires [resp. orthogonales] et $(H_k)_{k \in \mathbb{N}}$ [resp. $(S_k)_{k \in \mathbb{N}}$] une suite de matrices hermitiennes [resp. symétriques] définies positives. De la suite $(U_k)_{k \in \mathbb{N}}$ [resp. $(\Omega_k)_{k \in \mathbb{N}}$] dans le compact $\mathcal{U}_n(\mathbb{C})$ [resp. $\mathcal{O}_n(\mathbb{R})$] on peut extraire une sous-suite $(U_{\varphi(k)})_{k \in \mathbb{N}}$ [resp. $(\Omega_{\varphi(k)})_{k \in \mathbb{N}}$] qui converge vers une matrice $U \in \mathcal{U}_n(\mathbb{C})$ [resp. $\Omega \in \mathcal{O}_n(\mathbb{R})$]. De $H_k = U_k^{-1} A_k = U_k^* A_k$ [resp. $S_k = \Omega_k^{-1} A_k = {}^t \Omega_k A_k$] et de la continuité du produit matriciel, on déduit que la suite $(H_{\varphi(k)})_{k \in \mathbb{N}}$ [resp. $(S_{\varphi(k)})_{k \in \mathbb{N}}$] est convergente. La limite H [resp. S] de cette suite est une matrice hermitienne [resp. symétrique] positive et on a $A = UH$ [resp. $A = \Omega S$]. $\square$

Pour A de rang $r < n$, la décomposition ci-dessus n'est pas unique. En effet, on peut diagonaliser la matrice hermitienne [resp. symétrique] positive H [resp. S] dans une base orthonormée $(e_i)_{1 \leq i \leq n}$ avec $He_i = \lambda_i e_i$ [resp. $Se_i = \lambda_i e_i$] pour $1 \leq i \leq n$ où $\lambda_i = 0$ pour $1 \leq i \leq n - r$ et $\lambda_i > 0$ sinon (si A n'est pas inversible, il en est alors de même de H [resp. S] et 0 est valeur propre de H [resp. S]). Les Ue_i [resp. Ωe_i] sont alors uniquement déterminés pour $n - r + 1 \leq i \leq n$, mais pour $1 \leq i \leq n - r$ il n'y a pas unicité.

En désignant par $\mathcal{H}_n^{++}(\mathbb{C})$ [resp. $\mathcal{S}_n^{++}(\mathbb{R})$] l'ensemble des matrices complexes [resp. réelles] d'ordre n hermitiennes [resp. symétriques] définies positives, le théorème de décomposition polaire des matrices inversibles peut s'exprimer comme suit.

Théorème 3.6.

L'application $(U, H) \mapsto UH$ [resp. $(\Omega, S) \mapsto \Omega S$] réalise un homéomorphisme de $\mathcal{U}_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C})$ [resp. $\mathcal{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$] sur $GL_n(\mathbb{C})$ [resp. $GL_n(\mathbb{R})$].

Preuve. On fait la démonstration dans le cas des matrices réelles inversibles. Le cas complexe se traitant de manière analogue.

On sait que toute matrice $A \in GL_n(\mathbb{R})$ s'écrit de manière unique $A = \Omega S$ avec $\Omega \in \mathcal{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^{++}(\mathbb{R})$, donc l'application $\varphi : (\Omega, S) \mapsto \Omega S$ réalise une bijection de $\mathcal{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$ sur $GL_n(\mathbb{R})$. Cette application est continue car ses composantes sont des fonctions polynomiales des coefficients ω_{ij} de Ω et s_{rp} de S . Il reste à montrer que φ^{-1} est continue. Soit $(A_k)_{k \in \mathbb{N}}$ une suite de matrices dans $GL_n(\mathbb{R})$ qui converge vers A . On note $\varphi^{-1}(A_k) = (\Omega_k, S_k)$ pour tout $k \in \mathbb{N}$ et $\varphi^{-1}(A) = (\Omega, S)$. De la suite $(\Omega_k)_{k \in \mathbb{N}}$ dans le compact $\mathcal{O}_n(\mathbb{R})$, on peut extraire une sous-suite $(\Omega_{\varphi(k)})_{k \in \mathbb{N}}$ qui converge vers une matrice $\Omega' \in \mathcal{O}_n(\mathbb{R})$. De $S_k = {}^t \Omega_k A_k$, on déduit que la suite $(S_{\varphi(k)})_{k \in \mathbb{N}}$ converge vers $S' = {}^t \Omega' A$. La matrice S' est symétrique positive comme limite d'une suite de matrices symétriques positives et elle est définie puisque inversible. On a alors la décomposition polaire $A = \Omega' S'$. Cette dernière décomposition étant unique, on a nécessairement $\Omega' = \Omega$. On a donc ainsi montré que la suite $(\Omega_k)_{k \in \mathbb{N}}$ a une unique valeur d'adhérence dans le compact $\mathcal{O}_n(\mathbb{R})$, elle converge donc vers Ω et $(S_k)_{k \in \mathbb{N}} = ({}^t \Omega_k A_k)_{k \in \mathbb{N}}$ converge vers ${}^t \Omega A = S$. C'est-à-dire que la suite $((\Omega_k, S_k))_{k \in \mathbb{N}} = (\varphi^{-1}(A_k))_{k \in \mathbb{N}}$ converge vers $(\Omega, S) = \varphi^{-1}(A)$ et φ^{-1} est continue. $\square$

Théorème 3.7.

L'ensemble $GL_n(\mathbb{C})$ est connexe par arcs.

Preuve. Pour $A \neq B$ dans $GL_n(\mathbb{C})$, le polynôme $P(X) = \det((1 - X)A + XB)$ est de degré inférieur ou égal à n . Il a donc au plus n racines distinctes $z_1, \dots, z_p$ dans $\mathbb{C}$. On note $\Omega = \mathbb{C} \setminus \{z_1, \dots, z_p\}$. On a $\{0, 1\} \subset \Omega$ car $P(0) = \det(A) \neq 0$ et $P(1) = \det(B) \neq 0$. De plus parmi les droites $D_\alpha = \{z = \rho e^{i\alpha}; \rho \in \mathbb{R}\}$ et $\Delta_\beta = \{z = 1 + \rho e^{i\beta}; \rho \in \mathbb{R}\}$ avec $0 \leq \alpha, \beta < \pi$, il en existe une infinité qui ne rencontrent pas $\{z_1, \dots, z_p\}$ et telles que $D_\alpha \cap \Delta_\beta \neq \emptyset$ (figure 3.1). En choisissant deux telles droites qui se coupent en γ , le chemin formé de la juxtaposition des

segments $[0, \gamma]$ et $[\gamma, 1]$ permet de joindre continûment les points 0 et 1 dans Ω . On note $\varphi : [0, 1] \rightarrow \Omega$ ce chemin et on définit $\psi : [0, 1] \rightarrow \mathcal{M}_n(\mathbb{C})$ par :

$$\forall t \in [0, 1], \psi(t) = (1 - \varphi(t))A + \varphi(t)B$$

Cette application est continue et $P(\psi(t)) \neq 0$ pour tout $t \in [0, 1]$ ($\psi(t)$ est dans Ω), c'est-à-dire que ψ est une fonction continue qui relie A et B dans $GL_n(\mathbb{C})$ ($\psi(0) = A$ et $\psi(1) = B$). On a donc ainsi prouvé que $GL_n(\mathbb{C})$ est connexe par arcs. $\square$

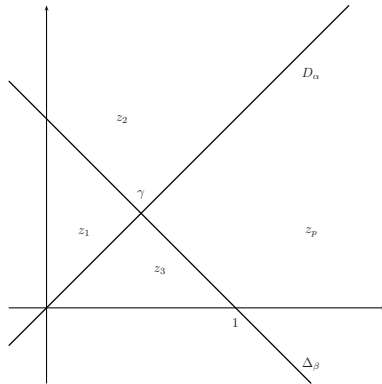


FIGURE 3.1 –

On peut aussi montrer la connexité de $GL_n(\mathbb{C})$ en utilisant le théorème de trigonalisation des matrices complexes (exercice 3.9).

L'ensemble $GL_n(\mathbb{R})$ n'est pas connexe (exercice 3.10).

Des résultats sur la réduction des matrices unitaires complexes et orthogonales réelles on déduit les résultats de connexité suivants.

Théorème 3.8.

Le sous groupe $U_n(\mathbb{C})$ de $GL_n(\mathbb{C})$ formé des matrices unitaires est connexe par arcs.

Preuve. Une matrice unitaire a toutes ses valeurs propres de module 1 et se diagonalise dans une base orthonormée. Pour tout $A \in U_n(\mathbb{C})$ il existe donc une matrice diagonale :

$$D = \begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & e^{i\theta_n} \end{pmatrix}$$

et une matrice unitaire $U \in U_n(\mathbb{C})$ telles que $A = UDU^*$. En posant, pour tout $t \in [0, 1]$, $\gamma(t) = U D(t) U^*$, où :

$$D(t) = \begin{pmatrix} e^{it\theta_1} & 0 & \cdots & 0 \\ 0 & e^{it\theta_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{it\theta_n} \end{pmatrix}$$

on définit un chemin continu dans $U_n(\mathbb{C})$ qui relie I_n et A . Ce qui prouve la connexité par arcs de $U_n(\mathbb{C})$. $\square$

Théorème 3.9.

Les composantes connexes du sous groupe $\mathcal{O}_n(\mathbb{R})$ de $GL_n(\mathbb{R})$ formé des matrices orthogonales sont $\mathcal{O}_n^+(\mathbb{R})$ et $\mathcal{O}_n^-(\mathbb{R})$.

Preuve. On sait que pour toute matrice orthogonale A , il existe une matrice Ω dans $\mathcal{O}_n(\mathbb{R})$ et une matrice :

$$D = \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \ddots & \vdots \\ 0 & 0 & R(\theta_1) & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & R(\theta_r) \end{pmatrix}$$

où pour tout $k \in \{1, \dots, r\}$, $R(\theta_k) = \begin{pmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{pmatrix}$ avec θ_k dans $]0, 2\pi[- \{\pi\}$, telles que $A = \Omega D {}^t\Omega$. Si de plus $A \in \mathcal{O}_n^+(\mathbb{R})$, l'entier q est alors nécessairement pair et la matrice D peut s'écrire $D = \begin{pmatrix} I_p & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{pmatrix}$ avec :

$$D_1 = \begin{pmatrix} R(\alpha_1) & 0 & \cdots & 0 \\ 0 & R(\alpha_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R(\alpha_{q'}) \end{pmatrix}$$

$$D_2 = \begin{pmatrix} R(\theta_1) & 0 & \cdots & 0 \\ 0 & R(\theta_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R(\theta_r) \end{pmatrix}$$

où $\alpha_j = \pi$ pour tout j . En posant, pour tout $t \in [0, 1]$, $\gamma(t) = \Omega D(t)^t \Omega$, où

$$D(t) = \begin{pmatrix} I_p & 0 & 0 \\ 0 & D_1(t) & 0 \\ 0 & 0 & D_2(t) \end{pmatrix}, \text{ avec :}$$

$$D_1(t) = \begin{pmatrix} R(t\alpha_1) & 0 & \cdots & 0 \\ 0 & R(t\alpha_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R(t\alpha_{q'}) \end{pmatrix}$$

$$D_2(t) = \begin{pmatrix} R(t\theta_1) & 0 & \cdots & 0 \\ 0 & R(t\theta_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R(t\theta_r) \end{pmatrix}$$

on définit un chemin continu dans $\mathcal{O}_n^+(\mathbb{R})$ qui relie I_n et A . Ce qui suffit à prouver la connexité par arcs de $\mathcal{O}_n^+(\mathbb{R})$.

Pour toute matrice $A \in \mathcal{O}_n^-(\mathbb{R})$ (par exemple $A = \begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix}$) l'application $M \mapsto AM$ réalise un homéomorphisme de $\mathcal{O}_n^+(\mathbb{R})$ sur $\mathcal{O}_n^-(\mathbb{R})$. On en déduit alors que $\mathcal{O}_n^-(\mathbb{R})$ est connexe par arcs.

On a $\mathcal{O}_n(\mathbb{R}) = \mathcal{O}_n^+(\mathbb{R}) \cup \mathcal{O}_n^-(\mathbb{R})$, avec $\mathcal{O}_n^+(\mathbb{R})$ et $\mathcal{O}_n^-(\mathbb{R})$ fermés (images réciproques de 1 et -1 respectivement par l'application déterminant) connexes disjoints dans $\mathcal{O}_n(\mathbb{R})$. Ce sont donc les composantes connexes de $\mathcal{O}_n(\mathbb{R})$. $\square$

3.3 Propriétés topologiques de l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$

Définition 3.2. Si $P(X) = \sum_{k=0}^n a_k X^k$ et $Q(X) = \sum_{k=0}^m b_k X^k$ sont deux polynômes non nuls dans $\mathbb{C}[X]$ avec $a_n \neq 0$ et $b_m \neq 0$, on appelle alors matrice de Sylvester de P et Q , la matrice de la famille de vecteurs $(P, XP, \dots, X^{m-1}P, Q, XQ, \dots, X^{n-1}Q)$ dans la base canonique de $\mathbb{C}_{n+m-1}[X]$. On note $S(P, Q)$ cette matrice, son déterminant est appelé le résultant de P et Q et est noté $\text{res}(P, Q)$.

Lemme 3.6 Soient P et Q deux polynômes non constants dans $\mathbb{C}[X]$. Ces polynômes ont une racine commune dans $\mathbb{C}$ si, et seulement si, il existe deux polynômes non nuls U et V tels que $\deg(U) < \deg(Q)$, $\deg(V) < \deg(P)$ et $UP + VQ = 0$.

Preuve. Si $\lambda \in \mathbb{C}$ est une racine commune à P et Q , il existe alors deux polynômes non nuls P_0 et Q_0 tels que $P(X) = (X - \lambda)P_0(X)$ et $Q(X) = (X - \lambda)Q_0(X)$.

En posant $U = Q_0$, $V = -P_0$, on a $UP + VQ = 0$ avec $\deg(U) = \deg(Q) - 1$ et $\deg(V) = \deg(P) - 1$. Réciproquement, supposons qu'il existe deux polynômes non nuls U, V dans $\mathbb{C}[X]$ tels que $UP = -VQ$ avec $\deg(U) < \deg(Q)$ et $\deg(V) < \deg(P)$. Si P et Q sont premiers entre eux, P divise alors V (théorème de Gauss), ce qui est impossible du fait que $\deg(V) < \deg(P)$. P et Q ne sont donc pas premiers entre eux et nécessairement ils ont une racine commune dans $\mathbb{C}$. $\square$

Théorème 3.10.

Deux polynômes non nuls P et Q ont une racine commune dans $\mathbb{C}$ si, et seulement si, $\text{res}(P, Q) = 0$.

Preuve. En reprenant la démonstration du lemme précédent, on pose :

$$U(X) = \sum_{k=0}^{m-1} u_k X^k, \quad V(X) = \sum_{k=0}^{n-1} v_k X^k$$

où n est le degré de P et m celui de Q . La condition $UP + VQ = 0$ avec U et V non nuls est équivalente à $\sum_{k=0}^{m-1} u_k X^k P(X) + \sum_{k=0}^{n-1} v_k X^k Q(X) = 0$, où les u_k et v_k ne sont pas tous nuls, ce qui équivaut à dire que la famille de vecteurs :

$$\{X^k P(X), X^j Q(X) \mid 0 \leq k \leq m-1, 0 \leq j \leq n-1\}$$

est liée dans $\mathbb{C}_{n+m-1}[X]$, ce qui est encore équivalent à dire que le résultant de P et Q est nul. En conclusion, P et Q ont une racine commune dans $\mathbb{C}$ si, et seulement si, $\text{res}(P, Q) = 0$. $\square$

On désigne par $\mathcal{D}'_n(\mathbb{C})$ l'ensemble des matrices $M \in \mathcal{M}_n(\mathbb{C})$ ayant n valeurs propres distinctes dans $\mathbb{C}$ et par $\mathcal{D}_n(\mathbb{C})$ l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{C})$.

Théorème 3.11.

L'ensemble $\mathcal{D}'_n(\mathbb{C})$ est l'intérieur de $\mathcal{D}_n(\mathbb{C})$.

Preuve. Une matrice $M \in \mathcal{M}_n(\mathbb{C})$ est dans $\mathcal{D}'_n(\mathbb{C})$ si, et seulement si, son polynôme caractéristique χ_M n'a que des racines simples dans $\mathbb{C}$, ce qui équivaut à dire que $\varphi(M) = \text{res}(\chi_M, \chi'_M) \neq 0$. L'ensemble $\mathcal{D}'_n(\mathbb{C})$ est donc un ouvert de $\mathcal{M}_n(\mathbb{C})$ comme image réciproque de l'ouvert $\mathbb{C}^*$ par l'application continue φ (φ est une fonction polynomiale des coefficients de M). Une matrice ayant n valeurs propres distinctes est diagonalisable, on a donc les inclusions $\mathcal{D}'_n(\mathbb{C}) \subset \mathcal{D}_n(\mathbb{C})$ et $\mathcal{D}'_n(\mathbb{C}) \subset \overset{\circ}{\mathcal{D}}_n(\mathbb{C})$ puisque $\mathcal{D}'_n(\mathbb{C})$ est ouvert. Supposons qu'il existe $A \in \overset{\circ}{\mathcal{D}}_n(\mathbb{C})$ ayant une valeur propre λ d'ordre supérieur ou égal à 2. On peut alors trouver une

matrice inversible P telle que $A = PDP^{-1}$ où :

$$D = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Pour tout entier $k > 0$, on pose :

$$\Delta_k = \begin{pmatrix} \lambda & \frac{1}{k} \\ 0 & \lambda \end{pmatrix}, \quad D_k = \begin{pmatrix} \Delta_k & 0 & \cdots & 0 \\ 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Le polynôme minimal de D_k est un multiple de celui de Δ_k , c'est-à-dire de $(X - \lambda)^2$ (si $P(D_k) = 0$, on a alors $P(\Delta_k) = 0$ et P est un multiple de π_{Δ_k}). En conséquence la matrice D_k et la matrice $A_k = PD_kP^{-1}$ ne sont pas diagonalisables (une matrice est diagonalisable si, et seulement si, son polynôme minimal est scindé à racines simples). Comme $A = \lim_{k \rightarrow +\infty} A_k$, on ne peut pas avoir $A \in \mathring{\mathcal{D}}_n(\mathbb{C})$. Donc toutes

les matrices de $\mathring{\mathcal{D}}_n(\mathbb{C})$ ont n valeurs propres distinctes et $\mathring{\mathcal{D}}_n(\mathbb{C}) \subset \mathcal{D}'_n(\mathbb{C})$. En définitive, $\mathring{\mathcal{D}}_n(\mathbb{C}) = \mathcal{D}'_n(\mathbb{C})$. $\square$

Théorème 3.12.

Les ensembles $\mathcal{D}'_n(\mathbb{C})$ et $\mathcal{D}_n(\mathbb{C})$ sont denses dans $\mathcal{M}_n(\mathbb{C})$.

Preuve. Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est semblable à une matrice triangulaire supérieure, c'est-à-dire qu'il existe une matrice P inversible et une matrice tri-

angulaire $T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{n-1,n-1} \\ 0 & \cdots & 0 & t_{nn} \end{pmatrix}$ telles que $A = PTP^{-1}$. On pose

alors :

$$\alpha = \begin{cases} 1 & \text{si } t_{ii} = t_{jj} \text{ pour tous } i \neq j \text{ dans } \{1, \dots, n\} \\ \inf_{1 \leq i, j \leq n, t_{ii} \neq t_{jj}} |t_{ii} - t_{jj}| & \text{sinon} \end{cases}$$

et on définit la suite de matrices $(T_k)_{k \geq 1}$ par $T_k = T + \Delta_k$, où :

$$\Delta_k = \begin{pmatrix} \frac{\alpha}{k} & 0 & \cdots & 0 \\ 0 & \frac{\alpha}{2k} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\alpha}{nk} \end{pmatrix}$$

Pour tout entier $k > 0$ la matrice T_k a toutes ses valeurs propres distinctes (si $t_{ii} + \frac{\alpha}{ik} = t_{jj} + \frac{\alpha}{jk}$ avec $t_{ii} \neq t_{jj}$, on a alors $|t_{ii} - t_{jj}| = \frac{\alpha}{k} \left| \frac{1}{i} - \frac{1}{j} \right| < \frac{\alpha}{k} \leq \alpha$, ce qui contredit la définition de α) et donc $T_k \in \mathcal{D}'_n(\mathbb{C})$ et elle est en particulier diagonalisable. On a alors, avec la continuité du produit matriciel, $A = \lim_{k \rightarrow +\infty} A_k$, où pour tout $k > 0$ la matrice $A_k = PT_kP^{-1}$ est dans $\mathcal{D}'_n(\mathbb{C})$ et diagonalisable. D'où la densité de $\mathcal{D}'_n(\mathbb{C})$ et $\mathcal{D}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$. $\square$

Pour $n \geq 2$, l'ensemble $\mathcal{D}_n(\mathbb{R})$ des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ n'est pas dense dans $\mathcal{M}_n(\mathbb{R})$ (exercice 3.13). De manière plus précise on peut montrer que l'adhérence de $\mathcal{D}_n(\mathbb{R})$ est l'ensemble $\mathcal{T}_n(\mathbb{R})$ des matrices trigonalisables de $\mathcal{M}_n(\mathbb{R})$ (exercice 3.14).

L'application qui associe à une matrice $A \in \mathcal{M}_n(\mathbb{C})$ son polynôme caractéristique $\chi_A \in \mathbb{C}_n[X]$ est continue (les coefficients de ce polynôme sont des fonctions polynomiales des a_{ij}). On peut alors se poser la question de savoir s'il en est de même pour l'application qui associe à une matrice son polynôme minimal. De la densité de $\mathcal{D}'_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$ on déduit que la réponse est négative.

Corollaire 3.5 : Pour $n \geq 2$, l'application qui associe à une matrice $A \in \mathcal{M}_n(\mathbb{C})$ son polynôme minimal n'est pas continue.

Preuve. Supposons que l'application $A \mapsto \pi_A$ qui associe à une matrice A de $\mathcal{M}_n(\mathbb{C})$ son polynôme minimal soit continue. Avec la densité de $\mathcal{D}'_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$, on peut écrire que $I_n = \lim_{k \rightarrow +\infty} A_k$ où $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices à valeurs propres deux à deux distinctes. Avec la continuité de $A \mapsto \pi_A$, on aurait alors $\pi_{I_n} = \lim_{k \rightarrow +\infty} \pi_{A_k}$. Mais pour toute matrice M dans $\mathcal{D}'_n(\mathbb{C})$ on a $\pi_M = \chi_M$ (polynôme caractéristique). On aurait alors avec la continuité de $A \mapsto P_A$:

$$(X - 1) = \pi_{I_n} = \lim_{k \rightarrow +\infty} \pi_{A_k} = \lim_{k \rightarrow +\infty} \chi_{A_k} = \chi_{I_n} = (X - 1)^n$$

ce qui est impossible pour $n \geq 2$. Donc, pour $n \geq 2$, l'application $A \mapsto \pi_A$ n'est pas continue. $\square$

3.4 Rayon spectral d'une matrice complexe

Définition 3.3. Le spectre d'une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est l'ensemble noté $\text{Sp}(A)$ de toutes ses valeurs propres complexes et son rayon spectral est le réel $\rho(A) = \max_{\lambda \in \text{Sp}(A)} |\lambda|$.

Le rayon spectral $\rho(A)$ est le rayon du plus petit disque centré en 0 du plan complexe contenant toutes les valeurs propres de la matrice A .

En utilisant le théorème 1.2, on a $\text{sp}(A^k) = \{\lambda^k \mid \lambda \in \text{sp}(A)\}$ pour tout $k \in \mathbb{N}$ et en conséquence :

$$\rho(A^k) = \max_{\lambda \in \text{sp}(A)} |\lambda^k| = \max_{\lambda \in \text{sp}(A)} |\lambda|^k = \rho(A)^k$$

puisque la fonction $t \mapsto t^k$ est croissante sur $\mathbb{R}^+$.

Lemme 3.7 Si $A \in \mathcal{M}_n(\mathbb{C})$ est une matrice normale, on a alors $\|A\|_2 = \rho(A)$.

Preuve. Une matrice normale $A \in \mathcal{M}_n(\mathbb{C})$ se diagonalise dans une base orthonormée, il existe donc des scalaires $\lambda_1, \dots, \lambda_n$ et une base orthonormée $(e_k)_{1 \leq k \leq n}$

de $\mathbb{C}^n$ tels que $Ae_k = \lambda_k e_k$ pour tout $k \in \{1, \dots, n\}$. Pour tout $x = \sum_{k=1}^n x_k e_k \in \mathbb{C}^n$

tel que $\|x\|_2^2 = \sum_{k=1}^n |x_k|^2 = 1$, on a :

$$\|Ax\|_2^2 = \sum_{k=1}^n |x_k|^2 |\lambda_k|^2 \leq \rho(A)^2 \sum_{k=1}^n |x_k|^2 = \rho(A)^2$$

On a donc $\|A\|_2 \leq \rho(A)$. Si $k \in \{1, \dots, n\}$ est tel que $\rho(A) = |\lambda_k|$, on a alors $\rho(A) = |\lambda_k| = \|Ae_k\|_2$ avec $\|e_k\|_2 = 1$. Donc $\|A\|_2 = \rho(A)$. $\square$

L'égalité $\|A\|_2 = \rho(A)$ est valable en particulier pour A complexe hermitienne ou unitaire et pour A réelle symétrique ou orthogonale.

Théorème 3.13.

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on a $\|A\|_2 = \sqrt{\|A^*A\|_2} = \sqrt{\rho(A^*A)}$.

Preuve. Pour $x \in \mathbb{C}^n$ tel que $\|x\|_2 = 1$, on a en utilisant l'inégalité de Cauchy-Schwarz :

$$\|Ax\|_2^2 = \langle x | A^*Ax \rangle \leq \|x\|_2 \|A^*Ax\|_2 \leq \|x\|_2 \|A^*A\|_2 \|x\|_2 = \|A^*A\|_2$$

Ce qui entraîne $\|A\|_2^2 \leq \|A^*A\|_2 \leq \|A\|_2 \|A^*\|_2$ et $\|A\|_2 \leq \|A^*\|_2$. En appliquant cette inégalité à A^* , on obtient $\|A^*\|_2 \leq \|A\|_2$ ce qui donne $\|A^*\|_2 = \|A\|_2$. On en déduit que $\|A\|_2^2 \leq \|A^*A\|_2 \leq \|A\|_2 \|A^*\|_2 = \|A\|_2^2$, donc $\|A\|_2^2 = \|A^*A\|_2$. La matrice A^*A étant normale (elle est hermitienne), on a aussi $\|A^*A\|_2 = \rho(A^*A)$, donc $\|A\|_2 = \sqrt{\|A^*A\|_2} = \sqrt{\rho(A^*A)}$. $\square$

Définition 3.4. Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, les racines carrées positives des valeurs propres de A^*A sont les valeurs singulières de A .

Lemme 3.8 En notant, pour tout réel $\delta > 0$, $D_\delta = \text{diag}(1, \delta, \dots, \delta^{n-1})$ dans $\mathcal{M}_n(\mathbb{R})$, on a $D_\delta^{-1}AD_\delta = (\delta^{j-i}a_{ij})_{1 \leq i, j \leq n}$ pour toute matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$.

Preuve. La multiplication à droite par la matrice diagonale D_δ a pour effet de multiplier la colonne numéro j par δ^{j-1} et la multiplication à gauche par $D_\delta^{-1} = D_{\frac{1}{\delta}}$ de diviser la ligne numéro i par δ^{i-1} . Il en résulte que le coefficient d'indice (i, j) de $D_\delta^{-1}AD_\delta$ est $\delta^{j-1}\delta^{1-i}a_{ij} = \delta^{j-i}a_{ij}$.

On peut aussi écrire en utilisant la base canonique $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de $\mathbb{K}^n$:

$$\begin{aligned} D_\delta^{-1} A D_\delta e_j &= D_\delta^{-1} A \delta^{j-1} e_j = \delta^{j-1} \sum_{i=1}^n a_{ij} D_{\frac{1}{\delta}} e_i = \delta^{j-1} \sum_{i=1}^n a_{ij} \frac{1}{\delta^{i-1}} e_i \\ &= \sum_{i=1}^n \delta^{j-i} a_{ij} e_i \end{aligned}$$

ce qui donne $(D_\delta^{-1} A D_\delta)_{ij} = \delta^{j-i} a_{ij}$. □

Lemme 3.9 Soit $A \in \mathcal{M}_n(\mathbb{C})$. Pour tout réel $\varepsilon > 0$ il existe $P_\varepsilon \in GL_n(\mathbb{C})$ telle que la matrice $T_\varepsilon = P_\varepsilon^{-1} A P_\varepsilon$ soit triangulaire supérieure avec :

$$T_\varepsilon = ((t_{ij}))_{1 \leq i, j \leq n}, \quad \max_{1 \leq i \leq n-1} \sum_{j=i+1}^n |t_{ij}| < \varepsilon$$

Preuve. Pour $A \in \mathcal{M}_n(\mathbb{C})$, il existe $P \in GL_n(\mathbb{C})$ telle que $T = P^{-1} A P$ soit triangulaire supérieure. En utilisant les notations du lemme précédent, on a pour tout réel $\delta > 0$:

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

La matrice T_δ est semblable à la matrice A et en notant $T_\delta = ((t'_{ij}))_{1 \leq i, j \leq n}$, on a $\lim_{\delta \rightarrow 0} t'_{ij} = 0$ pour $1 \leq i < j \leq n$, on peut donc choisir $\delta > 0$ tel que :

$$\forall i \in \{1, \dots, n-1\}, \quad \sum_{j=i+1}^n |t'_{ij}| < \varepsilon$$

Une autre façon de procéder est de considérer l'endomorphisme u de $\mathbb{C}^n$ canoniquement associé à la matrice A et une base $(e_k)_{1 \leq k \leq n}$ dans laquelle la matrice T de u est triangulaire supérieure. Pour tout réel $\delta > 0$, on note $\mathcal{B}_\delta = (e'_k)_{1 \leq k \leq n}$ la base de $\mathbb{C}^n$ définie par $e'_j = \delta^{j-1} e_j$ et on a :

$$\forall j \in \{1, \dots, n\}, \quad u(e'_j) = \delta^{j-1} \sum_{i=1}^j t_{ij} e_i = \sum_{i=1}^j \delta^{j-i} t_{ij} e'_i$$

Pour $\delta > 0$ assez petit la matrice de u dans $\mathcal{B}_\delta$ a alors la forme souhaitée. □

Théorème 3.14.

Soit $A \in \mathcal{M}_n(\mathbb{C})$.

1. Pour toute norme matricielle induite par une norme vectorielle, on a $\rho(A) \leq \|A\|$, l'inégalité pouvant être stricte et $\rho(A) \leq \inf_{k \in \mathbb{N}^*} \|A^k\|^{\frac{1}{k}}$.
2. Pour tout $\varepsilon > 0$, il existe une norme matricielle induite par une norme vectorielle telle que $\|A\| \leq \rho(A) + \varepsilon$.
3. $\rho(A) = \inf_{\|\cdot\| \in \mathcal{N}} \|A\|$, où $\mathcal{N}$ désigne l'ensemble de toutes les normes matricielles induites par une norme vectorielle (théorème de Householder).

Preuve.

1. Si $\lambda \in \text{Sp}(A)$ est telle que $\rho(A) = |\lambda|$ et x est un vecteur propre associé de norme 1 dans $(\mathbb{C}^n, \|\cdot\|)$, on a alors $\rho(A) = |\lambda| = \|\lambda x\| = \|Ax\| \leq \|A\|$. En prenant pour A une matrice nilpotente non nulle, on a $\rho(A) = 0 < \|A\|$. On en déduit que, pour tout $k \in \mathbb{N}^*$, on a $\rho(A)^k = \rho(A^k) \leq \|A^k\|$, soit $\rho(A) \leq \|A^k\|^{\frac{1}{k}}$ et en conséquence, $\rho(A) \leq \inf_{k \in \mathbb{N}^*} \|A^k\|^{\frac{1}{k}}$.
2. En utilisant le lemme 3.9, on peut trouver pour tout réel $\varepsilon > 0$, une matrice inversible P_ε telle que $T_\varepsilon = P_\varepsilon^{-1}AP_\varepsilon$ soit triangulaire supérieure avec :

$$T_\varepsilon = ((t_{ij}))_{1 \leq i, j \leq n}, \quad \max_{1 \leq i \leq n-1} \sum_{j=i+1}^n |t_{ij}| < \varepsilon$$

On associe à cette matrice P_ε la norme matricielle $M \mapsto \|M\|_P = \|P_\varepsilon^{-1}MP_\varepsilon\|_\infty$ (lemme 3.3) et on a :

$$\begin{aligned} \|A\|_P &= \|P_\varepsilon^{-1}AP_\varepsilon\|_\infty = \|T_\varepsilon\|_\infty \\ &= \max_{1 \leq i \leq n-1} \left(|t_{nn}|, |t_{ii}| + \sum_{j=i+1}^n |t_{ij}| \right) \leq \varepsilon + \max_{1 \leq i \leq n} |t_{ii}| \end{aligned}$$

soit $\|A\|_P \leq \rho(A) + \varepsilon$ puisque les t_{ii} sont les valeurs propres de A .

3. Résulte de ce qui précède. □

Le premier point du théorème précédent nous dit que l'application ρ est continue en 0.

On déduit aussi de ce point et de l'équivalence des normes sur $\mathcal{M}_n(\mathbb{C})$, déduire que pour toute norme sur $\mathcal{M}_n(\mathbb{C})$, il existe une constante $\alpha > 0$ telle que $\rho(A) \leq \alpha \|A\|$ pour tout A dans $\mathcal{M}_n(\mathbb{C})$. Mais on peut avoir $\rho(A) > \|A\|$. Par exemple, pour $n = 2$ avec la norme $A \mapsto \|A\| = \max_{1 \leq i, j \leq n} |a_{ij}|$ et la matrice

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \text{on a } \|A\| = \max\{|\cos(\theta)|, |\sin(\theta)|\} < 1 = \rho(A) \text{ pour } \theta \in \left]0, \frac{\pi}{2}\right[.$$

Avec l'exercice 3.24, on propose une autre démonstration du théorème de Householder.

Théorème 3.15. Gelfand

Quelle que soit la norme choisie N sur l'espace vectoriel $\mathcal{M}_n(\mathbb{C})$, on a $\rho(A) = \lim_{k \rightarrow +\infty} \left(N(A^k)^{\frac{1}{k}} \right)$ pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$.

Preuve. En reprenant les notations de la démonstration du deuxième point du théorème précédent, on a pour $A \in \mathcal{M}_n(\mathbb{C})$, $\varepsilon \in \mathbb{R}^{+,*}$ et $k \in \mathbb{N}^*$, en notant $\gamma_\varepsilon = \|P_\varepsilon\|_\infty \|P_\varepsilon^{-1}\|_\infty$:

$$\begin{aligned} \|A^k\|_\infty &= \|P_\varepsilon T_\varepsilon^k P_\varepsilon^{-1}\|_\infty \leq \|P_\varepsilon\|_\infty \|P_\varepsilon^{-1}\|_\infty \|T_\varepsilon^k\|_\infty \\ &\leq \gamma_\varepsilon \|T_\varepsilon\|_\infty^k < \gamma_\varepsilon (\rho(A) + \varepsilon)^k \end{aligned}$$

soit $\|A^k\|_\infty^{\frac{1}{k}} < \gamma_\varepsilon^{\frac{1}{k}} (\rho(A) + \varepsilon)$. Il en résulte que $\rho(A) \leq \|A^k\|_\infty^{\frac{1}{k}} < \gamma_\varepsilon^{\frac{1}{k}} (\rho(A) + \varepsilon)$ avec $\lim_{k \rightarrow +\infty} \gamma_\varepsilon^{\frac{1}{k}} = 1$, il existe donc un entier $k_0 \in \mathbb{N}^*$ tel que $\gamma_\varepsilon^{\frac{1}{k}} < 1 + \varepsilon$ pour tout $k \geq k_0$, ce qui nous donne $\rho(A) \leq \|A^k\|_\infty^{\frac{1}{k}} < (1 + \varepsilon)(\rho(A) + \varepsilon) = \rho(A) + \varepsilon(\rho(A) + \varepsilon + 1)$ pour tout $k \geq k_0$ et signifie que $\rho(A) = \lim_{k \rightarrow +\infty} \left(\|A^k\|_\infty^{\frac{1}{k}} \right)$. Comme $\rho(A)$ est un minorant de la suite $\left(\|A^k\|_\infty^{\frac{1}{k}} \right)_{k \in \mathbb{N}^*}$, c'est aussi la borne inférieure de cette suite. On a donc $\rho(A) = \lim_{k \rightarrow +\infty} \|A^k\|_\infty^{\frac{1}{k}} = \inf_{k \in \mathbb{N}^*} \|A^k\|_\infty^{\frac{1}{k}}$.

Toutes les normes sur $\mathcal{M}_n(\mathbb{C})$ étant équivalentes, il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que $\alpha \|X\|_\infty \leq N(X) \leq \beta \|X\|_\infty$ pour tout $X \in \mathcal{M}_n(\mathbb{C})$, ce qui nous donne pour tout $k \in \mathbb{N}^*$:

$$\alpha^{\frac{1}{k}} \|A^k\|_\infty^{\frac{1}{k}} \leq N(A^k)^{\frac{1}{k}} \leq \beta^{\frac{1}{k}} \|A^k\|_\infty^{\frac{1}{k}}$$

avec $\lim_{k \rightarrow +\infty} \alpha^{\frac{1}{k}} = \lim_{k \rightarrow +\infty} \beta^{\frac{1}{k}} = 1$ et $\lim_{k \rightarrow +\infty} \left(\|A^k\|_\infty^{\frac{1}{k}} \right) = \rho(A)$. On déduit alors du théorème d'encadrement que $\lim_{k \rightarrow +\infty} \left(N(A^k)^{\frac{1}{k}} \right) = \rho(A)$. $\square$

Le théorème de Gelfand peut aussi se montrer en utilisant la décomposition $D + N$ de Dunford (exercice 3.26).

Théorème 3.16.

L'application ρ qui associe à toute matrice de $\mathcal{M}_n(\mathbb{C})$ son rayon spectral est continue.

Preuve. On munit $\mathcal{M}_n(\mathbb{C})$ d'une norme matricielle induite par une norme vectorielle.

Si $(T_k)_{k \in \mathbb{N}}$ est une suite de matrices triangulaires supérieures qui converge vers une matrice T , cette matrice T est alors également triangulaire supérieure et ses valeurs propres sont les termes diagonaux $t_{ii} = \lim_{k \rightarrow +\infty} t_{ii}^{(k)}$, en notant $t_{ii}^{(k)}$ les termes diagonaux de T_k pour tout $k \in \mathbb{N}$ et tout i compris entre 1 et n . Avec la continuité

de l'application $x \in \mathbb{C}^n \mapsto \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$, on déduit que :

$$\rho(T) = \max_{1 \leq i \leq n} |t_{ii}| = \lim_{k \rightarrow +\infty} \max_{1 \leq i \leq n} |t_{ii}^{(k)}| = \lim_{k \rightarrow +\infty} \rho(T_k)$$

Soit $(A_k)_{k \in \mathbb{N}}$ une suite de matrices qui converge vers la matrice A dans $\mathcal{M}_n(\mathbb{C})$. On veut montrer que la suite $(\rho(A_k))_{k \in \mathbb{N}}$ converge vers $\rho(A)$ dans $\mathbb{R}$. Pour ce faire on va montrer que cette suite est bornée et admet $\rho(A)$ pour unique valeur d'adhérence.

Avec les inégalités $\rho(A_k) \leq \|A_k\|$ et la convergence de la suite $(A_k)_{k \in \mathbb{N}}$ on déduit que la suite $(\rho(A_k))_{k \in \mathbb{N}}$ est bornée dans $\mathbb{R}$ et on peut en extraire une sous-suite convergente $(\rho(A_{\varphi(k)}))_{k \in \mathbb{N}}$. En utilisant le théorème de Schur (corollaire 2.10), on peut trouver, pour tout entier naturel k , une matrice unitaire U_k telle que la matrice $T_k = U_k^* A_k U_k$ soit triangulaire supérieure. Dans le compact $\mathcal{U}_n(\mathbb{C})$, on peut extraire de $(U_{\varphi(k)})_{k \in \mathbb{N}}$ une sous-suite $(U_{\sigma(k)})_{k \in \mathbb{N}}$ qui converge vers une matrice unitaire U . La suite $(T_{\sigma(k)})_{k \in \mathbb{N}}$ converge alors vers la matrice $T = U^* A U$ qui est triangulaire supérieure semblable à la matrice A . On a alors :

$$\rho(A) = \rho(T) = \lim_{k \rightarrow +\infty} \rho(T_{\sigma(k)}) = \lim_{k \rightarrow +\infty} \rho(A_{\sigma(k)}) = \lim_{k \rightarrow +\infty} \rho(A_{\varphi(k)})$$

La suite réelle bornée $(\rho(A_k))_{k \in \mathbb{N}}$ admet donc $\rho(A)$ pour unique valeur d'adhérence, ce qui équivaut à dire qu'elle converge vers $\rho(A)$. $\square$

Avec l'exercice 3.25, on propose une autre démonstration de la continuité du rayon spectral. De ce théorème, on déduit que, pour tout réel $R > 0$, l'ensemble $\{A \in \mathcal{M}_n(\mathbb{C}) \mid \rho(A) < R\}$ est ouvert dans $\mathcal{M}_n(\mathbb{C})$ comme image réciproque de l'ouvert $] -\infty, R[$ par l'application continue ρ .

Lemme 3.10 Soient $\theta_1, \dots, \theta_q$ des réels deux à deux distincts dans $[0, 2\pi[$, $a_1, \dots, a_q$ des réels et $(u_k)_{k \in \mathbb{N}}$ la suite de nombres complexes définie par $u_k = \sum_{j=1}^q a_j e^{ik\theta_j}$ pour tout $k \in \mathbb{N}$. Si $\lim_{k \rightarrow +\infty} u_k = 0$, on a alors $a_j = 0$ pour tout j compris entre 0 et q .

Preuve. Comme $\lim_{k \rightarrow +\infty} u_k = 0$, la série entière $\sum u_k z^k$ a un rayon de convergence $R \geq 1$ et pour tout $z \in \mathbb{C}$ tel que $|z| < 1$, on a :

$$f(z) = \sum_{k=0}^{+\infty} u_k z^k = \sum_{j=1}^q a_j \sum_{k=0}^{+\infty} (z e^{i\theta_j})^k = \sum_{j=1}^q \frac{a_j}{1 - z e^{i\theta_j}}$$

ce qui nous donne, pour tout $j \in \{1, \dots, q\}$:

$$a_j = (1 - z e^{i\theta_j}) f(z) - (1 - z e^{i\theta_j}) \sum_{\substack{r=1 \\ r \neq j}}^q \frac{a_r}{1 - z e^{i\theta_r}}$$

En prenant $z = \rho e^{-i\theta_j}$, avec $0 < \rho < 1$, on obtient $a_j = \lim_{\rho \rightarrow 1^-} (1 - \rho) f(\rho e^{-i\theta_j})$ puisque les θ_r sont deux à deux distincts dans $[0, 2\pi[$. Comme $\lim_{k \rightarrow +\infty} u_k = 0$, pour

tout $\varepsilon > 0$ on peut trouver un entier k_0 tel que $|u_k| < \varepsilon$ pour tout $k > k_0$, ce qui donne, pour $|z| < 1$:

$$|f(z)| \leq \sum_{k=0}^{k_0} |u_k| + \varepsilon \sum_{k=k_0+1}^{+\infty} |z|^k = \sum_{k=0}^{k_0} |u_k| + \frac{\varepsilon |z|^{k_0+1}}{1-|z|} \leq \sum_{k=0}^{k_0} |u_k| + \frac{\varepsilon}{1-|z|}$$

et pour $z = \rho e^{-i\theta_j}$ avec $0 < \rho < 1$, on obtient :

$$(1-\rho) |f(\rho e^{-i\theta_j})| \leq (1-\rho) \left(\sum_{k=0}^{k_0} |u_k| \right) + \varepsilon < 2\varepsilon$$

pour ρ proche de 1. On a donc $\lim_{\rho \rightarrow 1^-} (1-\rho) f(\rho e^{-i\theta_j}) = 0$ et $a_j = 0$ pour tout j compris entre 1 et q . $\square$

Théorème 3.17.

Pour $A \in \mathcal{M}_n(\mathbb{C})$, les conditions suivantes sont équivalentes.

1. $\lim_{k \rightarrow +\infty} A^k = 0$;
2. pour toute valeur initiale x_0 , la suite $(x_k)_{k \in \mathbb{N}}$ définie par $x_{k+1} = Ax_k$, pour $k \geq 0$, converge vers le vecteur nul ;
3. $\rho(A) < 1$;
4. il existe au moins une norme matricielle induite telle que $\|A\| < 1$;
5. la matrice $I_n - A$ est inversible et la série de terme général A^k est convergente de somme $(I_n - A)^{-1}$;
6. la matrice $I_n - A$ est inversible et la série de terme général $\text{trace}(A^k)$ est convergente de somme $\text{trace}((I_n - A)^{-1})$;
7. $\lim_{k \rightarrow +\infty} \text{trace}(A^k) = 0$.

Preuve. (1) $\Rightarrow$ (2) Résulte de $\|x_k\| = \|A^k x_0\| \leq \|A^k\| \|x_0\|$.

(2) $\Rightarrow$ (3) Supposons qu'il existe une valeur propre λ de A telle que $|\lambda| \geq 1$. Si x_0 est un vecteur propre non nul associé à λ , en écrivant que $x_k = A^k x_0 = \lambda^k x_0$, on voit que la suite $(x_k)_{k \in \mathbb{N}}$ ne converge pas vers 0.

(3) $\Rightarrow$ (4) Soit $\varepsilon > 0$ tel que $\rho(A) + \varepsilon < 1$. Il suffit de prendre une norme matricielle induite telle que $\|A\| < \rho(A) + \varepsilon$ (théorème 3.14).

(4) $\Rightarrow$ (1) En prenant une norme matricielle induite qui vérifie $\|A\| < 1$ et en écrivant que $\|A^k\| \leq \|A\|^k$, on déduit que $\lim_{k \rightarrow +\infty} (A^k) = 0$.

On a donc montré que les assertions (1) à (4) sont équivalentes.

(3) $\Rightarrow$ (5) Si $\rho(A) < 1$, alors 1 n'est pas valeur propre de A et $I_n - A$ est inversible. En notant, pour tout entier p , $S_p = \sum_{k=0}^p A^k$, on a $(I_n - A) S_p = I_n - A^{p+1}$ avec $\lim_{p \rightarrow +\infty} (I_n - A^{p+1}) = I_n$. En utilisant la continuité du produit matriciel, on

déduit alors que :

$$\lim_{p \rightarrow +\infty} S_p = \lim_{p \rightarrow +\infty} (I_n - A)^{-1} (I_n - A^{p+1}) = (I_n - A)^{-1}$$

c'est-à-dire que $\sum_{k=0}^{+\infty} A^k = (I_n - A)^{-1}$.

(5) $\Rightarrow$ (6) La convergence de la série $\sum_{k=0}^{+\infty} A^k$ entraîne $\lim_{k \rightarrow +\infty} A^k = 0$ et en conséquence $\rho(A) < 1$. En notant $\lambda_1, \dots, \lambda_n$ les valeurs propres de A , on a pour tout entier p :

$$\sum_{k=0}^p \text{trace}(A^k) = \sum_{k=0}^p \sum_{j=1}^n \lambda_j^k = \sum_{j=1}^n \sum_{k=0}^p \lambda_j^k = \sum_{j=1}^n \frac{1 - \lambda_j^{p+1}}{1 - \lambda_j}$$

avec $|\lambda_j| < 1$ pour tout j . On déduit alors que la série de terme général $\text{trace}(A^k)$ est convergente avec :

$$\sum_{k=0}^{+\infty} \text{trace}(A^k) = \sum_{j=1}^n \frac{1}{1 - \lambda_j}$$

En considérant que les $\frac{1}{1 - \lambda_j}$, pour $1 \leq j \leq n$, sont toutes les valeurs propres de $(I_n - A)^{-1}$, on déduit que :

$$\sum_{k=0}^{+\infty} \text{trace}(A^k) = \text{trace}\left((I_n - A)^{-1}\right)$$

(6) $\Rightarrow$ (7) Résulte immédiatement du fait que le terme général d'une série convergente tend vers 0.

(7) $\Rightarrow$ (3) Supposons que $\rho(A) \geq 1$. On note $\lambda_1, \dots, \lambda_n$ les valeurs propres de A avec $|\lambda_1| = \dots = |\lambda_p| = \rho(A) > |\lambda_j|$ pour $j > p$ (dans le cas où $p < n$). On a alors :

$$\begin{aligned} \sum_{j=1}^p \left(\frac{\lambda_j}{\rho(A)} \right)^k &= \frac{1}{(\rho(A))^k} \sum_{j=1}^n \lambda_j^k - \sum_{j=p+1}^n \left(\frac{\lambda_j}{\rho(A)} \right)^k \\ &= \frac{1}{(\rho(A))^k} \text{trace}(A^k) - \sum_{j=p+1}^n \left(\frac{\lambda_j}{\rho(A)} \right)^k \xrightarrow{k \rightarrow +\infty} 0 \end{aligned}$$

En notant $\{e^{i\theta_1}, \dots, e^{i\theta_q}\}$ l'ensemble des valeurs prises par $\left\{ \frac{\lambda_1}{\rho(A)}, \dots, \frac{\lambda_p}{\rho(A)} \right\}$ avec les réels θ_j deux à deux distincts dans $[0, 2\pi[$, on a :

$$\sum_{j=1}^p \left(\frac{\lambda_j}{\rho(A)} \right)^k = \sum_{j=1}^q a_j e^{ik\theta_j} \xrightarrow{k \rightarrow +\infty} 0$$

les coefficients a_j étant des entiers strictement positifs, ce qui est impossible d'après le lemme 3.10. On a donc $\rho(A) < 1$. $\square$

3.5 Conditionnement d'une matrice

Quand on étudie un système linéaire de n équations à n inconnues à coefficients réels ou complexes, on peut se poser la question suivante : si $x \in \mathbb{K}^n$ est solution du système $Ax = b$, comment sera modifiée cette solution si les coefficients du second membre ou de la matrice sont modifiés ?

Considérons par exemple le système $Ax = b$, avec :

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$$

de solution $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. Si on modifie le second membre en $b + \delta b = \begin{pmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{pmatrix}$,

c'est-à-dire que b est donné avec une erreur relative de 0.3%, la solution devient alors $x + \delta x = \begin{pmatrix} 9.2 \\ -12.6 \\ 4.5 \\ -1.1 \end{pmatrix}$. On a donc une erreur relative sur x de l'ordre de

1000%. De même si on perturbe la matrice en prenant :

$$A + \delta A = \begin{pmatrix} 10 & 7 & 8.1 & 7.2 \\ 7.08 & 5.04 & 6 & 5 \\ 8 & 5.98 & 9.81 & 9 \\ 6.99 & 4.99 & 9 & 9.98 \end{pmatrix}$$

et, en gardant le second membre initial, la solution devient $x + \delta x = \begin{pmatrix} -8.1 \\ 137 \\ -34 \\ 22 \end{pmatrix}$.

Le théorème et la définition qui suivent permettent d'étudier plus en détail cette question.

Théorème 3.18.

Soient $x \mapsto \|x\|$ une norme sur $\mathbb{K}^n$, $A \mapsto \|A\|$ la norme matricielle induite, A une matrice dans $GL_n(\mathbb{K})$ et x dans $\mathbb{K}^n$ solution du système $Ax = b$. Si $x + \delta x$ est la solution du système perturbé $Ay = b + \delta b$, on a alors :

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta b\|}{\|b\|} \quad (3.5)$$

Si $x + \delta x$ est la solution du système perturbé $(A + \delta A)y = b$, on a alors :

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta A\|}{\|A\|} \quad (3.6)$$

Preuve. Supposons que $A(x + \delta x) = b + \delta b$ et $Ax = b$. On en déduit que $\delta x = A^{-1}\delta b$ et $\|\delta x\| \leq \|A^{-1}\| \|\delta b\|$. De même $b = Ax$ donne $\|b\| \leq \|A\| \|x\|$. L'inégalité (3.5) s'en déduit alors immédiatement.

De même, si $(A + \delta A)(x + \delta x) = b$ et $Ax = b$, on a alors $\delta x = -A^{-1}\delta A(x + \delta x)$ et $\|\delta x\| \leq \|A^{-1}\| \|\delta A\| \|x + \delta x\|$, d'où l'inégalité (3.6). $\square$

Ce qui nous conduit à poser la définition suivante.

Définition 3.5. Soit $A \mapsto \|A\|$ une norme matricielle induite par une norme vectorielle $x \mapsto \|x\|$. Si A est une matrice réelle ou complexe inversible, le conditionnement de A relativement à cette norme est alors la quantité $\text{cond}(A) = \frac{1}{\|A\| \|A^{-1}\|}$.

Le conditionnement n'est défini que pour une matrice inversible et dépend du choix d'une norme matricielle.

On notera cond_∞ , cond_1 et cond_2 les conditionnements associés respectivement aux trois normes classiques de $\mathbb{K}^n$.

Le mauvais conditionnement d'une matrice n'est pas lié à un déterminant voisin de 0. Dans le cas de l'exemple précédent, la matrice est symétrique définie positive de déterminant égal à 1.

Les valeurs propres de la matrice A sont 0.01, 0.84, 3.86 et 30.3. Le conditionnement, relativement à la norme euclidienne, est alors $\text{cond}_2(A) \cong \frac{0.01}{30.3} \cong 3.3 \cdot 10^{-4}$.

Le théorème suivant résume quelques propriétés du conditionnement.

Théorème 3.19.

Soit $A \mapsto \|A\|$ une norme matricielle induite par une norme vectorielle $x \mapsto \|x\|$. Pour toute matrice inversible A à coefficients réels ou complexes, on a $\text{cond}(A) \in]0, 1]$, $\text{cond}(A) = \text{cond}(A^{-1})$ et $\text{cond}(\alpha A) = \text{cond}(A)$ pour tout $\alpha \in \mathbb{K}^*$.

Preuve. En écrivant que $\|AA^{-1}\| = 1$, on déduit que $1 \leq \|A\| \|A^{-1}\|$, c'est-à-dire que le conditionnement est un réel non nul compris entre 0 et 1. Les autres propriétés sont des conséquences immédiates de la définition. $\square$

Un système sera bien conditionné si $\text{cond}(A)$ est voisin de 1 et mal conditionné si ce conditionnement est proche de 0.

Dans le cas de la norme hermitienne, le conditionnement d'une matrice peut se calculer en fonction de ses valeurs singulières (i. e. les valeurs propres de A^*A).

Théorème 3.20.

Soit A une matrice inversible à coefficients réels ou complexes. On a $\text{cond}_2(A) = \sqrt{\frac{\mu_{\min}}{\mu_{\max}}}$, où $\mu_{\min}$ [resp. $\mu_{\max}$] est la plus petite [resp. plus

grande] valeur propre de A^*A . En particulier, pour A normale, on a

$$\text{cond}_2(A) = \frac{\min_{1 \leq i \leq n} |\lambda_i|}{\max_{1 \leq i \leq n} |\lambda_i|}, \text{ où les } \lambda_i \text{ sont les valeurs propres de } A.$$

Preuve. On a $\|A\|_2^2 = \rho(A^*A) = \mu_{\max}$ et :

$$\|A^{-1}\|_2^2 = \rho\left((A^{-1})^* A^{-1}\right) = \rho\left((AA^*)^{-1}\right) = \frac{1}{\mu_{\min}}$$

ce qui donne $\text{cond}_2(A) = \frac{1}{\|A\|_2 \|A^{-1}\|_2} = \sqrt{\frac{\mu_{\min}}{\mu_{\max}}}.$

Si A est normale, il en est alors de même de A^{-1} et on a :

$$\|A\|_2 = \rho(A), \quad \|A^{-1}\|_2 = \rho(A^{-1}) = \frac{1}{\min_{1 \leq i \leq n} |\lambda_i|}$$

ce qui donne $\text{cond}_2(A) = \frac{1}{\|A\|_2 \|A^{-1}\|_2} = \frac{\min_{1 \leq i \leq n} |\lambda_i|}{\max_{1 \leq i \leq n} |\lambda_i|}.$ □

Pour A est unitaire complexe ou orthogonale réelle, on a $\|A\|_2 = \|A^{-1}\|_2 = 1$ et $\text{cond}_2(A) = 1$.

3.6 Quotient de Rayleigh-Ritz et Hausdorffien

Un outil intéressant pour l'étude des matrices hermitiennes est le quotient de Rayleigh-Ritz. Il nous permettra d'étudier le problème du conditionnement des valeurs propres dans le cas des matrices hermitiennes.

Pour ce paragraphe, $\mathbb{C}^n$ est muni de sa structure hermitienne canonique et on note S_1 la sphère unité de $(\mathbb{C}^n, \|\cdot\|_2)$.

Définition 3.6. Soit A une matrice complexe. On appelle quotient de Rayleigh-Ritz associé à cette matrice l'application :

$$R_A : x \in \mathbb{C}^n \setminus \{0\} \mapsto R_A(x) = \frac{\langle Ax | x \rangle}{\|x\|_2^2}$$

Définition 3.7. Soit A une matrice complexe. On appelle hausdorffien de A la partie de $\mathbb{C}$ définie par $\mathcal{H}(A) = R_A(\mathbb{C}^n \setminus \{0\})$.

Par linéarité, on vérifie facilement que le Hausdorffien d'une matrice A est aussi défini par :

$$\mathcal{H}(A) = R_A(S_1) = \{\langle Ax | x \rangle \mid x \in S_1\}$$

Le résultat qui suit nous donne une définition variationnelle des valeurs propres d'une matrice hermitienne.

Théorème 3.21. Rayleigh-Ritz

Soit A une matrice hermitienne de valeurs propres (réelles) $\lambda_1 \leq \dots \leq \lambda_n$ et $(e_k)_{1 \leq k \leq n}$ une base orthonormée de vecteurs propres associés avec, pour tout entier k compris entre 1 et n , $Ae_k = \lambda_k e_k$. On a :

$$\begin{cases} \lambda_k = \sup \{ R_A(x) \mid x \in \text{Vect} \{e_{k+1}, \dots, e_n\}^\perp \setminus \{0\} \} & (1 \leq k \leq n-1) \\ \lambda_n = \sup \{ R_A(x) \mid x \in \mathbb{C}^n \setminus \{0\} \} \\ \\ \lambda_1 = \inf \{ R_A(x) \mid x \in \mathbb{C}^n \setminus \{0\} \} \\ \lambda_k = \inf \{ R_A(x) \mid x \in \text{Vect} \{e_1, \dots, e_{k-1}\}^\perp \setminus \{0\} \} & (2 \leq k \leq n) \end{cases}$$

Preuve. Pour tout entier k compris entre 1 et n , on a $R_A(e_k) = \langle Ae_k \mid e_k \rangle = \lambda_k$.

On note $V_n = \mathbb{C}^n$ et pour k compris entre 1 et $n-1$, $V_k = \text{Vect} \{e_{k+1}, \dots, e_n\}^\perp$.

Soit k un entier compris entre 1 et n . Tout vecteur x dans V_k s'écrit $x = \sum_{j=1}^k x_j e_j$.

Si ce vecteur est non nul, on a alors $R_A(x) = \frac{\sum_{j=1}^k \lambda_j x_j^2}{\sum_{j=1}^k x_j^2} \leq \lambda_k \frac{\sum_{j=1}^k x_j^2}{\sum_{j=1}^k x_j^2} = \lambda_k$ Avec

$R_A(e_k) = \lambda_k$, on en déduit que $\lambda_k = \sup \{ R_A(x) \mid x \in V_k \setminus \{0\} \}$.

Les identités avec les bornes inférieures se montrent de manière analogue. $\square$

Corollaire 3.6 : Si A est une matrice hermitienne de valeurs propres $\lambda_1 \leq \dots \leq \lambda_n$, on a alors $\mathcal{H}(A) = [\lambda_1, \lambda_n]$.

Preuve. Du théorème de Rayleigh-Ritz, on déduit que $\mathcal{H}(A) \subset [\lambda_1, \lambda_n]$.

En écrivant tout $\lambda \in [\lambda_1, \lambda_n]$ sous la forme $\lambda = (1-t)\lambda_1 + t\lambda_n$ et en posant $x = \sqrt{1-t}e_1 + \sqrt{t}e_n$ (notations du théorème 3.21), on a $\|x\|_2 = 1$ et $R_A(x) = \lambda$. On a donc $[\lambda_1, \lambda_n] \subset \mathcal{H}(A)$.

Une autre façon de voir les choses est de dire que la sphère unité de $\mathbb{C}^n$ est compacte et connexe et son image par la fonction continue R_A est compacte et connexe dans $\mathbb{R}$, c'est donc un intervalle fermé borné. Les valeurs propres λ_1 et λ_n étant les bornes inférieure et supérieure de cette image, on a $\mathcal{H}(A) = [\lambda_1, \lambda_n]$. $\square$

Ce corollaire s'interprète en disant que, si A est une matrice hermitienne, pour tout vecteur non nul x dans $\mathbb{C}^n$, il existe au moins une valeur propre de A inférieure ou égale à $R_A(x)$ et au moins une valeur propre supérieure ou égale à $R_A(x)$.

Dans le cas d'une matrice hermitienne on a vu que le hausdorffien est compact convexe et contient le spectre de A . En fait ce résultat est général.

Lemme 3.11 Pour toute matrice complexe A et tout scalaire α , on a :

$$\mathcal{H}(\alpha A) = \alpha \mathcal{H}(A), \quad \mathcal{H}(A + \alpha I_n) = \mathcal{H}(A) + \{\alpha\}$$

Preuve. Il suffit de vérifier. $\square$

Théorème 3.22.

Pour toute matrice complexe A , le hausdorffien $\mathcal{H}(A)$ est une partie compacte et convexe de $\mathbb{C}$ qui contient le spectre de A .

Preuve. Le hausdorffien $\mathcal{H}(A)$ est compact comme image de la sphère unité S_1 (compacte dans $\mathbb{C}^n$) par l'application continue R_A .

Pour toute valeur propre λ de A , on peut trouver un vecteur propre associé unitaire x et on a $\lambda = \langle Ax | x \rangle = R_A(x) \in \mathcal{H}(A)$. On a donc $\text{sp}(A) \subset \mathcal{H}(A)$.

Montrer que $\mathcal{H}(A)$ est convexe revient à montrer que pour tous x, y dans S_1 le segment $[R_A(x), R_A(y)]$ est contenu dans $\mathcal{H}(A)$. Si $R_A(x) = R_A(y)$, ce segment est alors réduit au point $R_A(x)$ qui est bien dans $\mathcal{H}(A)$. On suppose donc que $R_A(x) \neq R_A(y)$. Dans ces conditions le système $\{x, y\}$ est libre (en effet $y = \lambda x$ entraîne $|\lambda| = 1$ et $R_A(y) = |\lambda|^2 R_A(x) = R_A(x)$) et pour tout réel t compris entre 0 et 1, le vecteur $z(t) = tx + (1-t)y$ est non nul. On peut donc poser $\varphi(t) = R_A(z(t))$ et on définit ainsi une fonction continue de $[0, 1]$ dans $\mathbb{C}$ à valeurs dans $\mathcal{H}(A)$.

On suppose dans un premier temps que $R_A(x) = 1$ et $R_A(y) = 0$.

La matrice A peut s'écrire $A = H_1 + iH_2$ avec H_1 et H_2 hermitiennes (prendre $H_1 = \frac{1}{2}(A + A^*)$ et $H_2 = \frac{1}{2i}(A - A^*)$) et pour tout z dans $\mathbb{C}^n$ on a :

$$\langle Az | z \rangle = \langle H_1 z | z \rangle + i \langle H_2 z | z \rangle$$

avec $\langle H_1 z | z \rangle$ et $\langle H_2 z | z \rangle$ réels. Du fait que $R_A(x) = 1$ et $R_A(y) = 0$ sont réels on déduit alors que $\langle H_2 x | x \rangle$ et $\langle H_2 y | y \rangle$ sont nuls et pour tout réel t compris entre 0 et 1 on a :

$$\langle H_2 z(t) | z(t) \rangle = 2 \text{Re} \langle H_2 x | y \rangle t(1-t)$$

Quitte à remplacer x par λx avec $|\lambda| = 1$ (ce qui ne change pas $R_A(x)$), on peut supposer que $\langle H_2 x | y \rangle$ est imaginaire pur de sorte que :

$$\forall t \in \mathbb{R}, \langle H_2 z(t) | z(t) \rangle = 0$$

et :

$$\forall t \in \mathbb{R}, \langle Az(t) | z(t) \rangle = \langle H_1 z(t) | z(t) \rangle \in \mathbb{R}$$

La fonction φ est donc à valeurs réelles, continue sur $[0, 1]$ avec $\varphi(0) = 0$ et $\varphi(1) = 1$. Le théorème des valeurs intermédiaires nous dit alors que φ prend toutes les valeurs entre 0 et 1, c'est-à-dire que :

$$[R_A(x), R_A(y)] = [0, 1] \subset \varphi([0, 1]) \subset \mathcal{H}(A)$$

Dans le cas général, avec $R_A(x) \neq R_A(y)$, on peut trouver des scalaires α, β tels que $\alpha R_A(x) + \beta = 1$ et $\alpha R_A(y) + \beta = 0$, ce qui peut aussi s'écrire en tenant compte de $\|x\|_2 = \|y\|_2 = 1$, $\langle (\alpha A + \beta I_n)x | x \rangle = 1$ et $\langle (\alpha A + \beta I_n)y | y \rangle = 0$. En notant $B = \alpha A + \beta I_n$, on a $R_B(x) = 1$, $R_B(y) = 0$ et ce qui précède nous dit que $[0, 1] \subset \mathcal{H}(B) = \alpha \mathcal{H}(A) + \{\beta\}$, c'est-à-dire que tout réel t compris entre 0 et 1 s'écrit $t = \alpha \mu + \beta$ avec $\mu \in \mathcal{H}(A)$.

En écrivant que $\alpha = \frac{1}{R_A(x) - R_A(y)}$, $\beta = -\frac{R_A(y)}{R_A(x) - R_A(y)}$, on obtient pour tout $t \in [0, 1]$, $t = \frac{\mu - R_A(y)}{R_A(x) - R_A(y)}$, $1 - t = \frac{R_A(x) - \mu}{R_A(x) - R_A(y)}$, ce qui entraîne que $tR_A(x) + (1 - t)R_A(y) = \mu \in \mathcal{H}(A)$. On a donc bien $[R_A(x), R_A(y)] \subset \mathcal{H}(A)$, ce qui achève de prouver que $\mathcal{H}(A)$ est convexe. $\square$

Corollaire 3.7 : *Soit A une matrice complexe de trace nulle. Il existe une matrice unitaire U telle U^*AU ait tous ses termes diagonaux nuls.*

Preuve. On procède par récurrence sur $n \geq 1$. Pour $n = 1$, c'est clair. On le suppose acquis pour $n - 1 \geq 1$ et on se donne une matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{C})$ de trace nulle. On note u l'endomorphisme de $\mathbb{C}^n$ défini par la matrice A dans la base canonique. Pour tout i compris entre 1 et n , $a_{ii} = \langle Ae_i | e_i \rangle$ est dans $\mathcal{H}(A)$, c'est-à-dire que $\mathcal{H}(A)$ contient tous les termes diagonaux de A . Avec la convexité de $\mathcal{H}(A)$ on déduit alors que :

$$0 = \frac{1}{n} \text{Tr}(A) = \frac{1}{n} \sum_{i=1}^n a_{ii} \in \mathcal{H}(A)$$

Il existe donc un vecteur unitaire f_1 tel que $\langle Af_1 | f_1 \rangle = 0$. On complète ce vecteur en une base orthonormée $(f_k)_{1 \leq k \leq n}$ de $\mathbb{C}^n$ et dans cette base la matrice de u est de la forme $A_1 = U_1^*AU_1 = \begin{pmatrix} 0 & a \\ b & C \end{pmatrix}$, où U_1 est une matrice unitaire, a une matrice à une ligne et $n - 1$ colonnes, b une matrice à $n - 1$ lignes et 1 colonne et C une matrice carrée d'ordre $n - 1$. Comme $0 = \text{Tr}(A) = \text{Tr}(A_1) = \text{Tr}(C)$, on peut appliquer l'hypothèse de récurrence à la matrice C et il existe une matrice unitaire U_2 d'ordre $n - 1$ telle que la matrice $U_2^*CU_2$ soit de trace nulle. En posant $U_3 = \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}$ on définit une matrice unitaire d'ordre n et la matrice :

$$U_3^*U_1^*AU_1U_3 = \begin{pmatrix} 0 & a' \\ b' & U_2^*CU_2 \end{pmatrix}$$

est unitairement semblable à la matrice A et de trace nulle. $\square$

3.7 Conditionnement des problèmes de valeurs propres

Dans ce paragraphe on va voir que les problèmes de valeurs propres sont bien conditionnés pour les matrices hermitiennes.

Le résultat qui suit nous donne une autre caractérisation variationnelle des valeurs propres d'une matrice hermitienne. Contrairement à la caractérisation de Rayleigh-Ritz elle ne fait pas intervenir les vecteurs propres.

Pour tout k compris entre 1 et n , on désigne par E_k l'ensemble des sous espaces vectoriels de dimension k de $\mathbb{C}^n$.

Théorème 3.23. Courant-Fischer

Soit A une matrice hermitienne de valeurs propres $\lambda_1 \leq \dots \leq \lambda_n$. Pour tout entier k compris entre 1 et n on a $\lambda_k = \inf \{\mu_A(V) \mid V \in E_k\}$ où $\mu_A(V) = \sup \{R_A(x) \mid x \in V \setminus \{0\}\}$.

Preuve. On désigne par $(e_k)_{1 \leq k \leq n}$ une base orthonormée de vecteurs propres de la matrice A , avec, pour tout entier k compris entre 1 et n , $Ae_k = \lambda_k e_k$.

On note $V_n = \mathbb{C}^n$ et pour k compris entre 1 et $n-1$, $V_k = \text{Vect} \{e_{k+1}, \dots, e_n\}^\perp$. On a vu que $\lambda_k = \sup \{R_A(x) \mid x \in V_k \setminus \{0\}\} = \mu_A(V_k)$. On déduit donc que $\lambda_k \geq \alpha_k = \inf \{\mu(V) \mid V \in E_k\}$.

Soit $V \in E_k$, on a :

$$\begin{aligned} \dim(V \cap V_{k-1}^\perp) &= \dim(V) + \dim(V_{k-1}^\perp) - \dim(V + V_{k-1}^\perp) \\ &= n + 1 - \dim(V + V_{k-1}^\perp) \geq 1 \end{aligned}$$

donc $V \cap V_{k-1}^\perp \neq \{0\}$ et pour tout $y \in V \cap V_{k-1}^\perp \setminus \{0\}$, on a :

$$\lambda_k = \inf \{R_A(x) \mid x \in V_{k-1}^\perp \setminus \{0\}\} \leq R_A(y) \leq \sup \{R_A(x) \mid x \in V \setminus \{0\}\}$$

c'est-à-dire que :

$$\forall V \in E_k, \lambda_k \leq \sup \{R_A(x) \mid x \in V \setminus \{0\}\} = \mu_A(V)$$

soit $\lambda_k \leq \alpha_k$ et $\lambda_k = \alpha_k$. □

On note $\mathcal{H}_n(\mathbb{C})$ l'ensemble des matrices hermitiennes.

Corollaire 3.8 : Soit $A : [a, b] \mapsto \mathcal{H}_n(\mathbb{C})$ une application continue. Si pour tout t dans $[a, b]$ on note $\lambda_1(t) \leq \dots \leq \lambda_n(t)$ les valeurs propres de $A(t)$ rangées dans l'ordre croissant, les fonctions λ_k sont alors continues de $[a, b]$ dans $\mathbb{R}$.

Preuve. Pour $t \in [a, b]$, soit $(e_k(t))_{1 \leq k \leq n}$ une base orthonormale de vecteurs propres de $A(t)$, avec $A(t)e_k(t) = \lambda_k(t)e_k(t)$. On note $V_k(t)$ l'espace vectoriel engendré par $(e_1(t), \dots, e_k(t))$ (l'orthogonal de $\{e_{k+1}(t), \dots, e_n(t)\}$). En utilisant le théorème de Courant-Fischer, on peut écrire que pour tous t, t_0 dans $[a, b]$ on a :

$$\begin{aligned} \lambda_k(t) &= \inf \{\mu_{A(t)}(V) \mid V \in E_k\} \\ &\leq \mu_{A(t)}(V_k(t_0)) = \sup \{R_{A(t)}(x) \mid x \in V_k(t_0) \setminus \{0\}\} \end{aligned}$$

avec $R_{A(t)}(x) = R_{A(t_0)}(x) + R_{A(t)-A(t_0)}(x)$, on a alors :

$$\begin{aligned} \lambda_k(t) &\leq \sup \{R_{A(t_0)}(x) + R_{A(t)-A(t_0)}(x) \mid x \in V_k(t_0) \setminus \{0\}\} \\ &\leq \sup \{R_{A(t_0)}(x) \mid x \in V_k(t_0) \setminus \{0\}\} + \alpha_k(t_0) \end{aligned}$$

où on a noté $\alpha_k(t_0) = \sup \{R_{A(t)-A(t_0)}(x) \mid x \in V_k(t_0) \setminus \{0\}\}$. Mais on a aussi $\sup \{R_{A(t_0)}(x) \mid x \in V_k(t_0) \setminus \{0\}\} = \lambda_k(t_0)$, donc $\lambda_k(t) \leq \lambda_k(t_0) + \alpha_k(t_0)$ avec :

$$\begin{aligned} \alpha_k(t_0) &\leq \sup \{R_{A(t)-A(t_0)}(x) \mid x \in \mathbb{C}^n \setminus \{0\}\} \\ &\leq \sup \left\{ \frac{\langle (A(t) - A(t_0))x | x \rangle}{\|x\|_2^2} \mid x \in \mathbb{C}^n \setminus \{0\} \right\} \\ &\leq \|A(t) - A(t_0)\|_2 \end{aligned}$$

On a donc montré que, pour tous t, t_0 dans $[a, b]$ on a :

$$\lambda_k(t) \leq \lambda_k(t_0) + \|A(t) - A(t_0)\|_2$$

En permutant les rôles de t et t_0 , on déduit que :

$$|\lambda_k(t) - \lambda_k(t_0)| \leq \|A(t) - A(t_0)\|_2$$

ce qui suffit à prouver la continuité de λ_k . $\square$

Ce résultat peut s'interpréter en disant que de petites perturbations sur les coefficients d'une matrice hermitienne n'engendreront que de petites perturbations sur les valeurs propres. Ce qui revient à dire que le problème de valeurs propres est bien conditionné dans ce cas.

Si les valeurs propres d'une fonction continue A définie sur $[a, b]$ et à valeurs dans $\mathcal{H}_n(\mathbb{C})$ définissent des fonctions continues, en général il n'en n'est pas de même des vecteurs propres (exercice 3.28).

Pour toute matrice A dans $\mathcal{H}_n(\mathbb{C})$, on note $\lambda_1(A) \leq \dots \leq \lambda_n(A)$ ses valeurs propres rangées dans l'ordre croissant.

Corollaire 3.9 (Weyl) : Soient A, B deux matrices hermitiennes. Pour tout entier k compris entre 1 et n on a :

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B)$$

En particulier pour B positive, on a $\lambda_k(A) \leq \lambda_k(A+B)$.

Preuve. On a $\lambda_k(A+B) = \inf \{\mu_{A+B}(V) \mid V \in E_k\}$, pour tout entier k compris entre 1 et n avec $\mu_{A+B}(V) = \sup \{R_{A+B}(x) \mid x \in V \setminus \{0\}\}$ pour tout $V \in E_k$. En écrivant que $R_{A+B}(x) = R_A(x) + R_B(x) \geq R_A(x) + \lambda_1(B)$, on déduit que :

$$\mu_{A+B}(V) \geq \mu_A(V) + \lambda_1(B) \geq \lambda_k(A) + \lambda_1(B)$$

et $\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B)$.

Remarquant que les valeurs propres de $-B$ sont les $\lambda_k(-B) = -\lambda_{n-k+1}(B)$ et en appliquant l'inégalité précédente au couple $(A+B, -B)$, on obtient :

$$\lambda_k(A+B) + \lambda_1(-B) \leq \lambda_k(A+B-B)$$

soit $\lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B)$.

Pour B positive, on a $\lambda_1(B) \geq 0$ et $\lambda_k(A) \leq \lambda_k(A+B) - \lambda_1(B) \leq \lambda_k(A+B)$.

$\square$

Corollaire 3.10 : Pour tout entier k compris entre 1 et n , l'application $A \mapsto \lambda_k(A)$ est continue de $\mathcal{H}_n(\mathbb{C})$ dans $\mathbb{R}$.

Preuve. Soient A, B dans $\mathcal{H}_n(\mathbb{C})$. On a :

$$-\rho(B) \leq \lambda_1(B) \leq \lambda_k(A+B) - \lambda_k(A) \leq \lambda_n(B) \leq \rho(B)$$

c'est-à-dire $|\lambda_k(A+B) - \lambda_k(A)| \leq \rho(B) = \|B\|_2$ et la continuité de λ_k en découle.

□

En écrivant l'ensemble $\mathcal{H}_n^{++}(\mathbb{C})$ des matrices hermitiennes définies positives sous la forme $\mathcal{H}_n^{++}(\mathbb{C}) = \bigcap_{k=1}^n \lambda_k^{-1}(\mathbb{R}^*)$ et en utilisant la continuité des applications λ_k , on déduit que $\mathcal{H}_n^{++}(\mathbb{C})$ est ouvert dans $\mathcal{H}_n(\mathbb{C})$.

3.8 Exercices

Exercice 3.1. Montrer que pour toute norme N sur $\mathcal{M}_n(\mathbb{K})$ il existe une constante réelle $\lambda > 0$ telle que :

$$\forall (A, B) \in (\mathcal{M}_n(\mathbb{K}))^2, N(AB) \leq \lambda N(A) N(B)$$

Solution. L'application bilinéaire $(A, B) \mapsto AB$ est continue de $(\mathcal{M}_n(\mathbb{K}))^2$ dans $\mathcal{M}_n(\mathbb{K})$ (en dimension finie toute application bilinéaire est continue), il existe donc une constante réelle $\lambda > 0$ telle que :

$$\forall (A, B) \in (\mathcal{M}_n(\mathbb{K}))^2, N(AB) \leq \lambda N(A) N(B)$$

On peut aussi travailler tout d'abord avec une norme sous-multiplicative (par exemple $\|\cdot\|_1$) puis utiliser le théorème d'équivalence des normes sur $\mathcal{M}_n(\mathbb{K})$.

Exercice 3.2. Montrer que la norme $N : A \mapsto \max_{1 \leq i, j \leq n} |a_{ij}|$ sur $\mathcal{M}_n(\mathbb{K})$, n'est pas sous-multiplicative.

Solution. Par exemple, dans le cas $n = 2$, avec $A = \begin{pmatrix} a & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} b & 0 \\ 1 & 0 \end{pmatrix}$, où $a > 1$, $b > 1$, on a $AB = \begin{pmatrix} ab+1 & 0 \\ 0 & 0 \end{pmatrix}$ et $N(AB) = ab+1 > N(A)N(B) = ab$.

Exercice 3.3. Montrer la densité de $GL_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ en utilisant le fait que toute matrice de rang r est équivalente à $A_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ (exercice 1.4).

Solution. Soit $A \in M_n(\mathbb{K})$ de rang r . Si $r = 0$, on a alors $A = 0 = \lim_{k \rightarrow +\infty} \frac{1}{k} I_n$. Si $r > 0$, il existe alors deux matrices inversibles P et Q telles que $A = PA_r Q$ et on a $A = \lim_{k \rightarrow +\infty} PM_k Q$, avec $M_k = \begin{pmatrix} I_r & 0 \\ 0 & \frac{1}{k} I_{n-r} \end{pmatrix}$. Dans tous les cas on peut écrire A comme limite d'une suite de matrices inversibles. Donc $GL_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$.

Exercice 3.4. Dédurre de la densité de $GL_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ que pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$, AB et BA ont même polynôme caractéristique.

Solution. Pour A est inversible, AB et $BA = A^{-1}(AB)A$ sont semblables et en conséquence, ont même polynôme caractéristique. Dans le cas général, on peut écrire que $A = \lim_{k \rightarrow +\infty} A_k$ où $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices inversibles. Les matrices $A_k B$ et BA_k ont donc même polynôme caractéristique et avec la continuité du produit matriciel et du déterminant, on peut écrire que, pour tout λ dans $\mathbb{C}$, on a :

$$\begin{aligned} \det(\lambda I_n - AB) &= \lim_{k \rightarrow +\infty} \det(\lambda I_n - A_k B) \\ &= \lim_{k \rightarrow +\infty} \det(\lambda I_n - BA_k) = \det(\lambda I_n - BA) \end{aligned}$$

c'est-à-dire que AB et BA ont même polynôme caractéristique.

Exercice 3.5. Montrer que $GL_n(\mathbb{Q})$ est dense dans $\mathcal{M}_n(\mathbb{R})$.

Solution. On utilise sur $\mathcal{M}_n(\mathbb{R})$ la norme $N : A \mapsto \max_{1 \leq i, j \leq n} |a_{ij}|$. Soient A dans $\mathcal{M}_n(\mathbb{R})$ et $\varepsilon \in \mathbb{R}^{+,*}$. Comme $GL_n(\mathbb{R})$ est dense dans $\mathcal{M}_n(\mathbb{R})$, il existe $P = ((p_{ij}))_{1 \leq i, j \leq n} \in GL_n(\mathbb{R})$ telle que $N(A - P) < \frac{\varepsilon}{2}$ et comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, il existe pour tous i, j compris entre 1 et n , une suite $\left(r_{i,j}^{(k)}\right)_{k \in \mathbb{N}}$ de nombres rationnels qui converge vers p_{ij} . De la continuité du déterminant, on déduit, en notant $R_k = \left(\left(r_{i,j}^{(k)}\right)\right)_{1 \leq i, j \leq n}$ pour tout $k \in \mathbb{N}$, qu'il existe un entier k_0 tel que $\det(R_k) \neq 0$ pour tout $k > k_0$ (puisque $\lim_{k \rightarrow +\infty} \det(R_k) = \det(P) \neq 0$). Prenant k_0 assez grand, on peut aussi avoir $\left|p_{ij} - r_{i,j}^{(k)}\right| < \frac{\varepsilon}{2}$ pour tout $k > k_0$ et pour tous i, j compris entre 1 et n , ce qui nous donne $N(P - R_{k_0+1}) < \frac{\varepsilon}{2}$ et $N(A - R_{k_0+1}) < \varepsilon$ (inégalité triangulaire) avec $R_{k_0+1} \in GL_n(\mathbb{Q})$. On a ainsi montré que $GL_n(\mathbb{Q})$ est dense dans $\mathcal{M}_n(\mathbb{R})$.

Exercice 3.6. Soient $A \in \mathcal{M}_n(\mathbb{C})$ et $A = UH$ une décomposition polaire de la matrice A avec U unitaire et H hermitienne positive. Montrer que la matrice A est normale si, et seulement si, U et H commutent.

Solution. Dans tous les cas, on a $A^*A = HU^*UH = H^2$. Si les matrices U et H commutent, on alors $AA^* = UHHU^* = H^2 = A^*A$, c'est-à-dire que la matrice A est normale. Réciproquement, on suppose que la matrice A est normale. La matrice $K = UHU^*$ est hermitienne positive ($\langle Kx | x \rangle = \langle H(U^*x) | U^*x \rangle \geq 0$ puisque H est positive) avec $K^2 = UHHU^* = AA^* = A^*A = H^2$. On a donc $H^2 = K^2$ avec H et K hermitiennes positives, ce qui entraîne $H = K$ du fait de l'unicité de la racine carrée hermitienne positive (corollaire 2.12). On a donc $UHU^* = H$ et $UH = HU$.

Exercice 3.7. Dédurre le théorème de Cayley-Hamilton de la densité de $\mathcal{D}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$.

Solution. Si $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable, il existe alors une matrice inversible

$$P \text{ et une matrice diagonale } D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \text{ telles que } A = PDP^{-1}.$$

Ce qui entraîne que :

$$\chi_A(X) = \prod_{k=1}^n (X - \lambda_k), \quad \chi_A(A) = P\chi_A(D)P^{-1} = 0$$

Une matrice quelconque $A \in \mathcal{M}_n(\mathbb{C})$ s'écrit $A = \lim_{k \rightarrow +\infty} A_k$, où $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices diagonalisables. Avec la continuité de l'application $M \mapsto \chi_M(M)$ de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$ (les composantes de cette application sont des fonctions polynomiales des m_{ij}), on en déduit que $\chi_A(A) = \lim_{k \rightarrow +\infty} \chi_{A_k}(A_k) = 0$.

Exercice 3.8. Soit φ une forme linéaire sur l'espace $\mathcal{M}_n(\mathbb{K})$ telle que $\varphi(AB) = \varphi(BA)$ pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$.

1. En notant $(E_{ij})_{1 \leq i, j \leq n}$ la base canonique de $\mathcal{M}_n(\mathbb{K})$, montrer que $\varphi(E_{ii}) = \varphi(E_{jj})$ pour tous i, j compris entre 1 et n . On note λ cette valeur commune.
2. Montrer que $\varphi(A) = \lambda \operatorname{Tr}(A)$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$ (on peut d'abord supposer que la matrice A est diagonalisable). Ce résultat est en fait valable pour tout corps commutatif $\mathbb{K}$ de caractéristique nulle.
3. Soit u un endomorphisme de l'espace $\mathcal{M}_n(\mathbb{K})$ tel que $u(I_n) = I_n$ et $u(AB) = u(BA)$ pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$. Montrer que u conserve la trace.

Solution. On rappelle que si $(e_i)_{1 \leq i \leq n}$ désigne la base canonique de $\mathbb{K}^n$, la matrice E_{ij} est alors définie par :

$$\forall k \in \{1, \dots, n\}, \quad E_{ij}e_k = \begin{cases} 0 & \text{si } k \neq j \\ e_i & \text{si } k = j \end{cases}$$

1. Pour $i \neq j$ compris entre 1 et n , on a $E_{ij}E_{ji} = E_{ii}$ et $E_{ji}E_{ij} = E_{jj}$, donc $\varphi(E_{ii}) = \varphi(E_{ij}E_{ji}) = \varphi(E_{ji}E_{ij}) = \varphi(E_{jj})$. On peut donc poser $\lambda = \varphi(E_{ii})$ pour tout i compris entre 1 et n .
2. Si D est une matrice diagonale, elle s'écrit alors $D = \sum_{i=1}^n \lambda_i E_{ii}$ et on a :

$$\varphi(D) = \sum_{i=1}^n \lambda_i \varphi(E_{ii}) = \lambda \sum_{i=1}^n \lambda_i = \lambda \operatorname{Tr}(D)$$

Si A est une matrice diagonalisable, elle s'écrit alors $A = PDP^{-1}$ avec P inversible, D diagonale et on a :

$$\varphi(A) = \varphi(PDP^{-1}) = \varphi(DP^{-1}P) = \varphi(D) = \lambda \operatorname{Tr}(D) = \lambda \operatorname{Tr}(A)$$

Si $A \in \mathcal{M}_n(\mathbb{K})$ est quelconque, elle peut alors s'écrire comme limite d'une suite $(A_k)_{k \in \mathbb{N}}$ de matrices diagonalisables et avec la continuité des formes linéaires φ et Tr , on déduit que $\varphi(A) = \lim_{k \rightarrow +\infty} \varphi(A_k) = \lambda \lim_{k \rightarrow +\infty} \operatorname{Tr}(A_k) = \lambda \operatorname{Tr}(A)$.

3. On définit la forme linéaire φ par $\varphi(A) = \operatorname{Tr}(u(A))$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$. On a $\varphi(AB) = \varphi(BA)$ pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$, donc $\varphi(A) = \lambda \operatorname{Tr}(A)$, soit $\operatorname{Tr}(u(A)) = \lambda \operatorname{Tr}(A)$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$. Enfin avec $u(I_n) = I_n$ on déduit que $\lambda = 1$.

Exercice 3.9. Montrer que $GL_n(\mathbb{C})$ est connexe par arcs en utilisant le fait que toute matrice complexe est semblable à une matrice triangulaire.

Solution. Pour toute matrice $A \in GL_n(\mathbb{C})$, il existe une matrice $P \in GL_n(\mathbb{C})$ et une matrice triangulaire supérieure $T = ((m_{ij}))_{1 \leq i, j \leq n}$ telles que $A = PTP^{-1}$. On note, pour tout $j \in \{1, 2, \dots, n\}$, $m_{jj} = \rho_j e^{i\theta_j}$ avec $\rho_j > 0$ et on définit un chemin continu $\varphi : [0, 1] \rightarrow GL_n(\mathbb{C})$ par :

$$\forall t \in [0, 1], \varphi(t) = \begin{pmatrix} \varphi_{11}(t) & \varphi_{12}(t) & \cdots & \varphi_{1n}(t) \\ 0 & \varphi_{22}(t) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \varphi_{n-1,n-1}(t) \\ 0 & \cdots & 0 & \varphi_{nn}(t) \end{pmatrix}$$

où :

$$\varphi_{ij}(t) = \begin{cases} t m_{ij} & \text{si } 1 \leq i < j \leq n \\ (1-t)e^{it\theta_j} + t m_{jj} & \text{si } i = j \end{cases}$$

On a alors $\varphi(0) = I_n$, $\varphi(1) = T$ et $\gamma : t \mapsto P\varphi(t)P^{-1}$ est un chemin continu qui relie la matrice identité à la matrice A dans $GL_n(\mathbb{C})$.

Exercice 3.10. Montrer que $GL_n(\mathbb{R})$ n'est pas connexe.

Solution. Si $GL_n(\mathbb{R})$ est connexe, son image par l'application continue déterminant est alors un connexe de $\mathbb{R}$ (théorème des valeurs intermédiaires). Or cette image est $\mathbb{R}^*$ non connexe dans $\mathbb{R}$. Donc $GL_n(\mathbb{R})$ n'est pas connexe.

Exercice 3.11. Montrer que pour tout $r \in \{0, 1, \dots, n\}$ l'ensemble $\mathcal{A}_r$ des matrices de $\mathcal{M}_n(\mathbb{C})$ de rang r est connexe par arcs.

Solution. Deux matrices sont de même rang si, et seulement si, elles sont équivalentes. Donc pour toutes matrices A, B dans $\mathcal{A}_r$ il existe P et Q dans $GL_n(\mathbb{C})$ telles que $B = PAQ$. Si γ_1 et γ_2 sont deux fonctions continues de $[0, 1]$ dans $GL_n(\mathbb{C})$ telles que $\gamma_1(0) = I_n$, $\gamma_2(0) = I_n$, $\gamma_1(1) = P$, $\gamma_2(1) = Q$, alors $\gamma : t \mapsto \gamma_1(t)A\gamma_2(t)$ est un chemin continu qui relie A et B dans $\mathcal{A}_r$. Ce qui prouve que $\mathcal{A}_r$ est connexe par arcs.

Exercice 3.12. Montrer que pour tout $r \in \{0, 1, \dots, n-1\}$ l'ensemble $\mathcal{A}_r$ des matrices de $\mathcal{M}_n(\mathbb{R})$ de rang r est connexe par arcs.

Solution. Pour $r = 0$, le résultat est clair. Pour A de rang $r \geq 1$, il existe P et Q dans $GL_n(\mathbb{R})$ telles que $A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = PA_rQ$. En fait, pour $r < n$, on peut prendre P et Q dans $GL_n^+(\mathbb{R})$. En effet si P est dans $GL_n^-(\mathbb{R})$, alors la

$$\text{matrice } P \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & -1 \end{pmatrix} = PJ \text{ est dans } GL_n^+(\mathbb{R}) \text{ et en remarquant que,}$$

pour $r < n$, $JA_r = A_r$, on a $A = PJA_rQ$. De même pour Q . On déduit alors la connexité de $\mathcal{A}_r$ de celle de $GL_n^+(\mathbb{R})$ comme dans l'exercice 3.11.

Exercice 3.13. Montrer que l'ensemble $\mathcal{D}_2(\mathbb{R})$ des matrices diagonalisables de $\mathcal{M}_2(\mathbb{R})$ n'est pas dense dans $\mathcal{M}_2(\mathbb{R})$.

Solution. L'application $\varphi : M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}) \mapsto (a-d)^2 + 4bc \in \mathbb{R}$ qui associe à une matrice le discriminant de son polynôme caractéristique (résultant de χ_M et χ'_M) est continue, donc $A = \lim_{k \rightarrow +\infty} A_k$ entraîne $\varphi(A) = \lim_{k \rightarrow +\infty} \varphi(A_k)$. Mais pour A_k dans $\mathcal{D}'_2(\mathbb{R})$ ou dans $\mathcal{D}_2(\mathbb{R})$, on a $\varphi(A_k) \geq 0$ et pour A à valeurs propres complexes non réelles, on a $\varphi(A) < 0$. Une telle matrice A ne peut donc être limite d'une suite de matrices de $\mathcal{D}'_2(\mathbb{R})$ ou de $\mathcal{D}_2(\mathbb{R})$.

Exercice 3.14. On désigne par $\mathcal{T}_n(\mathbb{R})$ l'ensemble des matrices trigonalisables de $\mathcal{M}_n(\mathbb{R})$ et par $\theta : \mathcal{T}_n(\mathbb{R}) \rightarrow \mathbb{R}^n$ l'application définie par $\theta(M) = (\lambda_1, \dots, \lambda_n)$ où $\lambda_1 \leq \dots \leq \lambda_n$ sont les valeurs propres de la matrice M .

1. Montrer que $\mathcal{T}_n(\mathbb{R})$ est fermé dans $\mathcal{M}_n(\mathbb{R})$ (on peut montrer que si $(T_k)_{k \in \mathbb{N}}$ est une suite de $\mathcal{T}_n(\mathbb{R})$ qui converge vers $T \in \mathcal{M}_n(\mathbb{R})$, la suite $(\theta(T_k))_{k \in \mathbb{N}}$ est alors bornée dans $\mathbb{R}^n$ et en déduire que le polynôme caractéristique de T est scindé sur $\mathbb{R}$).

2. Montrer que $\mathcal{T}_n(\mathbb{R})$ est l'adhérence de l'ensemble $D_n(\mathbb{R})$ des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$.

Solution. On munit $\mathbb{R}^n$ de la norme $\|\cdot\|_\infty$ et $\mathcal{M}_n(\mathbb{R})$ de la norme matricielle induite. On a alors :

$$\forall M \in T_n(\mathbb{R}), \|\theta(M)\|_\infty = \rho(M) = \max_{1 \leq k \leq n} |\lambda_k| \leq \|M\|_\infty$$

1. Si $(T_k)_{k \in \mathbb{N}}$ est une suite dans $\mathcal{T}_n(\mathbb{R})$ qui converge vers $T \in \mathcal{M}_n(\mathbb{R})$, la suite $(\theta(T_k))_{k \in \mathbb{N}}$ est alors bornée dans $\mathbb{R}^n$ et on peut en extraire une sous-suite $(\theta(T_{\varphi(k)}))_{k \in \mathbb{N}}$ qui converge vers $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$. D'autre part, le po-

lynôme caractéristique de $M \in \mathcal{T}_n(\mathbb{R})$ est de la forme $\chi_M(X) = \sum_{j=0}^n a_j X^j$,

les coefficients a_j étant des fonctions continues de $\theta(M)$ (fonctions polynomiales symétriques élémentaires des racines). De $\lim_{k \rightarrow +\infty} \theta(T_{\varphi(k)}) = \lambda$, on dé-

duit alors que $\lim_{k \rightarrow +\infty} \chi_{T_{\varphi(k)}}(X) = \prod_{j=1}^n (X - \lambda_j)$. L'application qui associe à une

matrice M son polynôme caractéristique est continue (ses composantes sont des fonctions polynomiales des m_{ij}), donc $\lim_{k \rightarrow +\infty} \chi_{T_{\varphi(k)}} = \chi_T$ et le polynôme

$\chi_T(X) = \prod_{j=1}^n (X - \lambda_j)$ est scindé dans $\mathbb{R}[X]$, ce qui entraîne que $T \in \mathcal{T}_n(\mathbb{R})$.

On a donc ainsi prouvé que $\mathcal{T}_n(\mathbb{R})$ est fermé dans $\mathcal{M}_n(\mathbb{R})$.

2. On a $\mathcal{D}_n(\mathbb{R}) \subset \mathcal{T}_n(\mathbb{R})$ avec $\mathcal{T}_n(\mathbb{R})$ fermé, donc $\overline{\mathcal{D}_n(\mathbb{R})} \subset \mathcal{T}_n(\mathbb{R})$. Une démonstration identique à celle du théorème 3.12 permet de montrer l'inclusion $\mathcal{T}_n(\mathbb{R}) \subset \overline{\mathcal{D}_n(\mathbb{R})}$, d'où l'égalité $\overline{\mathcal{D}_n(\mathbb{R})} = \mathcal{T}_n(\mathbb{R})$.

Exercice 3.15. Soient A et B deux matrices normales dans $\mathcal{M}_n(\mathbb{C})$. Montrer que $\rho(AB) \leq \rho(A)\rho(B)$.

Solution. On a $\rho(AB) \leq \|AB\|_2 \leq \|A\|_2 \|B\|_2 = \rho(A)\rho(B)$ pour A, B normales.

Exercice 3.16. Soient $\sum_{k=0}^{+\infty} a_k z^k$ une série entière de rayon de convergence $R > 0$ et $A \in \mathcal{M}_n(\mathbb{K})$. Montrer que pour $\rho(A) < R$, la série de terme général $a_k A^k$ est convergente et pour $\rho(A) > R$, elle est divergente.

Solution. Supposons que $\rho(A) < R$. On sait qu'il existe une norme matricielle $\|\cdot\|$ subordonnée à une norme vectorielle telle que $\rho(A) \leq \|A\| < R$ (théorème 3.14). On a alors $\sum_{k=0}^{+\infty} |a_k| \|A\|^k < +\infty$ et avec $\|A^k\| \leq \|A\|^k$ pour tout entier naturel k , on déduit que la série de terme général $a_k A^k$ est normalement

convergente, donc convergente, dans l'espace complet $\mathcal{M}_n(\mathbb{K})$. Pour $\rho(A) > R$, la suite $(|a_k| \rho(A)^k)_{k \in \mathbb{N}}$ n'est pas bornée. Avec $\rho(A)^k = \rho(A^k)$ et $|a_k| \|A^k\| \geq |a_k| \rho(A^k)$, on en déduit que la suite $(|a_k| \|A^k\|)_{k \in \mathbb{N}}$ n'est pas bornée et la série de terme général $a_k A^k$ est divergente.

Exercice 3.17. On désigne par $\mathcal{F}$ l'ensemble des matrices non inversibles de $\mathcal{M}_n(\mathbb{C})$. On munit $\mathcal{M}_n(\mathbb{C})$ de la norme matricielle induite par la norme hermitienne canonique de $\mathbb{C}^n$ et pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on note $d(A, \mathcal{F}) = \inf_{M \in \mathcal{F}} \|A - M\|_2$.

1. Montrer que pour toute matrice $M \in \mathcal{F}$, on peut trouver un vecteur unitaire x tel que $\frac{1}{\|A^{-1}\|_2} \leq \|(A - M)x\|_2$.

2. Soient $A \in GL_n(\mathbb{C})$ et $\lambda_1^2, \dots, \lambda_n^2$ les valeurs propres de la matrice A^*A avec $0 < \lambda_1 \leq \dots \leq \lambda_n$. On rappelle qu'on peut écrire $A = UDV^*$

$$\text{où } U, V \text{ sont deux matrices unitaires et } D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

(exercice 2.11).

(a) Montrer que $\|A^{-1}\|_2 = \frac{1}{\lambda_1}$.

(b) On pose $M_0 = UD_0V^*$, avec $D_0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$. Mon-

trer que $\|A - M_0\|_2 = \lambda_1$.

(c) Dédurre de ce qui précède que $d(A, \mathcal{F}) = \frac{1}{\|A^{-1}\|_2}$.

Solution.

1. Pour $M \in \mathcal{F}$, 0 est valeur propre de M et il existe un vecteur unitaire x tel que $Mx = 0$. On a alors :

$$\|(A - M)x\|_2 = \|Ax\|_2 = \frac{\|A^{-1}\|_2 \|Ax\|_2}{\|A^{-1}\|_2} \geq \frac{\|A^{-1}Ax\|_2}{\|A^{-1}\|_2} = \frac{1}{\|A^{-1}\|_2}$$

donc $\frac{1}{\|A^{-1}\|_2} \leq \|A - M\|_2$ pour tout $M \in \mathcal{F}$, soit $\frac{1}{\|A^{-1}\|_2} \leq d(A, \mathcal{F})$.

2. La matrice A^*A est hermitienne définie positive. Ces valeurs propres sont donc toutes réelles strictement positives et on peut les noter $\lambda_1^2, \dots, \lambda_n^2$.

(a) On a :

$$\|A^{-1}\|_2^2 = \|(A^{-1})^* A^{-1}\|_2 = \|(AA^*)^{-1}\|_2 = \rho((AA^*)^{-1}) = \frac{1}{\min_{1 \leq k \leq n} (\lambda_k^2)}$$

ce qui donne $\|A^{-1}\|_2 = \frac{1}{\lambda_1}$.

(b) On a $M_0 \in F$ et :

$$\begin{aligned} \|A - M_0\|_2^2 &= \rho((A - M_0)(A - M_0)^*) = \rho(U(D - D_0)^2 U^*) \\ &= \rho((D - D_0)^2) = \lambda_1^2 \end{aligned}$$

c'est-à-dire que $\|A - M_0\|_2 = \lambda_1 = \frac{1}{\|A^{-1}\|_2}$.

(c) On peut donc conclure à l'égalité $d(A, \mathcal{F}) = \frac{1}{\|A^{-1}\|_2}$.

Exercice 3.18. On désigne par A la matrice réelle d'ordre n supérieur ou égal à 2 définie par :

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

1. Calculer les valeurs propres de ${}^tAA$.

2. Calculer $\|A\|_2$.

Solution.

1. On a :

$${}^tAA = \begin{pmatrix} 2 & -1 & 0 & \cdots & -1 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ -1 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Les valeurs propres de la matrice symétrique réelles ${}^tAA$ sont réelles et le théorème de Gerschgorin-Hadamard nous dit que toute valeur propre λ de cette matrice est telle que $|\lambda - 2| \leq 2$. On peut donc écrire une telle valeur propre sous la forme $\lambda = 2(1 - \cos(\theta)) = 4 \sin^2\left(\frac{\theta}{2}\right)$ avec $\theta \in [0, \pi]$. De $\det(A) = 0$

(en développant suivant la première ligne) on déduit que 0 est valeur propre de ${}^tAA$, ce qui correspond à $\theta = 0$ ou $\theta = \pi$. Si λ est une valeur propre non nulle, on a alors $\theta \in]0, \pi[$ et un vecteur propre associé x de coordonnées x_k ($1 \leq k \leq n$) est défini par les relations de récurrence :

$$x_{k-1} + (\lambda - 2)x_k + x_{k+1} = 0 \quad (1 \leq k \leq n) \quad (3.7)$$

avec les conditions aux limites $x_0 = x_n$ et $x_{n+1} = x_1$. Tout revient à chercher les suites réelles $(x_k)_{k \in \mathbb{N}}$ périodiques de période n et vérifiant la récurrence (3.7). Le polynôme caractéristique de cette récurrence, $P(r) = r^2 - 2 \cos(\theta)r + 1$ a pour racines $r_1 = e^{i\theta}$ et $r_2 = e^{-i\theta}$. On obtient donc les solutions définies par :

$$\forall k \in \mathbb{N}, x_k = \alpha e^{ik\theta} + \beta e^{-ik\theta},$$

les coefficients α et β étant tels que $x_k \in \mathbb{R}$ et $x_{k+n} = x_n$ pour tout entier naturel k . Des conditions $x_k \in \mathbb{R}$ pour tout entier k on déduit que $(\alpha - \beta) \sin(k\theta) = 0$ et $\alpha = \beta$. La condition $x_0 = x_n$ donne $2\alpha \cos(n\theta) = 2\alpha$ et, si on s'intéresse à une solution non nulle on a nécessairement $\alpha \neq 0$ et $\cos(n\theta) = 1$, ce qui donne $\theta = j \frac{2\pi}{n}$ avec $0 \leq j \leq n-1$. On a donc ainsi montré que la matrice ${}^tAA$ a n valeurs propres simples données par :

$$\lambda_j = 4 \sin^2 \left(j \frac{\pi}{n} \right) \quad (0 \leq j \leq n-1)$$

2. On a :

$$\|A\|_2 = \sqrt{\rho({}^tAA)} = \begin{cases} 2 \text{ si } n = 2p \\ 2 \sin \left(p \frac{2\pi}{2p+1} \right) \text{ si } n = 2p+1 \end{cases}$$

Exercice 3.19. Montrer les inégalités suivantes pour toute matrice A dans $\mathcal{M}_n(\mathbb{C})$:

1. $\|A\|_2 \leq \sqrt{\|A\|_\infty \|A\|_1}$.
2. $\frac{\|A\|_\infty}{\sqrt{n}} \leq \|A\|_2 \leq \sqrt{n} \|A\|_\infty$.
3. $\frac{\|A\|_1}{\sqrt{n}} \leq \|A\|_2 \leq \sqrt{n} \|A\|_1$.

Solution.

1. On note $\lambda_1^2, \dots, \lambda_n^2$ les valeurs propres de A^*A avec $0 \leq \lambda_1 \leq \dots \leq \lambda_n$ et on a $\|A\|_2 = \lambda_n$. En désignant par e_n un vecteur propre de A associé à la valeur propre λ_n et unitaire pour $\|\cdot\|_\infty$, on a :

$$\|A\|_2^2 = |\lambda_n|^2 = \|\lambda_n^2 e_n\|_\infty = \|A^* A e_n\|_\infty \leq \|A^*\|_\infty \|A\|_\infty$$

Puis avec $\|A^*\|_\infty = \|A\|_1$, on déduit que $\|A\|_2 \leq \sqrt{\|A\|_\infty \|A\|_1}$. L'égalité étant réalisée pour $A = I_n$.

2. Pour tout x dans $\mathbb{C}^n$ on a $\|Ax\|_2^2 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \leq \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}| |x_j| \right)^2$ et, en utilisant l'inégalité de Cauchy-Schwarz, on déduit que :

$$\|Ax\|_2^2 \leq \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}|^2 \right) \left(\sum_{j=1}^n |x_j|^2 \right) = \|x\|_2^2 \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}|^2 \right)$$

Il en résulte que :

$$\|A\|_2^2 \leq \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}|^2 \right) \leq \sum_{i=1}^n \left(\sum_{j=1}^n |a_{ij}| \right)^2 \leq n \left(\max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \right)^2$$

soit $\|A\|_2 \leq \sqrt{n} \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \sqrt{n} \|A\|_\infty$. D'autre part il existe un indice i tel

que $\|A\|_\infty = \sum_{j=1}^n |a_{ij}|$ et en utilisant l'inégalité de Cauchy-Schwarz on obtient :

$$\|A\|_\infty = \sum_{j=1}^n |a_{ij}| \cdot 1 \leq \sqrt{n} \sqrt{\sum_{j=1}^n |a_{ij}|^2}$$

En désignant par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{C}^n$ on a $\sqrt{\sum_{j=1}^n |a_{ij}|^2} = \|A^* e_i\|_2$,

ce qui donne en définitive, $\|A\|_\infty \leq \sqrt{n} \|A^* e_i\|_2 \leq \sqrt{n} \|A^*\|_2 = \sqrt{n} \|A\|_2$.

3. Ces inégalités se déduisent des précédentes avec $\|A^*\|_\infty = \|A\|_1$, $\|A^*\|_2 = \|A\|_2$.

Exercice 3.20. Soient α, β des réels et $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice réelle d'ordre n supérieur ou égal à 3 définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, \dots, n\} \setminus \{i\} \end{cases}$$

1. Calculer le rayon spectral $\rho(\alpha, \beta)$ de $A(\alpha, \beta)$.
2. On suppose que $\alpha\beta \geq 0$. Comparer $\|A(\alpha, \beta)\|_\infty$, $\|A(\alpha, \beta)\|_1$ et $\|A(\alpha, \beta)\|_2$.

Solution.

1. On a vu (exercice 1.7) que les valeurs propres de la matrice $A(\alpha, \beta)$ sont données par $\lambda_1 = \beta + (n-1)\alpha$ et $\lambda_2 = \beta - \alpha$. Le rayon spectral de $A(\alpha, \beta)$ est donc $\rho(\alpha, \beta) = \max(|\beta + (n-1)\alpha|, |\beta - \alpha|)$, ce qui donne quatre possibilités.

(i) $\alpha \leq \beta$, $\beta > -(n-1)\alpha$. Alors :

$$\rho(\alpha, \beta) = \max(\beta + (n-1)\alpha, \beta - \alpha) = \begin{cases} \beta + (n-1)\alpha & \text{si } \alpha > 0 \\ \beta - \alpha & \text{si } \alpha \leq 0 \end{cases}$$

(ii) $\alpha \leq \beta$, $\beta \leq -(n-1)\alpha$. Alors :

$$\begin{aligned}\rho(\alpha, \beta) &= \max(-\beta - (n-1)\alpha, \beta - \alpha) \\ &= \begin{cases} -\beta - (n-1)\alpha & \text{si } \beta < -\frac{n-2}{2}\alpha \\ \beta - \alpha & \text{si } \beta \geq -\frac{n-2}{2}\alpha \end{cases}\end{aligned}$$

(iii) $\alpha > \beta$, $\beta > -(n-1)\alpha$. Alors :

$$\rho(\alpha, \beta) = \max(\beta + (n-1)\alpha, \alpha - \beta) = \begin{cases} \beta + (n-1)\alpha & \text{si } \beta > -\frac{n-2}{2}\alpha \\ \alpha - \beta & \text{si } \beta \leq -\frac{n-2}{2}\alpha \end{cases}$$

(iv) $\alpha > \beta$, $\beta \leq -(n-1)\alpha$. Alors :

$$\rho(\alpha, \beta) = \max(-\beta - (n-1)\alpha, \alpha - \beta) = \begin{cases} -\beta - (n-1)\alpha & \text{si } \alpha < 0, \\ \alpha - \beta & \text{si } \alpha \geq 0. \end{cases}$$

2. La matrice $A(\alpha, \beta)$ est symétrique réelle, donc $\rho(\alpha, \beta) = \|A(\alpha, \beta)\|_2$. Pour $\alpha\beta \geq 0$, on a :

$$\rho(\alpha, \beta) = |\beta| + (n-1)|\alpha| = \|A(\alpha, \beta)\|_2 = \|A(\alpha, \beta)\|_\infty = \|A(\alpha, \beta)\|_1$$

Exercice 3.21. Calculer le rayon spectral et le conditionnement pour la norme matricielle induite par la norme euclidienne de la matrice :

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

Solution. On a vu (exercice 1.9) que les valeurs propres de A sont :

$$\lambda_k = 2 - 2 \cos\left(\frac{k\pi}{n+1}\right) = 4 \sin^2\left(\frac{k\pi}{2(n+1)}\right) \quad (1 \leq k \leq n)$$

On a $\rho(A) = 4 \sin^2\left(\frac{n\pi}{2(n+1)}\right)$ et le conditionnement pour la norme matricielle

induite par la norme euclidienne par $\text{cond}_2(A) = \frac{\lambda_1}{\lambda_n} = \frac{\sin^2\left(\frac{\pi}{2(n+1)}\right)}{\sin^2\left(\frac{n\pi}{2(n+1)}\right)}$ (la matrice

A est symétrique définie positive). On peut remarquer que $\lim_{n \rightarrow +\infty} \text{cond}_2(A) = 0$, c'est-à-dire que la matrice A est mal conditionnée pour les grandes valeurs de n .

Exercice 3.22. Pour tout entier n supérieur ou égal à 2 l'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique avec le produit scalaire

$$(x, y) \mapsto \langle x | y \rangle = \sum_{i=1}^n x_i y_i \text{ de norme associée notée } \|\cdot\|_2. \text{ On désigne par } H_n$$

la matrice de Hilbert d'ordre n définie par $H_n = \left(\left(\frac{1}{i+j-1} \right) \right)_{1 \leq i, j \leq n}$.

1. Montrer que H_n est la matrice dans la base canonique de $\mathbb{R}_{n-1}[X]$ du produit scalaire défini par :

$$(P, Q) \mapsto \langle P | Q \rangle = \int_0^1 P(t) Q(t) dt$$

2. Qu'en déduit-on pour les valeurs propres de H_n .

On désigne par R_n le quotient de Rayleigh-Ritz associé à la matrice H_n et par ρ_n son rayon spectral.

Dans ce qui suit on écrit H_{n+1} sous la forme $H_{n+1} = \begin{pmatrix} H_n & C_n \\ {}^t C_n & \frac{1}{2n+1} \end{pmatrix}$

avec $C_n = \begin{pmatrix} \frac{1}{n+1} \\ \vdots \\ \frac{1}{2n} \end{pmatrix}$. Un vecteur x' de $\mathbb{R}^{n+1}$ est écrit sous la forme

$$x' = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \text{ avec } x \in \mathbb{R}^n \text{ et } x_{n+1} \in \mathbb{R}.$$

3. Montrer que la suite $(\rho_n)_{n \geq 2}$ est croissante.
4. Soient $x' \in \mathbb{R}^{n+1}$ tel que $\|x'\|_2 = 1$ et $x \neq 0$. Montrer que :

$$R_{n+1}(x') \leq \rho_n \|x\|_2^2 + 2|x_{n+1}| \|C_n\|_2 \|x\|_2 + \frac{1}{2n+1} x_{n+1}^2$$

5. On désigne par q la forme quadratique définie sur $\mathbb{R}^2$ par :

$$\forall (u, v) \in \mathbb{R}^2, q(u, v) = \rho_n u^2 + 2 \|C_n\|_2 uv + \frac{1}{2n+1} v^2$$

Calculer $\sup_{\substack{(u,v) \in \mathbb{R}^2 \\ u^2+v^2=1}} q(u, v)$.

6. Dédire de ce qui précède que $\rho_{n+1} \leq \rho_n + \|C_n\|_2 \leq \rho_n + \frac{1}{\sqrt{2n}}$.

Solution. On désigne par $(e_k)_{0 \leq k \leq n-1}$ la base canonique de $\mathbb{R}_{n-1}[X]$ définie par :

$$\forall i \in \{0, \dots, n-1\}, e_i(X) = X^i$$

1. La matrice du produit scalaire $(P, Q) \mapsto \langle P | Q \rangle$ dans la base canonique de $\mathbb{R}_{n-1}[X]$ est la matrice de coefficients :

$$\langle e_{i-1} | e_{j-1} \rangle = \int_0^1 t^{i+j-2} dt = \frac{1}{i+j-1} \quad (1 \leq i, j \leq n)$$

C'est bien la matrice de Hilbert d'ordre n .

2. On en déduit alors que la matrice réelle H_n est symétrique définie positive et ses valeurs propres sont réelles strictement positives.
3. La matrice H_n étant symétrique réelle de plus grande valeur propre ρ_n , on a $\rho_n = \sup_{x \in \mathbb{R}^n \setminus \{0\}} R_n(x)$ (théorème 3.21). Soient $x \in \mathbb{R}^n \setminus \{0\}$ et $x' \in \mathbb{R}^{n+1}$ défini par $x'_j = x_j$ pour j compris entre 1 et n et $x_{n+1} = 0$. On a $x' \neq 0$ et $R_n(x) = R_{n+1}(x') \leq \rho_{n+1}$. Le vecteur x étant quelconque dans $\mathbb{R}^n \setminus \{0\}$ on en déduit que $\rho_n \leq \rho_{n+1}$. La suite $(\rho_n)_{n \geq 2}$ est donc croissante.

4. On a $\langle H_{n+1}x' | x' \rangle = \langle H_nx + x_{n+1}C_n | x \rangle + x_{n+1} \left(\langle C_n | x \rangle + \frac{x_{n+1}}{2n+1} \right)$, soit :

$$\langle H_{n+1}x' | x' \rangle = \langle H_nx | x \rangle + 2x_{n+1} \langle C_n | x \rangle + \frac{x_{n+1}^2}{2n+1}$$

Et pour $x' \in \mathbb{R}^{n+1}$ tel que $\|x'\|_2 = 1$ et $x \neq 0$ cela s'écrit :

$$R_{n+1}(x') = R_n(x) \|x\|_2^2 + 2x_{n+1} \langle C_n | x \rangle + \frac{x_{n+1}^2}{2n+1}$$

En utilisant l'inégalité de Cauchy-Schwarz on a :

$$x_{n+1} \langle C_n | x \rangle \leq |x_{n+1}| |\langle C_n | x \rangle| \leq |x_{n+1}| \|C_n\|_2 \|x\|_2$$

et, avec $\rho_n = \sup_{x \in \mathbb{R}^n \setminus \{0\}} R_n(x)$, on aboutit à :

$$R_{n+1}(x') \leq \rho_n \|x\|_2^2 + 2|x_{n+1}| \|C_n\|_2 \|x\|_2 + \frac{1}{2n+1} x_{n+1}^2$$

5. La matrice de la forme quadratique q dans la base canonique de $\mathbb{R}^2$ est donnée par $A = \begin{pmatrix} \rho_n & \|C_n\|_2 \\ \|C_n\|_2 & \frac{1}{2n+1} \end{pmatrix}$. Ces valeurs propres sont :

$$\frac{1}{2} \left(\rho_n + \frac{1}{2n+1} \pm \sqrt{\left(\rho_n - \frac{1}{2n+1} \right)^2 + 4\|C_n\|_2^2} \right)$$

On en déduit que :

$$\sup_{\substack{(u,v) \in \mathbb{R}^2 \\ u^2+v^2=1}} q(u,v) = \frac{1}{2} \left(\rho_n + \frac{1}{2n+1} + \sqrt{\left(\rho_n - \frac{1}{2n+1} \right)^2 + 4\|C_n\|_2^2} \right)$$

6. Soit $x' \in \mathbb{R}^{n+1}$ tel que $R_{n+1}(x') = \rho_{n+1}$ et $\|x'\|_2 = 1$. On a $x' = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}$ avec $x \neq 0$ et $\rho_{n+1} \leq q(\|x\|_2, |x_{n+1}|) \leq \sup_{\substack{(u,v) \in \mathbb{R}^2 \\ u^2+v^2=1}} q(u,v)$ d'après ce qui précède.

En écrivant que :

$$\sqrt{\left(\rho_n - \frac{1}{2n+1}\right)^2 + 4\|C_n\|_2^2} \leq \left|\rho_n - \frac{1}{2n+1}\right| + 2\|C_n\|_2$$

et en remarquant que $\rho_n \geq \rho_2 = \frac{4+\sqrt{13}}{6} \geq 1$, on déduit que :

$$\left|\rho_n - \frac{1}{2n+1}\right| = \rho_n - \frac{1}{2n+1}$$

et $\rho_{n+1} \leq \rho_n + \|C_n\|_2$. Enfin avec :

$$\|C_n\|_2^2 = \sum_{k=1}^n \frac{1}{(n+k)^2} \leq \int_n^{2n} \frac{dt}{t^2} = \frac{1}{2n}$$

on déduit que $\rho_{n+1} \leq \rho_n + \frac{1}{\sqrt{2n}}$.

Exercice 3.23. On reprend les notations de l'exercice 3.22. On munit l'espace vectoriel $\mathbb{R}[X]$ du produit scalaire défini par :

$$(P, Q) \mapsto \langle P | Q \rangle = \int_0^1 P(t) Q(t) dt$$

la norme associée étant notée $\|\cdot\|$.

On note $(e_k)_{k \in \mathbb{N}}$ la base canonique de $\mathbb{R}[X]$ définie par :

$$\forall k \in \mathbb{N}, e_k(X) = X^k$$

Pour tout $n \in \mathbb{N}^*$, on désigne par δ_n la distance de e_n à $\mathbb{R}_{n-1}[X]$. On rappelle qu'il existe un unique polynôme $P \in \mathbb{R}_{n-1}[X]$ tel que $\delta_n = \|e_n - P\|$ et que ce polynôme est aussi caractérisé par $e_n - P \in (\mathbb{R}_{n-1}[X])^\perp$. On

note $P(X) = \sum_{j=0}^{n-1} a_j X^j$ et on associe à ce polynôme la fonction rationnelle

$$f \text{ définie par } f(t) = \frac{1}{t+n+1} - \sum_{j=0}^{n-1} \frac{a_j}{t+j+1}.$$

1. Calculer $f(k)$ pour tout entier k compris entre 0 et n .

2. En écrivant la fonction f sous la forme $f(t) = \frac{Q(t)}{\prod_{j=1}^{n+1} (t+j)}$, en déduire

la valeur de δ_n .

3. On désigne par μ_n la plus petite valeur propre de la matrice de Hilbert

$$H_n \text{ et on note } x = \begin{pmatrix} -a_0 \\ \vdots \\ -a_{n-1} \\ 1 \end{pmatrix}.$$

(a) Montrer que $\mu_{n+1} \leq \frac{\delta_n^2}{\|x\|_2^2}$.

(b) Montrer que $\mu_n \leq \frac{1}{2n-1} \frac{1}{\binom{2(n-1)}{n-1}^2} \leq \frac{1}{12} \frac{1}{15^{n-1}}$.

(c) Montrer que $\text{cond}_2(H_n) \leq \frac{1}{2(4+\sqrt{13})} \frac{1}{15^{n-1}}$.

Solution.

1. Pour tout entier naturel k on a :

$$f(k) = \int_0^1 t^{n+k} dt - \sum_{j=0}^{n-1} a_j \int_0^1 t^{j+k} dt = \langle e_n - P \mid e_k \rangle$$

De $e_n - P \in (\mathbb{R}_{n-1}[X])^\perp$ on déduit que $f(k) = 0$ pour tout entier k compris entre 1 et $n-1$ et $f(n) = \langle e_n - P \mid e_n \rangle = \|e_n - P\|^2 = \delta_n^2$.

2. Réduisant au même dénominateur, on a $f(t) = \frac{Q(t)}{\prod_{j=1}^{n+1} (t+j)}$ avec $Q \in \mathbb{R}_n[X]$.

La fonction f s'annulant en $0, 1, \dots, n-1$, il en est de même de Q et ce dernier s'écrit $Q(t) = \alpha_n \prod_{j=0}^{n-1} (t-j)$. De $\lim_{t \rightarrow -(n+1)} (t+n+1)f(t) = 1$, on déduit que

$$\alpha_n = \frac{(n!)^2}{(2n)!} \text{ et } \delta_n^2 = f(n) = \frac{(n!)^4}{(2n+1)((2n)!)^2}, \text{ soit } \delta_n = \frac{(n!)^2}{\sqrt{2n+1}(2n)!}.$$

3. On désigne par R_{n+1} le quotient de Rayleigh-Ritz associé à la matrice symétrique réelle H_{n+1} .

(a) On a $\mu_{n+1} = \inf_{y \in \mathbb{R}^{n+1} \setminus \{0\}} R_{n+1}(y) \leq R_{n+1}(x)$ (théorème 3.21) avec :

$$R_{n+1}(x) = \frac{\langle H_{n+1}x \mid x \rangle}{\|x\|_2^2} = \frac{t x H_{n+1} x}{\|x\|_2^2} = \frac{\|e_n - P\|^2}{\|x\|_2^2} = \frac{\delta_n^2}{\|x\|_2^2}$$

(H_{n+1} est la matrice dans la base canonique du produit scalaire de $\mathbb{R}_n[X]$).

On a donc bien $\mu_{n+1} \leq \frac{\delta_n^2}{\|x\|_2^2}$.

(b) Du fait que $\|x\|_2 \geq 1$ on déduit de ce qui précède que :

$$\mu_{n+1} \leq \delta_n^2 = \frac{(n!)^4}{(2n+1)((2n)!)^2} = \frac{1}{2n+1} \frac{1}{\binom{2n}{n}^2}$$

D'autre part avec $\frac{\delta_{n+1}^2}{\delta_n^2} = \frac{(n+1)^2}{4(2n+1)(2n+3)}$ et avec :

$$4(2n+1)(2n+3) = 16n^2 + 32n + 12 \geq 15(n^2 + 2n + 1)$$

on déduit que $\delta_{n+1}^2 \leq \frac{1}{15}\delta_n^2$ et $\delta_n^2 \leq \left(\frac{1}{15}\right)^{n-1} \delta_1^2 = \frac{1}{12} \left(\frac{1}{15}\right)^{n-1}$. On a donc $\mu_n \leq \frac{1}{12} \frac{1}{15^{n-1}}$.

(c) La matrice H_n étant symétrique réelle et la suite $(\rho_n)_{n \geq 2}$ étant croissante (exercice 3.22), on déduit que :

$$\text{cond}_2(H_n) = \frac{\mu_n}{\rho_n} \leq \frac{\mu_n}{\rho_2} \leq \frac{1}{2(4 + \sqrt{13})} \frac{1}{15^{n-1}}$$

Exercice 3.24. On propose une autre démonstration du théorème de Householder (théorème 3.14). Soit $A \in \mathcal{M}_n(\mathbb{C})$.

1. Montrer que pour tout $\varepsilon > 0$, l'application $x \mapsto \|x\|_\varepsilon = \sum_{k=0}^{+\infty} \frac{\|A^k x\|_\infty}{(\rho(A) + \varepsilon)^k}$ est bien définie et que c'est une norme sur $\mathbb{C}^n$.
2. Montrer que $\|A\|_\varepsilon < \rho(A) + \varepsilon$, puis en déduire que $\rho(A) = \inf_{\|\cdot\| \in \mathcal{N}} \|A\|$, où $\mathcal{N}$ désigne l'ensemble de toutes les normes matricielles induites par une norme vectorielle.

Solution.

1. En notant $u_k = \frac{\|A^k x\|_\infty}{(\rho(A) + \varepsilon)^k}$ et $v_k = \frac{\|A^k\|_\infty}{(\rho(A) + \varepsilon)^k}$, on a $0 \leq u_k \leq v_k \|x\|_\infty$ avec $\lim_{k \rightarrow +\infty} \sqrt[k]{v_k} = \lim_{k \rightarrow +\infty} \frac{\|A^k\|_\infty^{\frac{1}{k}}}{\rho(A) + \varepsilon} = \frac{\rho(A)}{\rho(A) + \varepsilon} \in [0, 1[$. On déduit alors du critère de Cauchy que la série à termes positifs $\sum v_k$ est convergente et il en est de même de $\sum u_k$. Avec $\|x\|_\infty \leq \|x\|_\varepsilon$, on déduit que $\|x\|_\varepsilon = 0$ si, et seulement si, $x = 0$. On vérifie facilement que les autres propriétés d'une norme sont bien vérifiées.
2. Pour tout $x \in \mathbb{C}^n \setminus \{0\}$, on a :

$$\begin{aligned} \|Ax\|_\varepsilon &= \sum_{k=0}^{+\infty} \frac{\|A^{k+1}x\|_\infty}{(\rho(A) + \varepsilon)^k} = (\rho(A) + \varepsilon) \sum_{k=1}^{+\infty} \frac{\|A^k x\|_\infty}{(\rho(A) + \varepsilon)^k} \\ &= (\rho(A) + \varepsilon) (\|x\|_\varepsilon - \|x\|_\infty) < \rho(A) + \varepsilon \end{aligned}$$

On a donc ainsi montré que pour tout $A \in \mathcal{M}_n(\mathbb{C})$ et tout $\varepsilon > 0$, il existe une norme matricielle $\|\cdot\|$ sur $\mathcal{M}_n(\mathbb{C})$ induite par une norme vectorielle $\|\cdot\|$ sur $\mathbb{C}^n$ telle que $\|A\| \leq \rho(A) + \varepsilon$. Comme $\rho(A) \leq \|A\|$ pour toute norme matricielle

induite par une norme vectorielle, cela se traduit par $\rho(A) = \inf_{\|\cdot\| \in \mathcal{N}} \|A\|$, où $\mathcal{N}$ désigne l'ensemble de toutes les normes matricielles induites par une norme vectorielle.

Exercice 3.25. On propose une autre démonstration de la continuité du rayon spectral.

1. En notant, pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, χ_A son polynôme caractéristique, montrer que l'application $\varphi : (A, z) \in \mathcal{M}_n(\mathbb{C}) \times \mathbb{C} \mapsto \chi_A(z)$ est lipschitzienne sur tout compact de $\mathcal{M}_n(\mathbb{C}) \times \mathbb{C}$.
2. Soient $B(0, R)$ une boule fermée de centre 0 et de rayon $R > 0$ dans $(\mathcal{M}_n(\mathbb{C}), N)$, où N est une norme sur $\mathcal{M}_n(\mathbb{C})$. Montrer qu'il existe une constante C_R telle que pour toutes matrices A, B dans $B(0, R)$, il existe une valeur propre $\mu \in \text{Sp}(B)$ telle que $|\rho(A) - \mu|^n \leq C_R N(B - A)$.
3. En déduire la continuité du rayon spectral sur $\mathcal{M}_n(\mathbb{C})$.

Solution.

1. Identifiant $\mathcal{M}_n(\mathbb{C}) \times \mathbb{C}$ à $\mathbb{C}^{n^2+1}$, la fonction φ est polynomiale, donc de classe $\mathcal{C}^\infty$ et en conséquence lipschitzienne sur tout compact de $\mathcal{M}_n(\mathbb{C}) \times \mathbb{C}$. En se donnant une norme N sur $\mathcal{M}_n(\mathbb{C})$, il existe donc pour tout compact K de $\mathcal{M}_n(\mathbb{C}) \times \mathbb{C}$ une constante $C_K > 0$ telle que $|\chi_A(z) - \chi_B(z')| \leq C_K (N(B - A) + |z' - z|)$ pour tous (A, z) et (B, z') dans $\mathcal{M}_n(\mathbb{C}) \times \mathbb{C}$.
2. En prenant $\lambda \in \text{Sp}(A)$ telle que $|\lambda| = \rho(A)$, on a :

$$|\chi_B(\lambda)| = \left| \prod_{\nu \in \text{Sp}(B)} (\lambda - \nu) \right| = |\chi_B(\lambda) - \chi_A(\lambda)| \leq C_R N(B - A)$$

où C_R est la constante C_K correspondante au compact $K = B(0, R) \times \{\lambda\}$ de la question précédente, ce qui nous donne pour $\mu \in \text{Sp}(B)$ telle que $|\lambda - \mu| = \min_{\nu \in \text{Sp}(B)} |\lambda - \nu|$:

$$|\lambda - \mu|^n \leq \prod_{\nu \in \text{Sp}(B)} |\lambda - \nu| \leq C_R N(B - A)$$

3. Avec les notations précédentes, on a $|\lambda| - |\mu| \leq |\lambda - \mu| \leq C_R^{\frac{1}{n}} (N(B - A))^{\frac{1}{n}}$, donc :

$$\rho(A) = |\lambda| \leq |\mu| + C_R^{\frac{1}{n}} (N(B - A))^{\frac{1}{n}} \leq \rho(B) + C_R^{\frac{1}{n}} (N(B - A))^{\frac{1}{n}}$$

A et B jouant des rôles symétriques, on a aussi $\rho(B) \leq \rho(A) + C_R^{\frac{1}{n}} (N(B - A))^{\frac{1}{n}}$, ce qui nous donne en définitive $|\rho(B) - \rho(A)| \leq C_R^{\frac{1}{n}} (N(B - A))^{\frac{1}{n}}$, pour toutes matrices A, B dans $B(0, R)$. La fonction ρ est donc höldérienne sur toute boule fermée $B(0, R)$ de $\mathcal{M}_n(\mathbb{C})$ et en conséquence continue sur $\mathcal{M}_n(\mathbb{C})$.

Exercice 3.26. En utilisant la décomposition de Dunford $A = D + N$ avec D diagonalisable, N nilpotente et $DN = ND$, montrer le théorème de Gelfand (théorème 3.15).

Solution. Par équivalence des normes sur $\mathcal{M}_n(\mathbb{C})$, il nous suffit de se placer dans le cas où $N = \|\cdot\|_\infty$. Pour $A \in \mathcal{M}_n(\mathbb{C})$, on a la décomposition de Dunford $A = D + N$ avec D diagonalisable de mêmes valeurs propres que A et N nilpotente qui commute avec D . Pour tout entier $k \geq n$, on a $N^k = 0$ (N est nilpotente d'ordre $r \leq n$) et :

$$A^k = (D + N)^k = \sum_{j=0}^n \binom{k}{j} D^{k-j} N^j = D^{k-n} \sum_{j=0}^n \binom{k}{j} D^{n-j} N^j$$

ce qui nous donne en notant $\alpha = \max_{0 \leq j \leq n} \|D^{n-j} N^j\|_\infty$:

$$\|A^k\|_\infty \leq \|D^{k-n}\|_\infty \sum_{j=0}^n \binom{k}{j} \|D^{n-j} N^j\|_\infty \leq \alpha \left(\sum_{j=0}^n \binom{k}{j} \right) \|D^{k-n}\|_\infty$$

Pour tout j compris entre 0 et n , on a $\binom{k}{j} \leq k(k-1)\cdots(k-j+1) \leq k^j \leq k^n$, donc $\|A^k\|_\infty \leq \alpha(n+1)k^n \|D^{k-n}\|_\infty$ et :

$$\rho(A) \leq \|A^k\|_\infty^{\frac{1}{k}} \leq (\alpha(n+1))^{\frac{1}{k}} k^{\frac{n}{k}} \|D^{k-n}\|_\infty^{\frac{1}{k}}$$

Comme D est diagonalisable, il existe $P \in GL_n(\mathbb{C})$ et $\Delta = \text{diag}(\lambda_1, \dots, \lambda_n)$ telles que $D = P\Delta P^{-1}$, de sorte que :

$$\|D^{k-n}\|_\infty = \|P\Delta^{k-n}P^{-1}\|_\infty \leq \|P\|_\infty \|P^{-1}\|_\infty \|\Delta^{k-n}\|_\infty = \beta \|\Delta^{k-n}\|_\infty$$

avec $\|\Delta^{k-n}\|_\infty = \|\text{diag}(\lambda_1^{k-n}, \dots, \lambda_n^{k-n})\|_\infty = \max_{1 \leq i \leq n} |\lambda_i|^{k-n} = (\rho(A))^{k-n}$ (D et A ont les mêmes valeurs propres). On a donc, pour tout $k \geq n$:

$$\rho(A) \leq \|A^k\|_\infty^{\frac{1}{k}} \leq (\alpha\beta(n+1))^{\frac{1}{k}} k^{\frac{n}{k}} (\rho(A))^{\frac{k-n}{k}}$$

avec $\lim_{k \rightarrow +\infty} (\alpha\beta(n+1))^{\frac{1}{k}} k^{\frac{n}{k}} = 1$ et $\lim_{k \rightarrow +\infty} (\rho(A))^{\frac{k-n}{k}} = \rho(A)$. On déduit alors du théorème d'encadrement que $\lim_{k \rightarrow +\infty} \left(\|A^k\|_\infty^{\frac{1}{k}} \right) = \rho(A)$.

Exercice 3.27. Pour toute matrice A dans $\mathcal{M}_n(\mathbb{C})$ on désigne par $\mathcal{H}(A)$ le hausdorffien de A .

1. Montrer que si A est diagonale, $\mathcal{H}(A)$ est alors l'enveloppe convexe du spectre de A .

2. Montrer que pour toute matrice A dans $\mathcal{M}_n(\mathbb{C})$ et toute matrice unitaire U , on a $\mathcal{H}(U^*AU) = \mathcal{H}(A)$.
3. Montrer que pour toute matrice normale A , $\mathcal{H}(A)$ est l'enveloppe convexe du spectre de A .

Solution.

1. Soit A une matrice diagonale de valeurs propres $\lambda_1, \dots, \lambda_n$. Pour tout vecteur x dans la sphère unité de $(\mathbb{C}^n, \|\cdot\|_2)$, on a $R_A(x) = \langle Ax \mid x \rangle = \sum_{i=1}^n |x_i|^2 \lambda_i$ avec $0 \leq |x_i|^2 \leq 1$ et $\sum_{i=1}^n |x_i|^2 = 1$. Il en résulte que $\mathcal{H}(A)$ est l'enveloppe convexe du spectre de A .
2. Pour tout vecteur x dans la sphère unité S_1 de $(\mathbb{C}^n, \|\cdot\|_2)$, on a :

$$\langle U^*AUx \mid x \rangle = \langle Ay \mid y \rangle$$

avec $y = Ux$ qui décrit tout S_1 quand x décrit S_1 pour U unitaire. On déduit alors que $\mathcal{H}(U^*AU) = \mathcal{H}(A)$.

3. Si A est une matrice normale, elle se diagonalise alors dans une base ortho-normée, c'est-à-dire qu'il existe une matrice unitaire U telle que la matrice $D = U^*AU$ soit diagonale. On a alors $\mathcal{H}(A) = \mathcal{H}(D)$ qui est égale à l'enveloppe convexe du spectre de D . Comme A et D ont même spectre on déduit que $\mathcal{H}(A)$ est l'enveloppe convexe du spectre de A . Dans le cas où la matrice A est hermitienne, on retrouve l'égalité $\mathcal{H}(A) = [\lambda_1, \lambda_n]$ où λ_1 est la plus petite valeur propre de A et λ_n la plus grande.

Exercice 3.28. Soit $A : [0, 1] \rightarrow S_2(\mathbb{R})$ définie par :

$$\forall t \in [0, 1], A(t) = \begin{pmatrix} a(t) & b(t) \\ b(t) & -a(t) \end{pmatrix}$$

avec :

$$a(t) = \begin{cases} 0 & \text{si } t = 0 \\ e^{-\frac{1}{t^2}} \cos\left(\frac{1}{t}\right) & \text{si } t \neq 0 \end{cases}, \quad b(t) = \begin{cases} 0 & \text{si } t = 0 \\ e^{-\frac{1}{t^2}} \sin\left(\frac{1}{t}\right) & \text{si } t \neq 0 \end{cases}$$

1. Montrer que A est continue et pour tout t dans $[0, 1]$ calculer les valeurs et vecteurs propres de $A(t)$.
2. Montrer que les vecteurs propres de $A(t)$ ne définissent pas des fonctions continues en 0.

Solution.

1. La fonction A est continue sur $[0, 1]$ et pour tout t dans $[0, 1]$, les valeurs propres de $A(t)$ sont données par :

$$\lambda_1(t) = \begin{cases} 0 & \text{si } t = 0 \\ e^{-\frac{1}{t^2}} & \text{si } t \neq 0 \end{cases}, \quad \lambda_2(t) = -\lambda_1(t)$$

Ce sont des fonctions continues sur $[0, 1]$. On peut leur associer les fonctions vecteurs propres suivantes :

$$\forall t \in]0, 1], e_1(t) = \begin{pmatrix} \cos\left(\frac{1}{2t}\right) \\ \sin\left(\frac{1}{2t}\right) \end{pmatrix}, e_2(t) = \begin{pmatrix} \sin\left(\frac{1}{2t}\right) \\ -\cos\left(\frac{1}{2t}\right) \end{pmatrix}$$

$$e_1(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Il est clair que ces fonctions vecteurs propres ne sont pas continues en 0.

Exercice 3.29. Soient A, B deux matrices hermitiennes et $[a, b]$ un intervalle réel. Pour tout $t \in [a, b]$ on pose $A(t) = A + tB$. Montrer que la fonction qui à tout réel t dans $[a, b]$ associe la plus petite valeur propre de $A(t)$ est concave.

Solution. Pour tout $t \in [a, b]$, on désigne par $(e_k(t))_{1 \leq k \leq n}$ une base orthonormée de vecteurs propres associée aux valeurs propres (réelles) $\lambda_1(t), \dots, \lambda_n(t)$ de $A(t)$ avec $A(t)e_k(t) = \lambda_k(t)e_k(t)$ pour tout $k = 1, \dots, n$. Pour tout $x = \sum_{j=1}^n x_j e_j(t)$ de norme 1 et tout $t \in [a, b]$, on a :

$$\langle A(t)x | x \rangle = \sum_{j=1}^n \lambda_j(t) x_j^2 \geq \lambda_1(t) \sum_{j=1}^n x_j^2 = \lambda_1(t)$$

C'est-à-dire que :

$$\forall t \in [a, b], \langle Ax | x \rangle + t \langle Bx | x \rangle \geq \lambda_1(t)$$

En prenant $x = e_1(t_0)$ pour $t_0 \in [a, b]$, on a :

$$\begin{cases} \lambda_1(t_0) = \langle Ae_1(t_0) | e_1(t_0) \rangle + t_0 \langle Be_1(t_0) | e_1(t_0) \rangle \\ \forall t \in [a, b], \lambda_1(t) \leq \langle Ae_1(t_0) | e_1(t_0) \rangle + t \langle Be_1(t_0) | e_1(t_0) \rangle \end{cases}$$

Ce qui peut s'écrire, en posant $b_0 = \langle Be_1(t_0) | e_1(t_0) \rangle$:

$$\forall t \in [a, b], \lambda_1(t) \leq \lambda_1(t_0) + b_0(t - t_0)$$

Pour u, v dans $[a, b]$ et $\theta \in [0, 1]$, on a alors :

$$\begin{cases} \lambda_1(u) \leq \lambda_1(u + \theta(v - u)) - \theta(v - u)b_{\theta, u, v} \\ \lambda_1(v) \leq \lambda_1(u + \theta(v - u)) + (1 - \theta)(v - u)b_{\theta, u, v} \end{cases}$$

et $(1 - \theta)\lambda_1(u) + \theta\lambda_1(v) \leq \lambda_1(u + \theta(v - u))$, c'est-à-dire que la fonction λ_1 est concave.

Chapitre 4

Matrices positives et irréductibles

Pour ce chapitre, $\mathbb{K}$ est le corps des réels ou des complexes.

Pour $A = ((a_{i,j}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$ et $x = (x_i)_{1 \leq i \leq m}$ dans $\mathbb{K}^m$, on note $(Ax)_i$ la composante numéro i du vecteur Ax .

Pour $A = ((a_{i,j}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$, on note $|A| = ((|a_{i,j}|))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{R})$.

Pour toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$ et pour $1 \leq p < q \leq n$, on note $A_{p,q} = ((a_{ij}))_{p \leq i,j \leq q}$ une sous-matrice principale de A .

On note $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ une matrice diagonale D dans $\mathcal{M}_n(\mathbb{K})$ de termes diagonaux $\lambda_1, \dots, \lambda_n$.

4.1 Matrices positives

Définition 4.1. Une matrice A dans $\mathcal{M}_{n,m}(\mathbb{R})$ est dite positive [resp. strictement positive] et on note $A \geq 0$ [resp. $A > 0$], si tous ses coefficients sont positifs ou nuls [resp. strictement positifs].

Si A, B sont deux matrices dans $\mathcal{M}_{n,m}(\mathbb{R})$ la notation $A \geq B$ [resp. $A > B$, ou $A \leq B$, ou $A < B$] signifie que la matrice $A - B$ est positive [resp. $A - B$ est strictement positive, ou $B - A$ est positive ou $B - A$ est strictement positive].

En considérant la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, on voit que les conditions $A \geq 0$ et $A \neq 0$ n'entraînent pas $A > 0$.

Lemme 4.1 Si $(z_k)_{1 \leq k \leq n} \in \mathbb{C}^n$ est tel que $\left| \sum_{k=1}^n z_k \right| = \sum_{k=1}^n |z_k|$, il existe alors un réel $\theta \in]-\pi, \pi]$ tel que $z_k = |z_k| e^{i\theta}$ pour tout k compris entre 1 et n .

Preuve. Dans le cas où tous les z_k sont nuls, n'importe quel réel θ convient. Dans le cas contraire, en désignant par I l'ensemble des indices k compris entre 1 et n tels que $z_k \neq 0$ et en écrivant pour tout $k \in I$, $z_k = \rho_k e^{i\theta_k}$ avec $\rho_k = |z_k| > 0$, $\theta_k \in]-\pi, \pi]$, on a :

$$\begin{cases} \left| \sum_{k \in I} z_k \right|^2 = \sum_{k \in I} |z_k|^2 + \sum_{j \neq k} \rho_j \rho_k \cos(\theta_j - \theta_k) \\ \left(\sum_{k \in I} |z_k| \right)^2 = \sum_{k \in I} |z_k|^2 + \sum_{j \neq k} \rho_j \rho_k \end{cases}$$

et l'égalité $\left| \sum_{k=1}^n z_k \right| = \sum_{k=1}^n |z_k|$ équivaut à $\sum_{\substack{(j,k) \in I^2 \\ j \neq k}} \rho_j \rho_k (1 - \cos(\theta_j - \theta_k)) = 0$. Tous

les termes de cette somme étant positifs ou nuls avec $\rho_j \rho_k > 0$, on en déduit que $\cos(\theta_j - \theta_k) = 1$ pour $j \neq k$ dans I , avec $-2\pi < \theta_j - \theta_k < 2\pi$, ce qui équivaut à $\theta_j = \theta_k$. En notant θ cette valeur commune on peut prendre $\theta_k = \theta$ pour les indices k tels que $\rho_k = 0$ et on a $z_k = \rho_k e^{i\theta} = |z_k| e^{i\theta}$ pour tout entier k compris entre 1 et n . $\square$

Avec les lemmes qui suivent on résume quelques propriétés élémentaires des matrices positives.

Lemme 4.2 Pour A, B dans $\mathcal{M}_{n,m}(\mathbb{K})$, on a :

1. $|A + B| \leq |A| + |B|$;
2. $|A|$ est positive et $|A| = 0$ si, et seulement si, $A = 0$;
3. pour tout scalaire λ , on a $|\lambda A| = |\lambda| |A|$;
4. pour A réelle strictement positive et $x \in \mathbb{R}^m$ positif non nul, Ax est strictement positif ;
5. s'il existe un vecteur x strictement positif dans $\mathbb{R}^m$ tel que $Ax = |A|x$, on a alors $A = |A|$.

Preuve. Le premier point se déduit des inégalités $|a_{i,j} + b_{i,j}| \leq |a_{i,j}| + |b_{i,j}|$ pour $1 \leq i \leq n$, $1 \leq j \leq m$ et les deux points suivants sont évidents.

Si $x \in \mathbb{R}^m$ est positif non nul, il existe alors un entier k compris entre 1 et m tel que $x_k > 0$ et pour tout i compris entre 1 et n , on a :

$$(Ax)_i = \sum_{j=1}^m a_{i,j} x_j \geq a_{i,k} x_k > 0$$

pour A strictement positive. On a donc $Ax > 0$ pour $A > 0$, $x \geq 0$ et $x \neq 0$.

Supposons qu'il existe un vecteur $x \in \mathbb{R}^m$ strictement positif tel que $Ax = |A|x$. Pour tout i compris entre 1 et n , on a $\sum_{j=1}^m (|a_{i,j}| - a_{i,j}) x_j = 0$ avec $x_j \in \mathbb{R}^{+,*}$ pour tout j , ce qui équivaut à $|a_{i,j}| = \operatorname{Re}(a_{i,j})$ et $\operatorname{Im}(a_{i,j}) = 0$, soit à $A = |A|$. $\square$

Lemme 4.3 Pour A dans $\mathcal{M}_{n,m}(\mathbb{K})$ et B dans $\mathcal{M}_{m,r}(\mathbb{K})$, on a :

1. $|AB| \leq |A||B|$;
2. pour A réelle positive et B réelle strictement positive, l'égalité $AB = 0$ équivaut à $A = 0$;
3. pour A réelle strictement positive et B telle que $|AB| = A|B|$, il existe des réels $\theta_1, \dots, \theta_r$ dans $]-\pi, \pi]$ tels que $B = |B|\Delta$, où $\Delta = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_r})$.

Preuve. Les inégalités $\left| \sum_{k=1}^m a_{i,k} b_{k,j} \right| \leq \sum_{k=1}^m |a_{i,k}| |b_{k,j}|$ pour $1 \leq i \leq n$, $1 \leq j \leq r$,

où $\left| \sum_{k=1}^m a_{i,k} b_{k,j} \right|$ est le coefficient d'indice (i, j) de $|AB|$ et $\sum_{k=1}^m |a_{i,k}| |b_{k,j}|$ celui de $|A||B|$, signifient que $|AB| \leq |A||B|$.

L'égalité $AB = 0$ est équivalente à $\sum_{k=1}^m a_{i,k} b_{k,j} = 0$ pour $1 \leq i \leq n$, $1 \leq j \leq r$ et pour A, B réelles positives cela équivaut à $a_{i,k} b_{k,j} = 0$ pour tout k compris entre 1 et m , équivalent à $a_{i,k} = 0$ si de plus B est strictement positive, c'est-à-dire que $A = 0$.

L'égalité $|AB| = A|B|$ avec A réelle strictement positive est équivalente à $\left| \sum_{k=1}^m a_{pk} b_{kq} \right| = \sum_{k=1}^m a_{pk} |b_{kq}|$ pour tout p compris entre 1 et n et tout q compris entre 1 et r . Pour p, q fixés, la suite de nombres complexes $(z_k)_{1 \leq k \leq m} = (a_{pk} b_{kq})_{1 \leq k \leq m}$ est telle que $\left| \sum_{k=1}^m z_k \right| = \sum_{k=1}^m |z_k|$, ce qui équivaut à l'existence d'un réel $\theta_{pq} \in]-\pi, \pi]$ tel que $z_k = e^{i\theta_{pq}} |z_k|$ pour tout k compris entre 1 et m (lemme 4.1). On a donc $a_{pk} b_{kq} = e^{i\theta_{pq}} |a_{pk} b_{kq}| = a_{pk} e^{i\theta_{pq}} |b_{kq}|$ avec $a_{pk} > 0$, ce qui équivaut à $b_{kq} = e^{i\theta_{pq}} |b_{kq}|$. En fixant p et en notant θ_q pour θ_{pq} , cela s'écrit $B = |B|\Delta$, où $\Delta = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_r})$. $\square$

Lemme 4.4 Si A, B dans $\mathcal{M}_{n,m}(\mathbb{R})$ et A', B' dans $\mathcal{M}_{m,r}(\mathbb{R})$ sont telles que $0 \leq A \leq B$ et $0 \leq A' \leq B'$, on a alors $0 \leq AA' \leq BB'$.

Preuve. Les matrices A, A' étant réelles positives, il en est de même de AA' . En écrivant que $BB' - AA' = B(B' - A') + (B - A)A'$, on déduit que si $A \leq B$, $A' \leq B'$, avec A', B positives, on a alors $AA' \leq BB'$. $\square$

Des lemmes précédents, on déduit facilement par récurrence que si A, B dans $\mathcal{M}_n(\mathbb{K})$ sont telles que $|A| \leq B$, on a alors $|A^k| \leq |A|^k \leq B^k$ pour tout entier naturel k .

Lemme 4.5 Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive.

1. Pour $B \in \mathcal{M}_n(\mathbb{C})$ telle que $|B| \leq A$, on a $\rho(B) \leq \rho(|B|) \leq \rho(A)$.
2. Pour toute sous-matrice principale $A_{p,q}$ de A , on a $\rho(A_{p,q}) \leq \rho(A)$. En particulier, on a $\max_{1 \leq i \leq n} a_{ii} \leq \rho(A)$.
3. Pour $B \in \mathcal{M}_n(\mathbb{R})$ telle que $0 \leq B < A$, on a $\rho(B) < \rho(A)$.
4. S'il existe $k \in \mathbb{N}^*$ tel que A^k soit strictement positive, on a alors $\rho(A) > 0$.

5. Si la somme des termes de chaque ligne [resp. colonne] de A est constante égale à α , on a alors $\rho(A) = \alpha = \|A\|_\infty$ [resp. $\rho(A) = \alpha = \|A\|_1$] et $\rho(A)$ est une valeur propre de A associée à un vecteur propre strictement positif.

6.

$$\min_{1 \leq i \leq n} \left(\sum_{j=1}^n a_{ij} \right) \leq \rho(A) \leq \max_{1 \leq i \leq n} \left(\sum_{j=1}^n a_{ij} \right)$$

$$\min_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij} \right) \leq \rho(A) \leq \max_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij} \right)$$

7. Pour tout $x \in \mathbb{R}^n$ strictement positif, on a $\min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$.
8. S'il existe $x \in \mathbb{R}^n$ strictement positif et α, β dans $\mathbb{R}^+$ tels que $\alpha x \leq Ax \leq \beta x$ [resp. $\alpha x < Ax < \beta x$], on a alors $\alpha \leq \rho(A) \leq \beta$ [resp. $\alpha < \rho(A) < \beta$].
9. Si la matrice A admet un vecteur propre strictement positif, la valeur propre associée est alors $\rho(A)$ et on a $\rho(A) = \sup_{x>0} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \inf_{x>0} \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$.

Preuve.

1. Pour tout $k \in \mathbb{N}$, on a $|B^k| \leq |B|^k \leq A^k$. Munissant $\mathcal{M}_n(\mathbb{C})$ de la norme $N : X = ((x_{ij}))_{1 \leq i, j \leq n} \mapsto \max_{1 \leq i, j \leq n} |x_{ij}|$, on a pour tout $k \in \mathbb{N}^*$:

$$N(B^k) = N(|B|^k) \leq N(|B|)^k \leq N(A)^k$$

et avec la croissance de $t \mapsto t^{\frac{1}{k}}$ sur $\mathbb{R}^+$, on en déduit que :

$$N(B^k)^{\frac{1}{k}} \leq N(|B|^k)^{\frac{1}{k}} \leq N(A)^{\frac{1}{k}}$$

puis faisant tendre k vers l'infini, il en résulte que $\rho(B) \leq \rho(|B|) \leq \rho(A)$ (théorème de Gelfand).

2. Pour $A \geq 0$, on a $0 \leq \tilde{A}_{p,q} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & A_{p,q} & 0 \\ 0 & 0 & 0 \end{pmatrix} \leq A$, ce qui nous donne

$$\rho(A_{p,q}) = \rho(\tilde{A}_{p,q}) \leq \rho(A).$$

3. Pour $\rho(B) = 0$, le résultat se déduit de $\rho(A) \geq \max_{1 \leq i \leq n} a_{ii} > 0$ (A est strictement positive). Pour tous i, j compris entre 1 et n , on a $0 \leq \frac{b_{ij}}{a_{ij}} < 1$, donc il existe un réel $\lambda > 0$ tel que $\frac{b_{ij}}{a_{ij}} < \lambda < 1$, soit $\frac{1}{\lambda} b_{ij} < a_{ij}$. On a donc $0 \leq \frac{1}{\lambda} B < A$ et $\rho\left(\frac{1}{\lambda} B\right) \leq \rho(A)$, ce qui entraîne pour $\rho(B) > 0$:

$$\rho(B) < \frac{1}{\lambda} \rho(B) = \rho\left(\frac{1}{\lambda} B\right) \leq \rho(A)$$

et en conséquence, $\rho(B) < \rho(A)$.

4. On a $\rho(A^k) = (\rho(A))^k$ et $A^k > 0$ donne $\rho(A^k) > 0$ (point précédent), donc $\rho(A) > 0$. On peut aussi dire que si $\rho(A) = 0$, on a alors $\rho(A^k) = (\rho(A))^k = 0$, donc toutes les valeurs propres de A^k sont nulles et $\text{Tr}(A^k) = 0$, ce qui est incompatible avec $A^k > 0$.

5. Dans le cas des lignes, les égalités $\sum_{j=1}^n a_{ij} = \alpha$ pour tout i compris entre 1 et n

nous disent que le vecteur $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$ est vecteur propre de la matrice

A associé à la valeur propre α . La matrice A et le réel α étant positifs, on a $\alpha \leq \rho(A) \leq \|A\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n a_{ij} \right) = \alpha$, soit $\alpha = \rho(A)$. En raisonnant

avec la transposée de la matrice A et en utilisant le fait qu'une matrice et sa transposée ont mêmes valeurs propres, on obtient le deuxième résultat en

considérant que $\|{}^t A\|_\infty = \|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n a_{ij} \right)$.

6. On sait déjà que $\rho(A) \leq \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$. On note $\alpha_i = \sum_{j=1}^n a_{ij}$ pour $1 \leq i \leq n$ et $\alpha = \min_{1 \leq i \leq n} (\alpha_i)$. Pour $\alpha = 0$, le résultat est évident. Pour $\alpha > 0$,

on a $\alpha_i > 0$ pour tout i et la matrice $B = \left(\left(\frac{\alpha}{\alpha_i} a_{ij} \right) \right)_{1 \leq i, j \leq n}$ est telle que

$0 \leq B \leq A$, $\sum_{j=1}^n b_{ij} = \alpha$ pour tout i compris entre 1 et n , ce qui nous donne

$\alpha = \rho(B) \leq \rho(A)$. En raisonnant avec ${}^t A$, considérant que $\rho({}^t A) = \rho(A)$ et $\|{}^t A\|_\infty = \|A\|_1$, on obtient le deuxième encadrement.

7. Pour $x = (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n$ strictement positif donné, la multiplication à droite de A par $D_x = \text{diag}(x_1, \dots, x_n)$ a pour effet de multiplier chaque colonne j de la matrice A par x_j et la multiplication à gauche par D_x^{-1} de diviser chaque ligne i de la matrice A par x_i , de sorte que $D_x^{-1} A D_x = \left(\left(\frac{x_j}{x_i} a_{ij} \right) \right)_{1 \leq i, j \leq n}$, ce qui nous donne :

$$\begin{aligned} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} &= \min_{1 \leq i \leq n} \sum_{j=1}^n \frac{x_j}{x_i} a_{ij} \leq \rho(D_x^{-1} A D_x) = \rho(A) \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n \frac{x_j}{x_i} a_{ij} = \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \end{aligned}$$

8. L'encadrement $\alpha x \leq Ax \leq \beta x$ équivaut à $\alpha x_i \leq (Ax)_i \leq \beta x_i$ pour tout i compris entre 1 et n , ce qui entraîne :

$$\alpha \leq \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \beta$$

On procède de même pour les inégalités strictes.

9. On a $\sup_{x>0} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \inf_{x>0} \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$. Si A admet un vecteur propre strictement positif x , la valeur propre α associée est alors réelle positive et avec $\alpha = \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \alpha$, on en déduit les égalités $\alpha = \rho(A) = \sup_{x>0} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \inf_{x>0} \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$.

□

Les inégalités $a_{ii} \leq \rho(A)$ ne sont pas nécessairement vérifiées si la matrice A n'est pas positive. Par exemple toute matrice $A = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}$ semblable à $J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a un rayon spectral nul et une telle matrice s'écrit, en notant $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible :

$$A = P^{-1}JP = \frac{1}{ad-bc} \begin{pmatrix} dc & d^2 \\ -c^2 & -dc \end{pmatrix}$$

En prenant $dc > 0$, on a $a_{11} > \rho(A) = 0$.

Du point 6. du lemme précédent, on déduit que si A est une matrice positive dans $\mathcal{M}_n(\mathbb{R})$ telle $\sum_{j=1}^n a_{ij} > 0$ [resp. $\sum_{i=1}^n a_{ij} > 0$] pour tout i [resp. j] compris entre 1 et n , on a alors $\rho(A) > 0$. En particulier une matrice strictement positive a son rayon spectral strictement positif.

4.2 Matrices strictement positives et théorème de Perron-Frobenius

Lemme 4.6 Soient $A \in \mathcal{M}_n(\mathbb{R})$ strictement positive et $x \in \mathbb{C}^n$ un vecteur propre non nul associé à une valeur propre λ telle que $|\lambda| = \rho(A)$. Dans ce cas, $\rho(A)$ est valeur propre de A avec $|x|$ comme vecteur propre associé, ce vecteur étant strictement positif et il existe un réel θ tel que $x = e^{i\theta} |x|$.

Preuve. On a $\rho(A) > 0$ du fait que $A > 0$ (point 6. du lemme 4.5).

De $Ax = \lambda x$ avec $|\lambda| = \rho(A)$, on déduit que $\rho(A)|x| = |Ax| \leq |A||x| = A|x|$, donc $y = A|x| - \rho(A)|x|$ est positif. Si ce vecteur est non nul, on a alors $Ay > 0$ car A est strictement positive (point 4. du lemme 4.2), ce qui signifie en notant $x' = A|x|$ que $\rho(A)x' < Ax'$ avec $x' > 0$ (le vecteur x est non nul) qui entraîne que $\rho(A) < \rho(A)$ (point 8. du lemme 4.5), soit une impossibilité. On a donc $y = 0$, soit $A|x| = \rho(A)|x|$, ce qui signifie que $\rho(A)$ est valeur propre de A avec $|x|$ comme vecteur propre associé. En écrivant que $|x| = \frac{1}{\rho(A)} A|x|$, on déduit que $|x| > 0$.

On a donc $A|x| = \rho(A)|x| = |Ax|$ et du lemme 4.3, point 3, avec $B = x$ dans $\mathcal{M}_{n,1}(\mathbb{R})$, on déduit qu'il existe $\theta \in]-\pi, \pi]$ tel que $x = e^{i\theta} |x|$. □

Théorème 4.1. Perron-Frobenius

Si $A \in \mathcal{M}_n(\mathbb{R})$ est strictement positive, $\rho(A)$ est alors l'unique valeur propre de A de module maximum, l'espace propre associé à $\rho(A)$ est une droite vectorielle engendrée par un vecteur strictement positif et $\rho(A)$ est valeur propre simple de A .

Preuve. Dans la démonstration du lemme précédent on a vu que si λ est une valeur propre de la matrice A telle que $|\lambda| = \rho(A)$ et si x est un vecteur propre non nul associé, on a alors $x = e^{i\theta} |x|$ avec $A|x| = \rho(A)|x|$. Le rayon spectral $\rho(A)$ est donc valeur propre de A . De plus, avec :

$$\lambda x = Ax = A(e^{i\theta} |x|) = e^{i\theta} A|x| = e^{i\theta} \rho(A) |x| = \rho(A) x$$

on déduit que $\lambda x = \rho(A)x$ avec $x \neq 0$, et $\lambda = \rho(A)$. Donc $\rho(A)$ est l'unique valeur propre de A de module maximal.

En notant $E_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho(A)$, tout vecteur non nul x dans $E_{\rho(A)}$ est tel que $|x| > 0$ (lemme précédent) et aucune des composantes de x n'est nulle. S'il existe deux vecteurs x, y linéairement indépendants dans $E_{\rho(A)}$, le vecteur $z = x_1 y - y_1 x$ est alors non nul dans $E_{\rho(A)}$ avec $z_1 = 0$, ce qui est impossible. On a donc $\dim(E_{\rho(A)}) = 1$.

Pour $n = 1$, il est clair que $\rho(A)$ est valeur propre simple de A . On suppose donc que $n \geq 2$. Si la multiplicité de $\rho(A)$ comme valeur propre de A est $m \geq 2$, en se donnant un générateur $x > 0$ de l'espace propre $E_{\rho(A)}$, il existe $y \in \mathbb{C}^n$ linéairement indépendant de x tel que $Ay = x + \rho(A)y$ (la matrice A est semblable

à une matrice de la forme $\begin{pmatrix} \begin{pmatrix} \rho(A) & 1 \\ 0 & \rho(A) \end{pmatrix} & 0 \\ 0 & B \end{pmatrix}$ puisque $\dim(E_{\rho(A)}) = 1$

avec $\rho(A)$ de multiplicité $m \geq 2$). En notant $\bar{y}$ le vecteur conjugué de y dans $\mathbb{C}^n$, on a $A\bar{y} = x + \rho(A)\bar{y}$ puisque A et x sont réels. Le vecteur $z = \frac{1}{2}(y + \bar{y})$ est alors réel tel que $Az = x + \rho(A)z$. Comme x est strictement positif, il existe un réel assez grand $\alpha > 0$ tel que $v = z + \alpha x$ soit strictement positif. On dispose alors d'un vecteur $v > 0$ tel que $Av = x + \rho(A)v > \rho(A)v$, ce qui nous donne $\rho(A) < \rho(A)$ (point 8. du lemme 4.5), soit une impossibilité. $\square$

Corollaire 4.1 : *Si $A \in \mathcal{M}_n(\mathbb{R})$ est strictement positive, il existe alors un unique vecteur propre associé à la valeur propre $\rho(A)$ dans le compact $K_1 = \{x \in \mathbb{R}^n \mid x \geq 0, \|x\|_1 = 1\}$.*

Preuve. On a vu que $E_{\rho(A)} = \mathbb{R}x$ avec $x > 0$. Le vecteur $v = \frac{1}{\|x\|_1}x$ est alors l'unique élément de $K_1 \cap E_{\rho(A)}$. $\square$

Le vecteur $x \in K_1 \cap E_{\rho(A)}$ est appelé, le vecteur de Perron de la matrice strictement positive A .

Corollaire 4.2 : *Si $A \in \mathcal{M}_n(\mathbb{R})$ est strictement positive, elle ne peut alors avoir deux vecteurs propres positifs linéairement indépendants.*

Preuve. Soit y un vecteur propre positif non nul associé à une valeur propre λ de A . On a $Ay > 0$ et il existe un indice i compris entre 1 et n tel que $(Ay)_i = \lambda y_i$ avec $y_i > 0$, ce qui entraîne que $\lambda > 0$ et $y = \frac{1}{\lambda} Ay > 0$. Avec le point 9. du lemme 4.5, on déduit que nécessairement $\lambda = \rho(A)$ et le théorème de Perron-Frobenius nous permet de conclure. $\square$

En utilisant la densité de l'ensemble des matrices strictement positives dans l'ensemble des matrices positives, on déduit le résultat suivant.

Corollaire 4.3 : *Si $A \in \mathcal{M}_n(\mathbb{R})$ est positive, $\rho(A)$ est alors valeur propre de A et il existe un vecteur propre associé non nul positif.*

Preuve. Pour tout $k \in \mathbb{N}^*$, on note $A_k = \left(\left(a_{ij} + \frac{1}{k} \right) \right)_{1 \leq i, j \leq n}$ et v_k est le vecteur de Perron de la matrice strictement positive A_k . On a $\lim_{k \rightarrow +\infty} A_k = A$ et avec la continuité du rayon spectral, on déduit que $\lim_{k \rightarrow +\infty} \rho(A_k) = \rho(A)$. D'autre part, la suite $(v_k)_{k \geq 1}$ étant dans le compact K_1 , on peut en extraire une sous suite $(v_{\varphi(k)})_{k \geq 1}$ convergente vers un vecteur $v \geq 0$ et on a :

$$Av = \lim_{k \rightarrow +\infty} A_{\varphi(k)} v_{\varphi(k)} = \lim_{k \rightarrow +\infty} \rho(A_{\varphi(k)}) v_{\varphi(k)} = \rho(A) v$$

c'est-à-dire que v est un vecteur propre non nul (puisque $\|v\|_1 = 1$) positif de A associé à la valeur propre $\rho(A)$. $\square$

Corollaire 4.4 : *Si $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice positive, on a alors $\rho(I_n + A) = 1 + \rho(A)$. De plus, s'il existe un entier $k \geq 1$ tel que A^k soit strictement positive, $\rho(A)$ est alors valeur propre simple de A .*

Preuve. Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a $\text{Sp}(I_n + A) = \{1 + \lambda \mid \lambda \in \text{Sp}(A)\}$, donc $\rho(I_n + A) \leq 1 + \rho(A)$. Si de plus A est positive, $\rho(A)$ est alors valeur propre de A , donc $1 + \rho(A)$ est valeur propre de $I_n + A$ et $1 + \rho(A) \leq \rho(I_n + A)$, d'où l'égalité.

En notant m la multiplicité algébrique de la valeur propre $\rho(A)$ de A , il existe une matrice inversible P telle que $T = P^{-1}AP$ soit triangulaire supérieure de diagonale $(\rho(A), \dots, \rho(A), \lambda_{m+1}, \dots, \lambda_n)$ avec $|\lambda_i| < \rho(A)$ pour tout entier i compris entre $m+1$ et n (dans le cas où $m < n$). La matrice $T^k = P^{-1}A^kP$ est alors triangulaire supérieure de diagonale $(\rho(A)^k, \dots, \rho(A)^k, \lambda_{m+1}^k, \dots, \lambda_n^k)$ et $\rho(A)^k = \rho(A^k)$ est alors valeur propre de A^k de multiplicité supérieure ou égale à m . Pour A^k strictement positive, cette multiplicité vaut 1, donc $m = 1$. $\square$

Une matrice positive telle que A^k soit strictement positive pour un entier $k \geq 1$ est un cas particulier de matrice positive irréductible.

Les théorèmes de Perron-Frobenius et de Gerschgorin-Hadamard nous permettent d'obtenir le résultat suivant de localisation des valeurs propres d'une matrice, où $D(a, r) = \{z \in \mathbb{C} \mid |z - a| \leq r\}$ désigne le disque fermé de centre $a \in \mathbb{C}$ et de rayon $r > 0$.

Corollaire 4.5 : Soient $A \in \mathcal{M}_n(\mathbb{C})$ et $B \in \mathcal{M}_n(\mathbb{R})$ positive telle que $|A| \leq B$. On a $\text{Sp}(A) \subset \bigcup_{i=1}^n D(a_{ii}, \rho(B) - b_{ii})$, ce qui signifie que toutes les valeurs propres de A sont dans la réunion des disques de centre a_{ii} et de rayon $\rho(B) - b_{ii}$.

Preuve. Supposons dans un premier temps que la matrice B soit strictement positive.

Si x est le vecteur de Perron de B , on a alors pour tout entier i compris entre 1 et n :

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \sum_{\substack{j=1 \\ j \neq i}}^n b_{ij} x_j = (\rho(B) - b_{ii}) x_i$$

soit $\frac{1}{x_i} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \rho(B) - b_{ii}$. D'autre part, le théorème de Gerschgorin-Hadamard

nous dit que si $\lambda \in \mathbb{C}$ est une valeur propre d'une matrice C , il existe alors un indice i compris entre 1 et n tel que $|\lambda - c_{ii}| \leq \sum_{j=1, j \neq i}^n |c_{ij}|$. La matrice A ayant les mêmes valeurs propres que la matrice $C_x = D_x^{-1} A D_x$, où $D_x = \text{diag}(x_1, \dots, x_n)$ (on a donc $C_x = \left(\left(\frac{x_j}{x_i} a_{ij} \right) \right)_{1 \leq i, j \leq n}$), on déduit que pour toute valeur propre $\lambda \in \mathbb{C}$ de A , il existe un indice i tel que :

$$|\lambda - a_{ii}| = |\lambda - c_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |c_{ij}| = \frac{1}{x_i} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| x_j \leq \rho(B) - b_{ii}$$

ce qui donne le résultat annoncé.

Dans le cas où la matrice B est positive, non strictement positive, en désignant pour tout réel $\varepsilon > 0$, par B_ε la matrice strictement positive $B_\varepsilon = ((b_{ij} + \varepsilon))_{1 \leq i, j \leq n}$,

on a $B_\varepsilon > |A|$, donc $\text{Sp}(A) \subset \bigcup_{i=1}^n D(a_{ii}, \rho(B_\varepsilon) - b_{ii} - \varepsilon)$. Avec la continuité du rayon spectral, on en déduit alors le résultat annoncé. $\square$

Lemme 4.7 Soit $A \in \mathcal{M}_n(\mathbb{R})$ strictement positive. Si x est un vecteur propre strictement positif de A associé à la valeur propre $\rho(A)$, il existe alors un unique vecteur y strictement positif qui est vecteur propre de ${}^t A$ associé à la valeur propre $\rho({}^t A) = \rho(A)$ tel que ${}^t y x = 1$.

Preuve. La matrice ${}^t A$ est strictement positive avec $\rho({}^t A) = \rho(A)$, l'espace propre associé à la valeur propre $\rho(A)$ de ${}^t A$ est donc une droite vectorielle dirigée par un vecteur $z > 0$. En posant $y = \frac{1}{{}^t z x} z$, on a $y > 0$, ${}^t A y = \rho(A) y$ et ${}^t y x = 1$. Réciproquement, si ${}^t A y = \rho(A) y$ avec $y > 0$, ${}^t y x = 1$, on a alors $y = \alpha z$ avec $\alpha = \frac{1}{{}^t z x}$, ce qui prouve l'unicité du vecteur y . $\square$

Avec les notations et hypothèses du lemme précédent, on désigne par $E_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho(A)$ de A (matrice strictement positive), par $E'_{\rho(A)}$ l'espace propre associé à la valeur propre $\rho({}^tA) = \rho(A)$ de ${}^tA$ et on note $L = x {}^ty = ((x_i y_j))_{1 \leq i, j \leq n}$.

Avec le lemme qui suit, on résume quelques propriétés de cette matrice L .

Lemme 4.8 *La matrice L est indépendante du choix du vecteur x et strictement positive de rang 1. On a $Lx = x$, ${}^tLy = y$ et pour tout entier naturel k non nul :*

$$L^k = L, \quad A^k L = L A^k = \rho(A)^k L, \quad \left(\frac{1}{\rho(A)} A - L \right)^k = \left(\frac{1}{\rho(A)} A \right)^k - L$$

L est la matrice de la projection sur la droite vectorielle $E_{\rho(A)}$ parallèlement à l'hyperplan $H = \{z \in \mathbb{C}^n \mid {}^t y z = 0\}$.

Preuve. On a $E_{\rho(A)} = \mathbb{C}x$, $E'_{\rho(A)} = \mathbb{C}y$, avec $x > 0$, $y > 0$ et ${}^t y x = 1$. Si (x', y') est un autre couple de vecteurs vérifiant ces propriétés, on a alors $x' = \alpha x$, $y' = \beta y$, avec $\alpha > 0$, $\beta > 0$, $1 = {}^t y' x' = \alpha \beta {}^t y x = \alpha \beta$ et $x' {}^t y' = \alpha \beta x {}^t y = x {}^t y = L$.

Les vecteurs x et y étant strictement positifs, il en est de même de la matrice L . De plus, pour j compris entre 1 et n , la colonne numéro j de L est $y_j x > 0$, il en résulte que L est de rang 1 avec $\text{Im}(L) = \mathbb{C}x = E_{\rho(A)}$.

Avec l'associativité du produit matriciel, on peut écrire :

$$\begin{aligned} Lx &= x ({}^t y x) = ({}^t y x) x = x \\ {}^t Ly &= y ({}^t x y) = ({}^t x y) y = {}^t ({}^t y x) y = y \end{aligned}$$

On a $L^2 = x ({}^t y x) {}^t y = ({}^t y x) (x {}^t y) = L$ et par récurrence $L^k = L$ pour tout $k \geq 1$.

De $Ax = \rho(A)x$, on déduit que $A^k x = \rho(A)^k x$ pour tout $k \geq 1$ et :

$$A^k L = A^k x {}^t y = \rho(A)^k x {}^t y = \rho(A)^k L$$

De même avec ${}^t A y = \rho(A)y$, on obtient :

$$L A^k = x {}^t y A^k = x {}^t \left(({}^t A)^k y \right) = x {}^t \left(\rho(A)^k y \right) = \rho(A)^k L$$

La dernière relation peut se montrer par récurrence sur $k \geq 1$. Le résultat est évident pour $k = 1$ et en le supposant acquis pour $k \geq 1$, on a :

$$\begin{aligned} \left(\frac{1}{\rho(A)} A - L \right)^{k+1} &= \left(\frac{1}{\rho(A)} A - L \right) \left(\left(\frac{1}{\rho(A)} A \right)^k - L \right) \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - \frac{1}{\rho(A)^k} L A^k - \frac{1}{\rho(A)} A L + L^2 \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - \frac{1}{\rho(A)^k} \rho(A)^k L - \frac{1}{\rho(A)} \rho(A) L + L \\ &= \left(\frac{1}{\rho(A)} A \right)^{k+1} - L \end{aligned}$$

On a $L^2 = L$, $\text{Im}(L) = E_{\rho(A)}$ et le noyau de L est donné par :

$$\text{Ker}(L) = \{z \in \mathbb{C}^n \mid x {}^t y z = 0\} = \{z \in \mathbb{C}^n \mid ({}^t y z) x = 0\} = H$$

c'est-à-dire que L est la projection sur $E_{\rho(A)}$ parallèlement à H . $\square$

Théorème 4.2.

Avec les notations et hypothèses du lemme 4.7, on a :

$$\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)} A \right)^k = L = x {}^t y$$

Preuve. Avec $\left(\frac{1}{\rho(A)} A \right)^k - L = \left(\frac{1}{\rho(A)} A - L \right)^k$ pour $k \geq 1$, il suffit de montrer que la matrice $\frac{1}{\rho(A)} A - L$ a un rayon spectral strictement inférieur à 1, ce qui revient à dire que $\rho(A - \rho(A)L) < \rho(A)$. Si λ est une valeur propre non nulle de $B = A - \rho(A)L$ et z est un vecteur propre non nul associé, on a alors $(A - \rho(A)L)z = \lambda z$ et avec $L(A - \rho(A)L) = 0$, on déduit que $\lambda Lz = 0$ et $Lz = 0$ pour $\lambda \neq 0$, ce qui entraîne que $\lambda z = Az - \rho(A)Lz = Az$, c'est-à-dire que λ est valeur propre de A avec z pour vecteur propre associé. On a donc montré que toute valeur propre λ non nulle de B est aussi valeur propre de A et donc $|\lambda| \leq \rho(A)$. Si λ est une valeur propre non nulle de B telle que $|\lambda| = \rho(A)$, on a alors $\lambda = \rho(A)$ (premier point du théorème de Perron-Frobenius 4.1) et tout vecteur propre associé z est aussi vecteur propre de A , on a donc $z = \alpha x$ avec $\alpha \in \mathbb{C}^*$ pour $z \neq 0$, ce qui entraîne :

$$\begin{aligned} \rho(A)z &= (A - \rho(A)L)z = (A - \rho(A)L)\alpha x \\ &= \alpha Ax - \alpha \rho(A)Lx = \alpha \rho(A)x - \alpha \rho(A)x = 0 \end{aligned}$$

en contradiction avec $\rho(A) > 0$ et $z \neq 0$.

On a donc $|\lambda| < \rho(A)$ pour toute valeur propre de B et $\rho(B) < \rho(A)$. D'où le résultat. $\square$

Avec ce résultat, on retrouve le fait que la matrice L ne dépend que de la matrice strictement positive A et pas du choix de x . On retrouve aussi le dernier point du théorème de Perron-Frobenius 4.1.

Théorème 4.3. Perron-Frobenius

Si $A \in \mathcal{M}_n(\mathbb{R})$ est strictement positive, $\rho(A)$ est alors valeur propre simple de A (l'espace propre associé est donc une droite vectorielle).

Preuve. On sait déjà que $\rho(A)$ est l'unique valeur propre de A de module maximum. Notons m sa multiplicité algébrique. Le théorème de trigonalisation sur $\mathbb{C}$ nous dit qu'il existe une matrice inversible P telle que $T = P^{-1}AP$ soit triangulaire supérieure de diagonale $(\rho(A), \dots, \rho(A), \lambda_{p+1}, \dots, \lambda_n)$ avec $|\lambda_i| < \rho(A)$ pour i compris entre $m+1$ et n (si $m < n$). En écrivant, pour tout entier naturel non nul k , que $\left(\frac{1}{\rho(A)} T \right)^k = P^{-1} \left(\frac{1}{\rho(A)} A \right)^k P$ et avec la continuité du produit matriciel,

le théorème précédent nous dit que $\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)} T \right)^k = P^{-1} L P = L'$ où L' est triangulaire supérieure de diagonale $(1, \dots, 1, 0, \dots, 0)$. On a donc $\text{rang}(L') \geq m$ et avec $\text{rang}(L') = \text{rang}(L) = 1$, on déduit que nécessairement $m = 1$. $\square$

Si la matrice A est positive, on a vu que $\rho(A)$ est valeur propre de A (corollaire 4.3), mais cette valeur propre n'est pas nécessairement simple (prendre par exemple la matrice identité). Dans le cas où il existe un entier naturel k tel que A^k soit strictement positive, $\rho(A)$ est valeur propre simple de A .

4.3 Matrices irréductibles

Si n est un entier naturel supérieur ou égal à 2, on désigne alors par $\mathfrak{S}_n$ le groupe des permutations de l'ensemble $\{1, \dots, n\}$ et par $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

Pour tout couple (i, j) d'entiers naturels, on note $\delta_{i,j}$ le symbole de Kronecker ($\delta_{ii} = 1$ et $\delta_{i,j} = 0$ pour $i \neq j$).

Définition 4.2. La matrice de permutation associée à $\sigma \in \mathfrak{S}_n$ est la matrice de passage P_σ de la base canonique de $\mathbb{K}^n$ à la base $\mathcal{B}_\sigma = (e_{\sigma(k)})_{1 \leq k \leq n}$.

On a donc, si P_σ est une matrice de permutation, $P_\sigma e_j = e_{\sigma(j)}$ pour tout entier j compris entre 1 et n , ce qui revient à dire que $P_\sigma = ((\delta_{i,\sigma(j)}))_{1 \leq i,j \leq n}$. En particulier, on a $P_{Id} = I_n$.

Une telle matrice de permutation étant la matrice de passage de la base canonique de $\mathbb{K}^n$ à une base orthonormée est unitaire, donc $P_\sigma^{-1} = {}^t P_\sigma$.

Une matrice de permutation étant unitaire dans $\mathcal{M}_n(\mathbb{C})$ est diagonalisable sur $\mathbb{C}$. Mais sur $\mathbb{R}$ ce résultat n'est plus valable. Par exemple, pour σ cycle d'ordre 3, le polynôme minimal de P_σ est $X^3 - 1$ et ce dernier n'est pas scindé sur $\mathbb{R}$, donc cette matrice n'est pas diagonalisable sur $\mathbb{R}$.

De plus il est facile de vérifier que pour toutes permutations σ, τ dans $\mathfrak{S}_n$, on a $P_\sigma P_\tau = P_{\sigma \circ \tau}$ (il suffit d'écrire que pour tout i compris entre 1 et n , on a $P_\sigma P_\tau e_i = P_\sigma e_{\tau(i)} = e_{\sigma \circ \tau(i)} = P_{\sigma \circ \tau} e_i$). On en déduit alors que $P_\sigma^{-1} = P_{\sigma^{-1}}$.

Pour toute permutation $\sigma \in \mathfrak{S}_n$ et tout vecteur $x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n$, on a $P_\sigma^{-1} x = \sum_{j=1}^n x_j e_{\sigma^{-1}(j)} = \sum_{i=1}^n x_{\sigma(i)} e_i$ (changement d'indice $i = \sigma^{-1}(j)$), ce qui signifie que $P_\sigma^{-1} x = (x_{\sigma(i)})_{1 \leq i \leq n}$ se déduit de x en faisant agir la permutation σ sur les composantes de x . On en déduit que pour toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$, $P_\sigma^{-1} A = ((a_{\sigma(i),j}))_{1 \leq i,j \leq n}$ se déduit de A en faisant agir la permutation σ sur les lignes de A .

La multiplication à droite d'une matrice A par une matrice de permutation P_σ va faire agir la permutation σ sur les colonnes de A . En effet, pour tout j compris entre 1 et n , on a $A P_\sigma e_j = A e_{\sigma(j)} = \sum_{i=1}^n a_{i,\sigma(j)} e_i$, donc $A P_\sigma = ((a_{i,\sigma(j)}))_{1 \leq i,j \leq n}$.

On a donc pour toute permutation $\sigma \in \mathfrak{S}_n$ et toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$, $P_\sigma^{-1}AP_\sigma = ((a_{\sigma(i),\sigma(j)}))_{1 \leq i,j \leq n}$, c'est-à-dire que $P_\sigma^{-1}AP_\sigma$ se déduit de A en faisant agir la permutation σ sur les lignes et les colonnes de A .

Définition 4.3. Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite réductible s'il existe une matrice de permutation P_σ telle que :

$$P_\sigma^{-1}AP_\sigma = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \quad (4.1)$$

où $B \in \mathcal{M}_p(\mathbb{K})$ avec $1 \leq p \leq n-1$. Une matrice non réductible est dite irréductible.

Exemples 4.1

1. Une matrice ayant tous ses coefficients non nuls est irréductible.
2. Une matrice ayant une ligne (ou une colonne) nulle est réductible. En effet si la ligne numéro i est nulle en transposant la ligne 1 avec la ligne i et la colonne 1 avec la colonne i , on obtient une matrice avec la première ligne nulle.
3. Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{K})$ est irréductible si, et seulement si, $bc \neq 0$. En effet $\mathfrak{S}_2 = \{\tau_{12}, I_d\}$ et $P_{I_d}^{-1}AP_{I_d} = A$ ou $P_{\tau_{12}}^{-1}AP_{\tau_{12}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est de la forme (4.1) si, et seulement si, $b = 0$ ou $c = 0$.

Une condition suffisante d'irréductibilité est donnée par le résultat suivant.

Lemme 4.9 Soit $A \in \mathcal{M}_n(\mathbb{K})$. S'il existe un entier naturel p tel que A^p ait tous ses coefficients non nuls, A est alors irréductible.

Preuve. Si A est réductible, il existe alors une matrice de permutation P_σ telle que $P_\sigma^{-1}AP_\sigma = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ et pour tout $p \geq 1$ on a $P_\sigma^{-1}A^pP_\sigma = \begin{pmatrix} B^p & 0 \\ C_p & D^p \end{pmatrix}$, ce qui signifie que la matrice A^p est également réductible, elle a donc au moins un coefficient nul. $\square$

Exemple 4.1 La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ est irréductible puisque $A^2 > 0$.

Cette condition n'est pas nécessaire comme le montre l'exemple de la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Une définition équivalente de la notion de matrice réductible est donnée par le résultat suivant.

Théorème 4.4.

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est réductible si, et seulement si, il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$.

Preuve. Soit $A \in \mathcal{M}_n(\mathbb{K})$ réductible. Il existe une permutation σ telle que $a_{\sigma(i), \sigma(j)} = 0$ pour $1 \leq i \leq p$ et $p+1 \leq j \leq n$, où p est un entier compris entre 1 et $n-1$. En notant $I = \{\sigma(1), \dots, \sigma(p)\}$, $J = \{\sigma(p+1), \dots, \sigma(n)\}$, on définit une partition non triviale de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$. Réciproquement supposons qu'il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$. En notant p le cardinal de I , on a $1 \leq p \leq n-1$ et il existe $\sigma \in \mathfrak{S}_n$ telle que $I = \{\sigma(1), \dots, \sigma(p)\}$, $J = \{\sigma(p+1), \dots, \sigma(n)\}$, ce qui donne :

$$P_\sigma^{-1} A P_\sigma = ((a_{\sigma(i), \sigma(j)}))_{1 \leq i, j \leq n} = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

c'est-à-dire que A est réductible. $\square$

De ce théorème, on peut déduire qu'une matrice est irréductible si, et seulement si, sa transposée l'est.

Une autre définition équivalente de la notion de matrice réductible est donnée par le résultat suivant.

Théorème 4.5.

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est réductible si, et seulement si, il existe une partie non triviale J de $\{1, \dots, n\}$ telle que le sous-espace vectoriel de $\mathbb{K}^n$, $V_J = \text{Vect}\{e_j \mid j \in J\}$, soit stable par A .

Preuve. Dire que $A \in \mathcal{M}_n(\mathbb{K})$ est réductible équivaut à dire qu'il existe une partition non triviale (I, J) de $\{1, \dots, n\}$ telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$, ce qui revient à dire que pour tout $j \in J$ on a $Ae_j = \sum_{i=1}^n a_{ij} e_i = \sum_{i \in J} a_{ij} e_i \in V_J$, ce qui est encore équivalent à dire que V_J est stable par A . $\square$

Les propriétés élémentaires suivantes nous seront utiles.

Lemme 4.10 *Si A est réductible, il en est alors de même de $|A|$ et de $I_n + A$.*

Preuve. Cela résulte de $P_\sigma^{-1} A P_\sigma = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ qui entraîne que :

$$P_\sigma^{-1} |A| P_\sigma = \begin{pmatrix} |B| & 0 \\ |C| & |D| \end{pmatrix} \text{ et } P_\sigma^{-1} (I_n + A) P_\sigma = \begin{pmatrix} I_p + B & 0 \\ C & I_{n-p} + D \end{pmatrix}$$

puisque l'action de P_σ est seulement de permuter des lignes et des colonnes. $\square$

On a vu que si A est une matrice strictement positive dans $\mathcal{M}_n(\mathbb{R})$, $\rho(A)$ est alors l'unique valeur propre de A de module maximum, cette valeur propre est simple, l'espace propre associé est une droite vectorielle engendrée par un vecteur

strictement positif et une telle matrice ne peut avoir deux vecteurs propres positifs linéairement indépendants.

Ces résultats ne s'étendent pas au cas des matrices positives, mais pour les matrices positives qui sont de plus irréductibles, on a des résultats analogues.

Lemme 4.11 *Soient $A \in \mathcal{M}_n(\mathbb{R})$ positive irréductible et $y \in \mathbb{R}^n$ un vecteur positif non nul. Si y est strictement positif, il en est alors de même de $z = (I_n + A)y$. Si y a au moins une composante nulle, le nombre de coordonnées nulles de z est alors strictement inférieur au nombre de coordonnées nulles de y . Dans tous les cas, le vecteur $(I_n + A)^{n-1}y$ est strictement positif.*

Preuve. Les composantes du vecteur $z = (I_n + A)y$ sont données par :

$$z_i = \sum_{j=1}^n a_{ij}y_j + y_i \quad (1 \leq i \leq n)$$

Si la matrice A et le vecteur y sont positifs, on a alors $z_i \geq y_i \geq 0$ et z est positif. Avec $z \geq y$, on déduit que z est strictement positif si y l'est.

En supposant que y a au moins une composante nulle, de $0 \leq y_i \leq z_i$, on déduit que $z_i = 0$ entraîne $y_i = 0$. Le nombre de coordonnées nulles du vecteur z est donc inférieur ou égal à celui de y .

Supposons que z et y ont le même nombre de composantes nulles. En notant J_y l'ensembles des indices compris entre 1 et n tels que $y_i = 0$, on a $z_i > 0$ pour $i \notin J_y$ et en conséquence $z_i = y_i = 0$ pour tout $i \in J_y$, avec $z_i = \sum_{j \notin J_y} a_{ij}y_j$ et $y_j > 0$ pour

$j \notin J_y$. On a donc en tenant compte du fait que les coefficients a_{ij} sont positifs ou nuls, $a_{ij} = 0$ pour $i \in J_y$ et $j \notin J_y$ avec J_y de cardinal compris entre 1 et $n - 1$ (y a au moins une composante nulle et n'est pas le vecteur nul) ce qui revient à dire que la matrice A est réductible. En conclusion le nombre de composantes nulles de z est strictement inférieur à celui de y , si la matrice positive A est irréductible.

Si le vecteur y est positif non nul, il a alors au moins une coordonnée strictement positive et ce qui précède nous dit que le vecteur $(I_n + A)y$ a au moins deux coordonnées strictement positives. Par récurrence on en déduit que le vecteur $(I_n + A)^{n-1}y$ a au moins n coordonnées strictement positives, ce qui revient à dire qu'il est strictement positif. $\square$

Théorème 4.6.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive. Cette matrice est irréductible si, et seulement si, la matrice $(I_n + A)^{n-1}$ est strictement positive.

Preuve. Du lemme 4.11, on déduit que si A est positive irréductible, alors pour tout j compris entre 1 et n le vecteur $(I_n + A)^{n-1}e_j$ est strictement positif, ce qui équivaut à dire que la matrice $(I_n + A)^{n-1}$ est strictement positive. Réciproquement si la matrice $(I_n + A)^{n-1}$ est strictement positive, elle est alors irréductible ainsi que A (lemme 4.10). $\square$

On peut en fait montrer que $A \in \mathcal{M}_n(\mathbb{K})$ est irréductible si, et seulement si, $(I_n + |A|)^{n-1}$ est strictement positive (voir [9], théorème 6.2.23).

Théorème 4.7. Perron-Frobenius

Si $A \in \mathcal{M}_n(\mathbb{R})$ est positive et irréductible, $\rho(A)$ est alors strictement positif, c'est une valeur propre simple de A et l'espace propre associé est une droite vectorielle engendrée par un vecteur strictement positif.

Preuve. La matrice A étant irréductible n'a pas de ligne nulle, on a donc puis-
qu'elle est positive $\sum_{j=1}^n a_{ij} > 0$ pour tout i compris entre 1 et n , ce qui entraîne
 $\rho(A) > 0$ (point 6. du lemme 4.5). Avec le théorème de trigonalisation sur $\mathbb{C}$, on
voit que si $\rho(A)$ est valeur propre de multiplicité supérieure ou égal à 2, il en est
alors de même de $1 + \rho(A)$ comme valeur propre de $I_n + A$. Mais $I_n + A$ positive
et $(I_n + A)^{n-1}$ strictement positive entraîne que $\rho(I_n + A) = 1 + \rho(A)$ est valeur
propre simple de $I_n + A$ (corollaire 4.4). En conséquence $\rho(A)$ est valeur propre
simple de A . L'espace propre associé est donc de dimension 1 et on sait qu'il peut
être engendré par un vecteur positif x (corollaire 4.3). De $Ax = \rho(A)x$, on dé-
duit que $(I_n + A)^{n-1}x = (1 + \rho(A))^{n-1}x$ et avec $(I_n + A)^{n-1}x > 0$ (point 4. du
lemme 4.2), $(1 + \rho(A))^{n-1} > 0$, on déduit que $x > 0$. $\square$

Corollaire 4.6 : *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ positive irréductible ne peut pas posséder deux vecteurs propres positifs linéairement indépendants.*

Preuve. Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive et irréductible. On sait déjà qu'une telle
matrice possède un vecteur propre positif associé à la valeur propre $\rho(A)$. Si
 x est un vecteur propre positif non nul associé à une valeur propre λ de A , il
existe alors un indice i compris entre 1 et n tel que $(Ax)_i = \lambda x_i$ avec $x_i > 0$ et
 $(Ax)_i \geq 0$, ce qui entraîne que $\lambda \geq 0$. On a aussi $(I + A)^{n-1}x = (1 + \lambda)^{n-1}x$, avec
 $(I + A)^{n-1}x > 0$ (puisque $(I + A)^{n-1} > 0$ et $x \geq 0$ non nul) et $(1 + \lambda)^{n-1} > 0$,
ce qui entraîne $x > 0$. En définitive un vecteur propre positif non nul de A est
nécessairement strictement positif.

Si x et y sont deux vecteurs propres non nuls de A associés aux valeurs propres
respectives λ et μ , on a alors $x > 0$, $y > 0$, $\lambda \geq 0$ et $\mu \geq 0$. Supposons $\lambda \leq \mu$. En
posant $\alpha = \max_{1 \leq i \leq n} \frac{y_i}{x_i}$, on a $\alpha x - y \geq 0$ et il existe un indice i compris entre 1 et n
tel que $\alpha x_i - y_i = 0$. Si $\alpha x - y$ est non nul, on a alors $(I + A)^{n-1}(\alpha x - y) > 0$, soit
 $\alpha(1 + \lambda)^{n-1}x > (1 + \mu)^{n-1}y$ et en particulier $\alpha(1 + \lambda)^{n-1}x_i > (1 + \mu)^{n-1}y_i$,
avec $\alpha x_i = y_i > 0$, ce qui entraîne $(1 + \lambda)^{n-1} > (1 + \mu)^{n-1}$ en contradiction avec
 $\lambda \leq \mu$. On a donc $\alpha x = y$, c'est-à-dire que les vecteurs x et y sont liés et $\lambda = \mu$. $\square$

On a vu (point 1. du lemme 4.5) que si A, B dans $\mathcal{M}_n(\mathbb{C})$ sont telles que
 $|B| \leq A$, on a alors $\rho(B) \leq \rho(|B|) \leq \rho(A)$. Avec le théorème qui suit on s'intéresse
au cas d'égalité $\rho(B) = \rho(A)$.

Lemme 4.12 *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice positive et irréductible. Si x est un
vecteur propre positif non nul tel que $\rho(A)x \leq Ax$, on a alors $Ax = \rho(A)x$ et x
est strictement positif.*

Preuve. Si la matrice A est positive irréductible, il en est alors de même de sa
transposée et le théorème de Perron-Frobenius nous assure l'existence d'un vecteur

strictement positif y tel que ${}^tAy = \rho(A)y$ et pour tout vecteur x dans $\mathbb{R}^n$, on a ;

$${}^ty(Ax - \rho(A)x) = \rho(A){}^tyx - \rho(A){}^tyx = 0$$

Si de plus on suppose que $\rho(A)x \leq Ax$, on a ${}^ty(Ax - \rho(A)x) = 0$ avec ${}^ty > 0$ et $Ax - \rho(A)x \geq 0$, ce qui équivaut à $Ax - \rho(A)x = 0$ (point 2. du lemme 4.3 en transposant). Le vecteur x est donc un vecteur propre positif non nul de A associé à $\rho(A)$, il est donc strictement positif (théorème de Perron-Frobenius). $\square$

Lemme 4.13 *Si $A \in \mathcal{M}_n(\mathbb{R})$ est positive irréductible et $B \in \mathcal{M}_n(\mathbb{K})$ telle que $|B| \leq A$, on a alors $\rho(B) \leq \rho(A)$. Si $\rho(B) = \rho(A)$, on a alors $|B| = A$. En particulier pour B positive, $B \leq A$ et $B \neq A$ entraîne $\rho(B) < \rho(A)$.*

Preuve. On sait déjà que $\rho(B) \leq \rho(A)$ (point 1. du lemme 4.5). Supposons que $\rho(B) = \rho(A)$. Dans ce cas, il existe une valeur propre de B de plus grand module de la forme $\lambda = e^{i\theta}\rho(A)$ et si x est un vecteur propre non nul associé dans $\mathbb{C}^n$, on a alors $\rho(A)|x| = |\lambda x| = |Bx| \leq |B||x| \leq A|x|$, soit $\rho(A)|x| \leq A|x|$ avec A positive irréductible et $|x|$ positif non nul. Le lemme précédent nous dit alors que $A|x| = \rho(A)|x|$ et que $|x|$ est strictement positif. On a donc $|Bx| = |B||x| = A|x|$, soit $(A - |B|)|x| = 0$ avec $A - |B| \geq 0$ et $|x| > 0$, ce qui entraîne $A = |B|$. $\square$

Théorème 4.8.

Soient $A \in \mathcal{M}_n(\mathbb{R})$ positive et irréductible et $B \in \mathcal{M}_n(\mathbb{K})$ telle que $|B| \leq A$. On a $\rho(B) \leq \rho(A)$. Si $\rho(B) = \rho(A)$ et si $\lambda = e^{i\theta}\rho(A)$ est une valeur propre de module maximum de B , il existe alors des réels $\theta_1, \dots, \theta_n$ tels que $B = e^{i\theta}\Delta B\Delta^{-1}$, où $\Delta = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$.

Preuve. Avec $\rho(B) = \rho(A)$, en reprenant la démonstration qui précède, si x est non nul tel que $Bx = \lambda x$, on a alors $A|x| = \rho(A)|x|$ avec $|x| > 0$ et $A = |B|$. Pour tout entier k compris entre 1 et n , on peut écrire $x_k = |x_k|e^{i\theta_k}$, avec $|x_k| > 0$ et $\theta_k \in]-\pi, \pi]$ et en désignant par Δ la matrice diagonale de termes diagonaux $e^{i\theta_1}, \dots, e^{i\theta_n}$, on a $x = \Delta|x|$. L'égalité $Bx = \lambda x$ s'écrit alors $B\Delta|x| = e^{i\theta}\rho(A)\Delta|x|$, ce qui donne $A|x| = \rho(A)|x| = e^{-i\theta}\Delta^{-1}B\Delta|x|$. En notant $C = e^{-i\theta}\Delta^{-1}B\Delta$, on a $|C| = |B| = A$ et l'égalité précédente s'écrit $|C||x| = C|x|$ avec $|x| > 0$, ce qui équivaut à $C = |C|$ (point 5. du lemme 4.2), soit à $e^{-i\theta}\Delta^{-1}B\Delta = A$ ou encore $B = e^{i\theta}\Delta B\Delta^{-1}$. $\square$

4.4 Matrices primitives

Le théorème 4.2 n'est pas valable si on suppose seulement la matrice A positive. L'exemple de la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ nous montre que l'hypothèse A positive et irréductible n'est pas suffisante non plus. C'est la notion de matrice primitive qui va nous permettre de prolonger ce théorème.

Définition 4.4. *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite primitive, si elle est positive irréductible et a une unique valeur propre de module maximale.*

Du théorème de Perron-Frobenius 4.7, on déduit que si la matrice A est primitive, $\rho(A)$ est alors l'unique valeur propre de module maximal et que cette valeur propre est simple. Les matrices strictement positives sont des cas particuliers de matrices primitives (théorème de Perron-Frobenius 4.1).

Il est facile de vérifier qu'une matrice est primitive si, et seulement si, sa transposée l'est.

Théorème 4.9.

Si $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice primitive, son rayon spectral est alors strictement positif et on a $\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)} A \right)^k = x {}^t y$, où x, y sont des vecteurs strictement positifs tels que $Ax = \rho(A)x$, ${}^t Ay = \rho(A)y$, ${}^t xy = 1$.

Preuve. La matrice A étant positive irréductible, le théorème de Perron-Frobenius 4.7 nous dit que $\rho(A)$ est strictement positif, que c'est une valeur propre simple et que l'espace propre associé est une droite vectorielle engendrée par un vecteur x strictement positif. La matrice ${}^t A$ ayant les mêmes propriétés, il existe un vecteur $z > 0$ tel que ${}^t Az = \rho(A)z$ et le vecteur $y = \frac{1}{{}^t z x} z$ est tel que ${}^t Ay = \rho(A)y$ et ${}^t xy = 1$. On peut donc définir la matrice $L = x {}^t y$ et cette matrice vérifie toutes les conditions du lemme 4.8. On a alors $\left(\frac{1}{\rho(A)} A \right)^k - L = \left(\frac{1}{\rho(A)} A - L \right)^k$ pour tout entier $k \geq 1$ et il s'agit de montrer que $\rho(A - \rho(A)L) < \rho(A)$. Si λ est une valeur propre non nulle de $B = A - \rho(A)L$ et z est un vecteur propre non nul associé, on a alors $(A - \rho(A)L)z = \lambda z$ et avec $L(A - \rho(A)L) = 0$, $\lambda \neq 0$, on déduit que $Lz = 0$, ce qui entraîne que $\lambda z = Az - \rho(A)Lz = Az$, c'est-à-dire que λ est valeur propre de A avec z pour vecteur propre associé et donc $|\lambda| \leq \rho(A)$. Si $|\lambda| = \rho(A)$, on a alors $\lambda = \rho(A)$ (A est primitive) et tout vecteur propre associé z est aussi vecteur propre de A , on a donc $z = \alpha x$ avec $\alpha \in \mathbb{C}^*$, ce qui entraîne que :

$$\rho(A)z = (A - \rho(A)L)z = (A - \rho(A)L)\alpha x = 0$$

qui est en contradiction avec $\rho(A) > 0$ et $z \neq 0$. On a donc $|\lambda| < \rho(A)$ pour toute valeur propre de B , ce qui donne le résultat annoncé. $\square$

Le résultat qui suit nous donne une définition équivalente de la notion de matrice primitive.

Théorème 4.10.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive. Cette matrice est primitive si, et seulement si, il existe un entier naturel p tel que A^p soit strictement positive.

Preuve. Si $A \in \mathcal{M}_n(\mathbb{R})$ est primitive, l'égalité $\lim_{k \rightarrow +\infty} \left(\frac{1}{\rho(A)} A \right)^k = L > 0$ entraîne alors que $A^p > 0$ pour p assez grand. Réciproquement si $A \in \mathcal{M}_n(\mathbb{R})$ est positive avec $A^p > 0$, elle est alors irréductible (lemme 4.9). Si λ est une valeur propre de A telle que $|\lambda| = \rho(A)$, λ^p alors est valeur propre de A^p avec $|\lambda^p| = \rho(A)^p$, ce qui entraîne $\lambda^p = \rho(A)^p$ puisque A^p est strictement positive. Si

$\lambda \neq \rho(A)$, il existe alors deux vecteurs x et y linéairement indépendants tels que $Ax = \rho(A)x$ et $Ay = \lambda y$, ce qui entraîne $A^p x = \rho(A)^p x$ et $A^p y = \lambda^p y = \rho(A)^p z$, c'est-à-dire que x et y sont dans l'espace propre de A^p associé à la valeur propre $\rho(A)^p$, mais on sait que cet espace propre est une droite vectorielle, les vecteurs y et z sont donc liés. On aboutit ainsi à une contradiction. La seule possibilité est donc $\lambda = \rho(A)$, ce qui achève de prouver que A est primitive. $\square$

Le résultat qui suit nous donne un critère plus pratique pour vérifier qu'une matrice positive irréductible est primitive.

Théorème 4.11.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ positive irréductible. Si tous les termes diagonaux de A sont strictement positifs, A^{n-1} est alors strictement positive et A est primitive.

Preuve. On désigne par B la matrice dans $\mathcal{M}_n(\mathbb{R})$ définie par $b_{ij} = a_{ij}$ pour $i \neq j$ et $b_{ii} = 0$, soit $B = A - \Delta$ où Δ est la matrice diagonale de termes diagonaux a_{ii} . La matrice B est irréductible comme A . En effet si B est réductible, il existe alors une partition (I, J) non triviale de $\{1, \dots, n\}$ telle $b_{ij} = 0$ pour $(i, j) \in I \times J$, ce qui entraîne $a_{ij} = 0$ pour $(i, j) \in I \times J$ puisque $i \neq j$ dans ce cas, et donc A est réductible. En notant α le plus petit des a_{ii} pour i compris entre 1 et n , on a $\alpha > 0$ et $A \geq \alpha I_n + B = \alpha \left(I_n + \frac{1}{\alpha} B \right)$, donc $A^{n-1} > \alpha^{n-1} \left(I_n + \frac{1}{\alpha} B \right)^{n-1} > 0$

du fait que $\frac{1}{\alpha} B$ est irréductible. $\square$

On peut montrer le résultat suivant (voir [9], corollaire 8.5.9).

Théorème 4.12. Wielandt

Soit A une matrice positive dans $\mathcal{M}_n(\mathbb{R})$. Cette matrice est primitive si, et seulement si, A^{n^2-2n+2} est strictement positive.

4.5 Matrices stochastiques et bistochastiques

Définition 4.5. On dit qu'une matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$ est stochastique si tous ses coefficients sont positifs et si :

$$\forall i \in \{1, \dots, n\}, \sum_{j=1}^n a_{i,j} = 1$$

Exemple 4.2 Une matrice de permutation est stochastique.

Si on note e le vecteur de $\mathbb{R}^n$ dont toutes les composantes sont égales à 1, une matrice positive $A \in \mathcal{M}_n(\mathbb{R})$ est stochastique si, et seulement si, $Ae = e$, ce qui revient à dire que 1 est valeur propre de A et que e est un vecteur propre associé.

De cette remarque, on peut déduire que le produit de deux matrices stochastiques est une matrice stochastique.

D'autre part si $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices stochastiques qui converge vers une matrice A , cette limite est alors également stochastique. L'ensemble $\mathcal{P}_n(\mathbb{R})$ des matrices stochastiques est donc fermé dans $\mathcal{M}_n(\mathbb{R})$ et étant borné (on a $\|A\|_\infty = 1$ pour toute matrice stochastique), on en déduit que c'est un compact.

Enfin il est facile de vérifier que $\mathcal{P}_n(\mathbb{R})$ est convexe.

Du point 5. du lemme 4.5, on déduit que le rayon spectral d'une matrice stochastique vaut 1 et que c'est une valeur propre. De plus e étant un vecteur propre strictement positif associé à la valeur propre $1 = \rho(A)$, on déduit du point 9. du lemme 4.5 que si x est un vecteur propre strictement positif d'une matrice stochastique, la valeur propre associée est alors 1. Ce même lemme nous dit que

$$\sup_{x > 0} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = \inf_{x > 0} \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} = 1.$$

On résume ces propriétés avec le théorème suivant.

Théorème 4.13.

L'ensemble $\mathcal{P}_n(\mathbb{R})$ des matrices stochastiques d'ordre n est convexe, compact et stable par la multiplication des matrices. Toute matrice stochastique a un rayon spectral égal 1 et ce rayon spectral est la seule valeur propre admettant des vecteurs propres strictement positifs.

Si A est une matrice stochastique primitive (par exemple strictement positive), $1 = \rho(A)$ est alors l'unique valeur propre dominante de A et cette valeur propre est simple. En désignant par f un vecteur propre strictement positif de ${}^tA$ associé à la valeur propre 1 tel que ${}^tfe = \sum_{i=1}^n f_i = 1$, on a d'après le théorème 4.9 :

$$\lim_{k \rightarrow +\infty} A^k = e {}^t f = \begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ f_1 & f_2 & \cdots & f_n \\ \vdots & \vdots & \ddots & \vdots \\ f_1 & f_2 & \cdots & f_n \end{pmatrix}$$

Définition 4.6. *On appelle matrice doublement stochastique, une matrice stochastique A telle que ${}^tA$ soit aussi stochastique.*

Si A est une matrice doublement stochastique, e est vecteur propre de A et ${}^tA$ associé à la valeur propre 1 et si de plus A est primitive, le résultat précédent donne alors $f = \frac{1}{n}e$ et :

$$\lim_{k \rightarrow +\infty} A^k = \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

Pour A stochastique irréductible, la suite de matrices $(A^k)_{k \geq 1}$ n'a pas nécessairement de limite comme le montre l'exemple de la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Par contre si on suppose de plus que tous les coefficients diagonaux de A sont strictement positifs, cette matrice est alors primitive (théorème 4.11) et le résultat précédent s'applique.

Les matrices de permutation sont des matrices doublement stochastiques et elles suffisent à décrire l'ensemble de toutes ces matrices, c'est ce que nous dit le théorème de Birkhoff qui suit. La démonstration de ce théorème nécessite quelques résultats sur les polyèdres convexes de $\mathbb{R}^n$.

On rappelle qu'une partie C de $\mathbb{R}^n$ est convexe, si pour tout couple (a, b) d'éléments de C , le segment $[a, b] = \{(1-t)a + tb \mid 0 \leq t \leq 1\}$ est contenu dans C .

Exemples 4.2

1. Une intersection d'ensembles convexes est convexe.
2. L'adhérence d'un convexe est convexe. En effet si C est convexe et $a = \lim_{k \rightarrow +\infty} a_k$, $b = \lim_{k \rightarrow +\infty} b_k$ sont dans l'adhérence $\overline{C}$ de C , les a_k et b_k étant dans C , on a alors pour tout $t \in [0, 1]$, $(1-t)a + tb = \lim_{k \rightarrow +\infty} ((1-t)a_k + tb_k) \in \overline{C}$.
3. Si u est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, alors pour tout convexe C dans $\mathbb{R}^n$ [resp. dans $\mathbb{R}^m$] l'image directe [resp. l'image réciproque] de C par u est un convexe de $\mathbb{R}^m$ [resp. $\mathbb{R}^n$].
4. En particulier, si φ est une forme linéaire non nulle sur $\mathbb{R}^n$, alors pour tout réel α l'ensemble $H_\alpha = \varphi^{-1}(\alpha) = \{x \in \mathbb{R}^n \mid \varphi(x) = \alpha\}$ est convexe dans $\mathbb{R}^n$ comme image réciproque du convexe $\{\alpha\}$ de $\mathbb{R}$ par l'application linéaire φ . Un tel convexe H_α est l'hyperplan affine d'équation $\varphi(x) = \alpha$.
5. Pour toute forme linéaire non nulle φ sur $\mathbb{R}^n$, les ensembles $H_\alpha^+ = \varphi^{-1}([\alpha, +\infty[)$ et $H_\alpha^- = \varphi^{-1}(]-\infty, \alpha])$ [resp. $H_\alpha^{+,*} = \varphi^{-1}([\alpha, +\infty[)$ et $H_\alpha^{-,*} = \varphi^{-1}(]-\infty, \alpha])$] sont convexes dans H_α . De tels convexes sont les demi-espaces fermés [resp. ouverts] limités par H_α . En dimension finie, l'application $x \mapsto \varphi(x) - \alpha$ est continue et en conséquence les demi-espaces H_α^+ et H_α^- [resp. $H_\alpha^{+,*}$ et $H_\alpha^{-,*}$] sont bien des fermés [resp. ouverts] de $\mathbb{R}^n$.

Définition 4.7. Un polyèdre dans $\mathbb{R}^n$ est une partie bornée de E qui peut s'écrire comme intersection d'un nombre fini de demi-espaces fermés.

Pour $n = 2$, on retrouve la notion de polygone.

Un polyèdre est fermé et convexe comme intersection d'ensembles fermés et convexes. Étant fermé et borné, il est compact dans $\mathbb{R}^n$.

Exemple 4.3 L'ensemble $P = \left\{ x \in (\mathbb{R}^+)^n \mid \|x\|_1 = \sum_{i=1}^n x_i = 1 \right\}$ est un polyèdre convexe (exercice 4.4).

Définition 4.8. Si X est une partie de $\mathbb{R}^n$, on appelle alors enveloppe convexe de X , l'intersection de tous les convexes de $\mathbb{R}^n$ qui contiennent X .

Cette enveloppe convexe est aussi le plus petit convexe de $\mathbb{R}^n$ contenant X .

On note $\text{Cv}(X)$ l'enveloppe convexe de X . C'est un convexe de $\mathbb{R}^n$.

Une définition équivalente de la notion d'enveloppe convexe est donnée par le résultat suivant.

Théorème 4.14.

Si X est une partie non vide de $\mathbb{R}^n$, l'enveloppe convexe de X est alors l'ensemble des combinaisons linéaires convexes d'éléments de X , c'est-à-dire qu'un vecteur x est dans $\text{Cv}(X)$ si et seulement si il existe des vecteurs $x_1, \dots, x_p$ dans X et des réels positifs ou nuls $\lambda_1, \dots, \lambda_p$ tels que $\sum_{i=1}^p \lambda_i = 1$ et $x = \sum_{i=1}^p \lambda_i x_i$, ce qui peut aussi s'exprimer en disant que x est barycentre de points de X affectés de coefficients positifs.

Preuve. Notons $\mathcal{B}(X)$ l'ensemble des combinaisons linéaires convexes d'éléments de X . On montre tout d'abord que $\mathcal{B}(X)$ est convexe, c'est donc un convexe de E contenant X et en conséquence il contient $\text{Cv}(X)$.

Si $x = \sum_{i=1}^p \lambda_i x_i$ et $y = \sum_{i=1}^q \mu_i y_i$ sont dans $\mathcal{B}(X)$, avec $\lambda_i \geq 0$, $x_i \in X$ pour $1 \leq i \leq p$, $\mu_i \geq 0$, $y_i \in X$ pour $1 \leq i \leq q$ et $\sum_{i=1}^p \lambda_i = \sum_{i=1}^q \mu_i = 1$, alors pour tout $t \in [0, 1]$, le vecteur $z = (1-t)x + ty$ s'écrit $z = \sum_{i=1}^p (1-t)\lambda_i x_i + \sum_{i=1}^q t\mu_i y_i$, avec $(1-t)\lambda_i > 0$, $t\mu_j \geq 0$, pour $1 \leq i \leq p$, $1 \leq j \leq q$ et :

$$\sum_{i=1}^p (1-t)\lambda_i + \sum_{i=1}^q t\mu_i = (1-t) + t = 1$$

ce qui signifie que $z \in \mathcal{B}(X)$.

On montre ensuite par récurrence sur $p \geq 1$ que toute combinaison linéaire convexe de p éléments de X est dans $\text{Cv}(X)$. Pour $p = 1$, le résultat découle de $X \subset \text{Cv}(X)$. Pour $p = 2$, si x_1, x_2 sont dans X et λ_1, λ_2 sont des réels positifs tels que $\lambda_1 + \lambda_2 = 1$, $\lambda_1 x_1 + \lambda_2 x_2$ est alors dans $\text{Cv}(X)$ puisque cet ensemble est convexe. Supposons le résultat acquis pour $p \geq 2$ et soient $x_1, \dots, x_{p+1}$ dans X ,

$\lambda_1, \dots, \lambda_{p+1}$ dans $\mathbb{R}^+$ tels que $\sum_{i=1}^{p+1} \lambda_i = 1$. Notons $\lambda = \sum_{i=1}^p \lambda_i$. Pour $\lambda = 0$, tous les

λ_i pour $1 \leq i \leq p$, sont nuls et $\lambda_{p+1} = 1$, de sorte que $\sum_{i=1}^{p+1} \lambda_i x_i = x_{p+1}$ est dans

$\text{Cv}(X)$. Pour $\lambda \neq 0$, on a par hypothèse de récurrence, $x' = \sum_{i=1}^p \frac{\lambda_i}{\lambda} x_i \in \text{Cv}(X)$ et

$\sum_{i=1}^{p+1} \lambda_i x_i = \lambda x' + \lambda_{p+1} x_{p+1} \in \text{Cv}(X)$ puisque $\lambda \geq 0$, $\lambda_{p+1} \geq 0$ et $\lambda + \lambda_{p+1} = 1$. On a donc ainsi montré que $\text{Cv}(X) = \mathcal{B}(X)$. $\square$

Le théorème de Carathéodory qui suit, nous permet de préciser que dans le résultat qui précède, on peut toujours avoir $p \leq n + 1$.

Théorème 4.15. Carathéodory

Si X est une partie non vide dans $\mathbb{R}^n$, tout élément de l'enveloppe convexe de X est alors combinaison linéaire convexe de $p \leq n + 1$ éléments de X .

Preuve. On sait déjà que tout élément de $\text{Cv}(X)$ s'écrit $x = \sum_{i=1}^p \lambda_i x_i$ avec $x_i \in X$,

$\lambda_i \in \mathbb{R}^+$ et $\sum_{i=1}^p \lambda_i = 1$. Pour $p \leq n + 1$, il n'y a rien à montrer. Supposons donc que $p > n + 1$. La famille de vecteurs $(x_i - x_1)_{2 \leq i \leq p}$ formé de $p - 1 > n$ vecteurs est alors liée et il existe des réels $\mu_2, \dots, \mu_p$ non tous nuls tels que $\sum_{i=2}^p \mu_i (x_i - x_1) = 0$, ce qui

peut aussi s'écrire en posant $\mu_1 = -\sum_{i=2}^p \mu_i$, $\sum_{i=1}^p \mu_i x_i = 0$ avec $\sum_{i=1}^p \mu_i = 0$. On peut

alors écrire, pour tout réel positif t , $x = \sum_{i=1}^p (\lambda_i - t\mu_i) x_i$, avec $\sum_{i=1}^p (\lambda_i - t\mu_i) = 1$.

Comme les coefficients μ_i sont tous non nuls de somme nulle, il en existe au moins un qui est strictement positif et on peut poser :

$$t_0 = \min \left\{ \frac{\lambda_i}{\mu_i} \mid 1 \leq i \leq p, \mu_i > 0 \right\} = \frac{\lambda_k}{\mu_k}$$

pour un indice k compris entre 1 et p . En notant $\delta_i = \lambda_i - t_0 \mu_i$, pour $1 \leq i \leq p$, on a $\delta_k = 0$, $\delta_i \geq 0$ pour $1 \leq i \leq p$ (pour $\mu_i > 0$, on a $\frac{\lambda_i}{\mu_i} \geq t_0$, soit $\delta_i \geq 0$ et

pour $\mu_i \leq 0$, $\delta_i \geq \lambda_i \geq 0$) et $\sum_{i=1}^p \delta_i = 1$. On a donc $x = \sum_{i=1, i \neq k}^p \delta_i x_i$, avec $\delta_i \geq 0$,

$\sum_{i=1, i \neq k}^p \delta_i = 1$, c'est-à-dire que x est combinaison linéaire convexe de $p - 1$ éléments de X . Une récurrence descendante nous permet alors de conclure. $\square$

Corollaire 4.7 : *L'enveloppe convexe d'un compact de $\mathbb{R}^n$ est compacte.*

Preuve. On désigne par Δ le compact de $\mathbb{R}^{n+1}$ défini par :

$$\Delta = \left\{ \lambda \in (\mathbb{R}^+)^{n+1} \mid \|\lambda\|_1 = \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$$

et, pour X compact non vide dans $\mathbb{R}^n$, par φ l'application définie par :

$$\forall (\lambda, x) \in \Delta \times X^{n+1}, \quad \varphi(\lambda, x) = \sum_{i=1}^{n+1} \lambda_i x_i$$

Le théorème de Carathéodory nous dit que l'image de φ est exactement l'enveloppe convexe de X , donc $\text{Cv}(X) = \varphi(\Delta \times X^{n+1})$ est compacte dans E comme image du compact $\Delta \times X^{n+1}$ (produit de compacts) par l'application continue φ . $\square$

Corollaire 4.8 : *L'enveloppe convexe d'une partie bornée de $\mathbb{R}^n$ est bornée.*

Preuve. Si X est une partie bornée de $\mathbb{R}^n$, elle est alors contenue dans une partie compacte Y et $\text{Cv}(X)$ est bornée car contenue dans le compact $\text{Cv}(Y)$. $\square$

Pour la démonstration du théorème de Birkhoff, nous aurons également besoins du théorème de projection sur un convexe fermé et du théorème de Krein-Milman qui suivent.

Pour ce qui suit, $\mathbb{R}^n$ est muni de sa structure euclidienne canonique.

Théorème 4.16. Projection orthogonale sur un convexe fermé

Soit C une partie non vide de $\mathbb{R}^n$ convexe et fermée. Pour tout $x \in \mathbb{R}^n$, il existe un unique $y \in \mathbb{R}^n$ tel que :

$$\|x - y\|_2 = \inf_{z \in C} \|x - z\|_2 \quad (4.2)$$

Ce vecteur $y \in \mathbb{R}^n$ est également caractérisé par :

$$\forall z \in C, \quad \langle x - y \mid z - y \rangle \leq 0 \quad (4.3)$$

Preuve. L'ensemble $\{\|x - z\|_2 \mid z \in C\}$ est une partie non vide minorée de $\mathbb{R}$, elle admet donc une borne inférieure $\delta = \inf_{z \in C} \|x - z\|_2$. Par définition de cette borne inférieure, on peut construire une suite $(y_k)_{k \geq 1}$ d'éléments de C telle que :

$$\forall k \geq 1, \quad \delta^2 \leq \|x - y_k\|_2^2 \leq \delta^2 + \frac{1}{k} \quad (4.4)$$

Pour $q > p \geq 1$, on a :

$$\begin{aligned} \|y_q - y_p\|_2^2 &= \|(y_q - x) + (x - y_p)\|_2^2 \\ &= 2 \left(\|y_q - x\|_2^2 + \|x - y_p\|_2^2 \right) - \|(y_q - x) - (x - y_p)\|_2^2 \end{aligned}$$

avec $\|(y_q - x) - (x - y_p)\|_2^2 = 4 \left\| x - \frac{1}{2}(y_p + y_q) \right\|_2^2 \geq 4\delta^2$ puisque $\frac{1}{2}(y_p + y_q) \in C$ (qui est convexe). On a donc, pour $q > p \geq 1$:

$$\|y_q - y_p\|_2^2 \leq 2 \left(\|y_q - x\|_2^2 + \|x - y_p\|_2^2 \right) - 4\delta^2 \leq 2 \left(\frac{1}{q} + \frac{1}{p} \right) - 2\delta^2 \leq \frac{4}{p}$$

et il en résulte que la suite $(y_k)_{k \geq 1}$ est de Cauchy dans l'espace complet $\mathbb{R}^n$, elle est donc convergente et sa limite y est dans C qui est fermé.

Et en faisant tendre k vers l'infini dans (4.4), on obtient $\|x - y\|_2 = \delta$. On a donc l'existence de $y \in C$ vérifiant (4.2) et il reste à montrer l'unicité.

Si z est un autre élément de C vérifiant (4.2), on a alors :

$$\begin{aligned} \|y - z\|_2^2 &= \|(y - x) + (x - z)\|_2^2 \\ &= 2 \left(\|y - x\|_2^2 + \|x - z\|_2^2 \right) - 4 \left\| x - \frac{1}{2}(y + z) \right\|_2^2 \\ &= 4\delta^2 - 4 \left\| x - \frac{1}{2}(y + z) \right\|_2^2 \leq 4\delta^2 - 4\delta^2 = 0 \end{aligned}$$

puisque $\frac{1}{2}(y + z) \in C$ (qui est convexe) et nécessairement $z = y$.

Soit y l'élément de C vérifiant (4.2), cet ensemble étant convexe, pour tout z dans C et tout $t \in]0, 1]$, le vecteur $v = (1 - t)y + tz$ est dans C et on a :

$$\|x - y\|_2^2 \leq \|x - v\|_2^2 = \|x - y\|_2^2 - 2 \langle x - y | z - y \rangle t + \|z - y\|_2^2 t^2$$

ce qui équivaut à :

$$\forall z \in C, \forall t \in]0, 1], 2 \langle x - y | z - y \rangle \leq \|z - y\|_2^2 t$$

En faisant tendre t vers 0^+ , on aboutit à :

$$\forall z \in C, \langle x - y | z - y \rangle \leq 0$$

Réciproquement supposons que $u \in C$ vérifie (4.3). Pour tout $z \in C$, on peut écrire :

$$\begin{aligned} \|x - z\|_2^2 &= \|(x - u) + (u - z)\|_2^2 \\ &= \|x - u\|_2^2 - 2 \langle x - u | z - u \rangle + \|u - y\|_2^2 \geq \|x - u\|_2^2 \end{aligned}$$

ce qui équivaut à dire que $\|x - u\|_2 = \inf_{z \in C} \|x - z\|_2$ et nécessairement $u = y$. $\square$

Avec les hypothèses et notations du théorème précédent, la borne inférieure $\delta = \inf_{z \in C} \|x - z\|_2$ est la distance $d(x, C)$ de x à C . Le vecteur $y \in C$ réalisant cette distance est la meilleure approximation de $x \in \mathbb{R}^n$ par des éléments du convexe C . Ce vecteur y étant également caractérisé par (4.3) est de ce fait aussi appelé la projection orthogonale de x sur C et noté $y = p_C(x)$. L'application p_C ainsi définie de $\mathbb{R}^n$ sur C est la projection orthogonale de $\mathbb{R}^n$ sur C .

Corollaire 4.9 : Soit C une partie convexe fermée non vide de $\mathbb{R}^n$ distincte de $\mathbb{R}^n$. Si pour x dans $\mathbb{R}^n \setminus C$, on désigne par D la droite vectorielle dirigée par $x - p_C(x)$ et par H l'hyperplan affine passant par x et orthogonal à D , soit :

$$H = x + D^\perp = \{z \in \mathbb{R}^n \mid \langle x - p_C(x) \mid x - z \rangle = 0\}$$

cet hyperplan contient alors x et C est contenu dans le demi-espace ouvert :

$$H^{+,*} = \{z \in \mathbb{R}^n \mid \langle x - p_C(x) \mid x - z \rangle > 0\}$$

Preuve. On a bien $x \in H$ et pour tout $z \in C$, on a :

$$\begin{aligned} \langle x - p_C(x) \mid x - z \rangle &= \langle x - p_C(x) \mid (x - p_C(x)) + (p_C(x) - z) \rangle \\ &= \|x - p_C(x)\|_2^2 - \langle x - p_C(x) \mid z - p_C(x) \rangle \\ &\geq \|x - p_C(x)\|_2^2 > 0 \end{aligned}$$

puisque $x \notin C$. □

On note $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$ et $(e_k^*)_{1 \leq k \leq n}$ sa base duale (définie par $e_i^* \left(\sum_{k=1}^n x_k e_k \right) = x_i$ pour $1 \leq i \leq n$).

Si X est une partie non vide de $\mathbb{R}^n$, sa frontière est alors $\text{Fr}(X) = \overline{X} \setminus \overset{\circ}{X}$.

Définition 4.9. Soit C un convexe dans $\mathbb{R}^n$ non vide et distinct de $\mathbb{R}^n$. On dit qu'un hyperplan affine H est un hyperplan d'appui de C si $H \cap C$ est non vide et C est contenu dans l'un des demi-espaces fermés limités par H .

On rappelle que pour toute forme linéaire φ sur $\mathbb{R}^n$, on peut trouver un unique vecteur $v \in \mathbb{R}^n$ tel que pour tout $x \in \mathbb{R}^n$, on ait $\varphi(x) = \langle x \mid v \rangle$ ($v = \sum_{k=1}^n \varphi(e_k) e_k$) et ce vecteur est non nul pour φ est non nulle.

Lemme 4.14 Soit C un convexe dans $\mathbb{R}^n$ non vide et distinct de $\mathbb{R}^n$. Si H est un hyperplan d'appui de C , tout point de $H \cap C$ est alors un point frontière de C .

Preuve. Soit H un hyperplan d'appui de C d'équation $\varphi(x) = \alpha$.

Soit $a \in H \cap C$ et supposons que $\varphi(x) = \langle x \mid v \rangle \geq \alpha$ pour tout x dans C . Si a n'est pas dans la frontière de C , il existe alors un réel $\varepsilon > 0$ tel que la boule ouverte $B(a, \varepsilon)$ de centre a et de rayon ε soit contenue dans C . Pour tout réel $t > 0$ assez petit, on a $a - tv \in C$ et :

$$\varphi(a - tv) = \langle a \mid v \rangle - t\|v\|_2^2 < \langle a \mid v \rangle = \alpha$$

en contradiction avec $\varphi(a - tv) \geq \alpha$. On a donc $a \in \text{Fr}(C)$. □

Exemple 4.4 Soit $C = \bigcap_{i=1}^p H_i^+ \subset \mathbb{R}^n$ un polyèdre convexe, où $H_i^+ = \varphi_i^{-1}([\alpha_i, +\infty[)$, les φ_i étant des formes linéaires non nulles sur E et les α_i des réels. Si $x \in C$

est tel que $\varphi_i(x) > \alpha_i$ pour tout i compris entre 1 et p , avec la continuité des applications φ_i , on déduit alors qu'il existe un réel $\varepsilon > 0$ tel que la boule ouverte $B(x, \varepsilon)$ de centre x et de rayon ε soit contenue dans $C = \bigcap_{i=1}^p H_i^+$ et en conséquence x est dans l'intérieur de C , donc $x \notin \text{Fr}(C)$. On a donc ainsi montré que pour tout $x \in \text{Fr}(C)$, il existe un indice i compris entre 1 et p tel que $\varphi_i(x) = \alpha_i$ et H_i est un hyperplan d'appui de C qui contient x . C'est-à-dire que tout point de la frontière de C est contenu dans un hyperplan d'appui.

En fait ce résultat est valable pour tout convexe fermé dans $\mathbb{R}^n$ euclidien.

Lemme 4.15 *Si C est un convexe fermé dans $\mathbb{R}^n$ non vide et distinct de $\mathbb{R}^n$, tout point de la frontière de C est alors contenu dans un hyperplan d'appui de C .*

Preuve. Soit a dans la frontière de C . Pour tout entier naturel non nul k on peut trouver un élément x_k dans la boule ouverte $B\left(a, \frac{1}{k}\right)$ de centre a et de rayon $\frac{1}{k}$ qui n'appartient pas à C . On note $y_k = p_C(x_k)$ la projection de x_k sur C , $z_k = \frac{1}{\|x_k - y_k\|_2} (x_k - y_k)$ et $H_k = \{z \in E \mid \langle z_k \mid x_k - z \rangle = 0\}$ l'hyperplan affine passant par x_k et orthogonal à z_k . Le corollaire 4.9 nous dit alors que C est contenu dans le demi-espace ouvert $H_k^{+,*} = \{z \in E \mid \langle z_k \mid x_k - z \rangle > 0\}$. Chaque vecteur z_k étant dans la sphère unité de E qui est compacte, on peut extraire de la suite $(z_k)_{k \geq 1}$ une sous-suite $(z_{\varphi(k)})_{k \geq 1}$ qui converge vers un vecteur v de norme 1. En considérant que la suite $(x_k)_{k \geq 1}$ converge vers a et en utilisant la continuité du produit scalaire, on déduit que pour tout z dans C on a :

$$\langle v \mid a - z \rangle = \lim_{k \rightarrow +\infty} \langle z_{\varphi(k)} \mid x_{\varphi(k)} - z \rangle \geq 0$$

c'est-à-dire que C est contenu dans le demi-espace $H^+ = \{z \in E \mid \langle v \mid a - z \rangle \geq 0\}$, le vecteur a étant dans l'hyperplan H d'équation $\langle v \mid a - z \rangle = 0$. Cet hyperplan H est donc un hyperplan d'appui de C . $\square$

Par analogie à la notion de sommet d'un polygone dans $\mathbb{R}^2$, on définit de manière plus générale les sommets, ou points extrémaux, d'un convexe de la manière suivante.

Définition 4.10. *Soit C un convexe non vide de $\mathbb{R}^n$. On dit qu'un point a de C est un point extrême si tout segment dans C qui contient a admet ce point pour extrémité.*

Dire que a dans le convexe C est extrême équivaut à dire que si $a \in [x, y]$ avec x, y dans C , on a alors $a = x$ ou $a = y$, ce qui est encore équivalent à dire que si $a = (1 - t)x + ty$ avec x, y dans C et $0 < t < 1$, on a alors $a = x = y$.

Une définition équivalente de point extrême d'un convexe est donnée par le résultat suivant.

Lemme 4.16 *Soit C un convexe non vide de $\mathbb{R}^n$. Un point a de C est extrême si, et seulement si, $C \setminus \{a\}$ est convexe.*

Preuve. Soit $a \in C$ extrême et x, y dans $C \setminus \{a\}$. L'ensemble C étant convexe, on a $[x, y] \subset C$ et $a \in [x, y]$ entraîne $x = a$ ou $y = a$, ce qui est exclu. On a donc $[x, y] \subset C \setminus \{a\}$. On a donc ainsi montré que $C \setminus \{a\}$ est convexe. Réciproquement, supposons $C \setminus \{a\}$ convexe et soit $[x, y]$ un segment dans C qui contient a . Si x, y sont tous deux dans $C \setminus \{a\}$ qui est convexe, on a alors $a \in [x, y] \subset C \setminus \{a\}$, ce qui est impossible. On a donc $x = a$ ou $y = a$, ce qui prouve que a est extrême dans C . $\square$

Exemple 4.5 Les points extrêmes de $P = \left\{ x \in (\mathbb{R}^+)^n \mid \|x\|_1 = \sum_{i=1}^n x_i = 1 \right\}$ sont les vecteurs de la base canonique (exercice 4.6).

De manière plus générale, un convexe compact admet des points extrêmes.

Lemme 4.17 Un convexe compact non vide dans $\mathbb{R}^n$ a des points extrêmes.

Preuve. Soit C un convexe compact non vide dans $\mathbb{R}^n$. L'application e_1^* étant continue sur le compact C , elle y est bornée et atteint sa borne inférieure, on peut donc poser $t_1 = \inf_{x \in C} e_1^*(x)$ (t_1 est la plus petite des premières composantes

d'éléments de C). L'ensemble $C_1 = \left\{ x = t_1 e_1 + \sum_{i=2}^n x_i e_i \mid x \in C \right\}$ est alors un compact non vide de E (il est fermé et borné) et on peut poser $t_2 = \inf_{x \in C_1} e_2^*(x)$.

En continuant ainsi de suite, on construit un vecteur $u = \sum_{i=1}^n t_i e_i$ dans C et on vérifie qu'il est extrême. Si $u = (1-t)x + ty$ avec x, y dans C et t dans $]0, 1[$, de $t_1 = (1-t)x_1 + ty_1$ avec $t_1 \leq x_1$, $t_1 \leq y_1$ et $0 < t < 1$, on déduit que nécessairement $t_1 = x_1 = y_1$. Puis par récurrence, vu la construction des t_k , on déduit que $t_k = x_k = y_k$ pour tout k compris entre 1 et n . On a donc $u = x = y$, ce qui prouve que u est extrême dans C . $\square$

Lemme 4.18 Si C est un convexe compact non vide dans $\mathbb{R}^n$, pour tout hyperplan d'appui H de C , $C \cap H$ (qui est convexe compact et non vide) admet alors des points extrêmes qui sont aussi des points extrêmes de C .

Preuve. Notons $H = \{x \in \mathbb{R}^n \mid \varphi(x) = \alpha\}$ un hyperplan d'appui de C . On a $\varphi(x) \geq \alpha$ pour tout $x \in C$ et H contient un point frontière de C . L'intersection $C \cap H$ est alors un convexe compact non vide et il admet des points extrêmes (lemme 4.17). Soit x un point extrême de $C \cap H$. Si il existe y, z dans C et t dans $]0, 1[$ tels que $x = (1-t)y + tz$, on a alors :

$$\alpha = \varphi(x) = (1-t)\varphi(y) + t\varphi(z) > (1-t)\alpha + t\alpha = \alpha$$

si $\varphi(y) > \alpha$ ou $\varphi(z) > \alpha$, ce qui est impossible. On a donc $\varphi(y) = \alpha$ et $\varphi(z) = \alpha$, c'est-à-dire que y et z sont dans $C \cap H$ et $y = z = x$ puisque x est un point extrême de $C \cap H$. En conclusion x est un point extrême de C . $\square$

Théorème 4.17. Krein-Milman

Tout compact convexe dans $\mathbb{R}^n$ est l'enveloppe convexe de ses points extrémaux.

Preuve. Soit C un convexe compact non vide dans $\mathbb{R}^n$. On note $S(C)$ l'enveloppe convexe de l'ensemble des points extrémaux de C . On a $S(C) \subset C$.

Supposons qu'il existe $a \in C$ qui n'est pas dans $S(C)$. On a alors $a \notin \overline{S(C)}$, ce dernier ensemble étant convexe (l'adhérence d'un convexe est convexe) et fermé dans E . On peut donc utiliser le corollaire 4.9 pour dire qu'il existe un hyperplan affine d'équation $\varphi(x) = \alpha$ contenant a et tel que $\varphi(x) > \alpha$ pour tout $x \in \overline{S(C)}$. L'image de C par φ est convexe (image d'un convexe par l'application linéaire φ) et compacte (image du compact C par l'application continue φ) dans $\mathbb{R}$, c'est donc un intervalle réel $[u, v]$ qui contient α (puisque $\varphi(a) = \alpha$). Désignons par K l'hyperplan affine d'équation $\varphi(x) = u$. On a $K \cap C \neq \emptyset$ et $\varphi(x) \geq u$ pour tout $x \in C$ car $[u, v] = \varphi(C)$, c'est-à-dire que K est un hyperplan d'appui de C et le lemme 4.18 nous dit que K contient des points extrémaux de C , si x est l'un de ces points, il est alors dans $S(C)$ et $\varphi(x) > \alpha \geq u$ en contradiction avec $\varphi(x) = u$ (x est dans K). On a donc $C \subset S(C)$ et $C = S(C)$. $\square$

On note $\mathcal{B}_n(\mathbb{R})$ l'ensemble des matrices bistochastiques dans $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{E}_n(\mathbb{R})$ le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ constitué des matrices $A = ((a_{ij}))_{1 \leq i, j \leq n}$ véri-

fiant $\sum_{k=1}^n a_{ik} = \sum_{j=1}^n a_{kj} = 0$ pour tous i, j compris entre 1 et n .

Lemme 4.19 *L'espace vectoriel $\mathcal{E}_n(\mathbb{R})$ est de dimension égale à $(n-1)^2$.*

Preuve. On note φ l'application linéaire qui associe à toute matrice $X = ((x_{i,j}))_{1 \leq i, j \leq n}$ dans $\mathcal{E}_n(\mathbb{R})$ la matrice $Y = ((y_{i,j}))_{1 \leq i, j \leq n-1}$. Si $X \in \mathcal{E}_n(\mathbb{R})$ est telle que $\varphi(X) = 0$, on a alors $x_{ij} = 0$ pour tous i, j compris entre 1 et $n-1$, $x_{in} = -\sum_{k=1}^{n-1} x_{ik} = 0$ pour i compris entre 1 et $n-1$, et $x_{nn} = -\sum_{k=1}^{n-1} x_{kn} = 0$, c'est-à-dire que $X = 0$ et φ est injective.

Pour toute matrice $Y \in \mathcal{M}_{n-1}(\mathbb{R})$, en posant :

$$x_{in} = -\sum_{k=1}^{n-1} y_{ik}, \quad x_{nj} = -\sum_{k=1}^{n-1} y_{kj}, \quad (1 \leq i, j \leq n-1)$$

on a $\sum_{i=1}^{n-1} x_{i,n} = -\sum_{i=1}^{n-1} \sum_{k=1}^{n-1} y_{ik} = -\sum_{k=1}^{n-1} \sum_{i=1}^{n-1} y_{ik} = \sum_{k=1}^{n-1} x_{nk}$, de sorte que la matrice :

$$X = \begin{pmatrix} y_{11} & \cdots & y_{1,n-1} & x_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ y_{n-1,1} & \cdots & y_{n-1,n-1} & x_{n-1,n} \\ x_{n1} & \cdots & x_{n,n-1} & -\sum_{i=1}^{n-1} x_{i,n} \end{pmatrix}$$

est dans $\mathcal{E}_n(\mathbb{R})$ avec $\varphi(X) = Y$. L'application φ est donc surjective. Il en résulte que c'est un isomorphisme et $\dim(\mathcal{E}_n(\mathbb{R})) = \dim(\mathcal{M}_{n-1}(\mathbb{R})) = (n-1)^2$. $\square$

Lemme 4.20 $\mathcal{B}_n(\mathbb{R})$ est un polyèdre convexe dans $\mathcal{M}_n(\mathbb{R})$ et les points extrémaux de ce polyèdre sont les matrices de permutations.

Preuve. Il est facile de vérifier que $\mathcal{B}_n(\mathbb{R})$ est fermé et borné, donc compact. Dire que $A = ((a_{i,j}))_{1 \leq i,j \leq n}$ est dans $\mathcal{B}_n(\mathbb{R})$ équivaut à :

$$a_{ij} \geq 0, \sum_{k=1}^n a_{ik} = 1, \sum_{j=1}^n a_{kj} = 1 \quad (1 \leq i, j \leq n)$$

En notant $(E_{ij}^*)_{1 \leq i,j \leq n}$ la base duale de la base canonique de $\mathcal{M}_n(\mathbb{R})$ (définie par $E_{ij}^*(A) = a_{ij}$), et pour $1 \leq i, j \leq n$, L_i^* et C_j^* les formes linéaires définies sur $\mathcal{M}_n(\mathbb{R})$ par $L_i^* = \sum_{k=1}^n E_{ik}^*$, $C_j^* = \sum_{k=1}^n E_{kj}^*$, on déduit que $A \in \mathcal{B}_n(\mathbb{R})$ équivaut à :

$$E_{ij}^*(A) \geq 0, L_i^*(A) = 1, C_j^*(A) = 1 \quad (1 \leq i, j \leq n)$$

Il en résulte que $\mathcal{B}_n(\mathbb{R})$ est un polyèdre convexe dans $\mathcal{M}_n(\mathbb{R})$.

Soit A une matrice de permutation et supposons que $A = (1-t)X + tY$ avec X, Y dans $\mathcal{B}_n(\mathbb{R})$ et $0 < t < 1$. Sur chaque ligne i de la matrice A il y a un seul coefficient non nul $a_{ij} = 1$ et pour $k \neq j$, on a $0 = a_{ik} = (1-t)x_{ik} + ty_{ik}$ qui entraîne que $x_{ik} = y_{ik} = 0$ et $x_{ij} = y_{ij} = 1$ puisque tous les coefficients d'une matrice bistochastique sont positifs ou nuls et la somme des termes d'une même ligne vaut 1. On a donc $X = Y = A$. On a donc ainsi montré que les matrices de permutation sont des points extrémaux de $\mathcal{B}_n(\mathbb{R})$.

Pour la réciproque, on procède par récurrence sur $n \geq 2$. Pour $n = 2$, une matrice stochastique est de la forme $A = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$ avec $0 \leq a \leq 1$. Si cette matrice est de permutation c'est alors un point extrême de $\mathcal{B}_2(\mathbb{R})$, sinon on a $0 < a < 1$. Supposons, ce qui n'est pas restrictif, que $0 < a \leq 1-a < 1$, soit $0 < a \leq \frac{1}{2}$. En posant :

$$\begin{aligned} X &= A + a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ Y &= A - a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2a & 1-2a \\ 1-2a & 2a \end{pmatrix} \end{aligned}$$

on a $X \in \mathcal{B}_2(\mathbb{R})$, $Y \in \mathcal{B}_2(\mathbb{R})$, $A = \frac{1}{2}(X + Y)$, avec $A \neq X$, $A \neq Y$, ce qui signifie que A n'est pas extrême. Le résultat est donc montré pour $n = 2$. Supposons le acquis pour $n-1 \geq 2$ et soit $A \in \mathcal{B}_n(\mathbb{R})$ un élément extrême. On montre tout d'abord que la matrice A a au plus $2n-1$ coefficients non nuls. Supposons que A ait au moins $2n$ coefficients non nuls que nous notons a_{i_k, j_k} avec $1 \leq k \leq 2n$ et les couples (i_k, j_k) deux à deux distincts. On désigne par H le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ de dimension $2n$ engendré par les matrices E_{i_k, j_k} , pour $1 \leq k \leq 2n$,

où $(E_{ij})_{1 \leq i, j \leq n}$ désigne la base canonique de $\mathcal{M}_n(\mathbb{R})$. On a $H \cap \mathcal{E}_n(\mathbb{R}) \neq \{0\}$ à cause des dimensions. Il existe donc une matrice $B \in H \cap \mathcal{E}_n(\mathbb{R}) \setminus \{0\}$. Si pour tout réel t , on note $C_t = A + tB$, on a alors $c_{ij} = a_{ij}$ pour $(i, j) \neq (i_k, j_k)$ et $c_{i_k, j_k} = a_{i_k, j_k} + tb_{i_k, j_k} > 0$ pour t voisin de 0 (on a $a_{i_k, j_k} > 0$ pour tout k), puis avec $B \in \mathcal{E}_n(\mathbb{R})$, on déduit que $\sum_{k=1}^n c_{ik} = \sum_{k=1}^n a_{ik} = 1$, $\sum_{k=1}^n c_{kj} = \sum_{k=1}^n a_{kj} = 1$ pour tous i, j . Pour $t \in]-\varepsilon, \varepsilon[\setminus \{0\}$ avec $\varepsilon > 0$ assez petit, les matrices C_t et C_{-t} sont donc dans $\mathcal{B}_n(\mathbb{R})$ et $A = \frac{1}{2}(C_t + C_{-t})$ avec $A \neq C_t$, $A \neq C_{-t}$, en contradiction avec A extrême. La matrice A a donc au plus $2n - 1$ termes non nuls et il existe nécessairement une ligne d'indice i de cette matrice avec un seul coefficient a_{ij} non nul, ce coefficient valant 1. La matrice A étant bistochastique, tous les coefficients de la colonne j , excepté celui en ligne i , sont nuls. La matrice A' extraite de A en supprimant la ligne i et la colonne j est alors dans $\mathcal{B}_{n-1}(\mathbb{R})$ et extrême. En effet si $A' \in [B', C']$ avec B', C' dans $\mathcal{B}_{n-1}(\mathbb{R})$, on a alors $A \in [B, C]$, où B, C sont dans $\mathcal{B}_n(\mathbb{R})$ telles que $b_{ij} = c_{ij} = 1$ et B', C' sont extraites de B, C en supprimant la ligne i et la colonne j et $A = B$ ou $A = C$, ce qui entraîne $A' = B'$ ou $A' = C'$. Avec l'hypothèse de récurrence, on déduit alors que A' est une matrice de permutation dans $\mathcal{M}_{n-1}(\mathbb{R})$ et A est une matrice de permutation dans $\mathcal{M}_n(\mathbb{R})$. $\square$

De ce résultat et des théorèmes de Krein-Milman et de Carathéodory (théorèmes 4.17 et 4.15), on déduit enfin le résultat suivant.

Théorème 4.18. Birkhoff

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est doublement stochastique si, et seulement si, elle s'écrit $A = \sum_{k=1}^p \lambda_k P_k$, où $p \leq (n - 1)^2 + 1$, les P_k sont des matrices de permutation et les λ_k des réels positifs tels que $\sum_{k=1}^p \lambda_k = 1$.

Preuve. L'ensemble des matrices bistochastiques étant un polyèdre convexe dans l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$, le théorème de Krein-Milman nous dit que c'est l'enveloppe convexe de ses points extrémaux, donc de l'ensemble des matrices de permutations.

Toute matrice $A \in \mathcal{B}_n(\mathbb{R})$ s'écrit donc $A = \sum_{k=1}^q \mu_k P_k$, où les P_k sont des

matrices de permutation et les μ_k des réels positifs tels que $\sum_{k=1}^q \mu_k = 1$. En écrivant

$A - I_n = \sum_{k=1}^q \mu_k (P_k - I_n)$, on déduit que $A - I_n$ est dans $\text{Cv}(X) \subset \mathcal{E}_n(\mathbb{R})$ formé des matrices $P_\sigma - I_n$, où σ parcourt l'ensemble des permutations de $\{1, \dots, n\}$.

Le théorème de Carathéodory nous dit alors que $A - I_n = \sum_{k=1}^p \lambda_k (P_k - I_n)$, où

$p \leq \dim(\mathcal{E}_n(\mathbb{R})) + 1 = (n-1)^2 + 1$, les P_k sont des matrices de permutation et les λ_k des réels positifs tels que $\sum_{k=1}^p \lambda_k = 1$, ce qui équivaut à $A = \sum_{k=1}^p \lambda_k P_k$. $\square$

4.6 Exercices

Exercice 4.1. Pour $n \geq 2$, on note $J_n = \{1, \dots, n\}$, $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$ et on désigne par $A \in \mathcal{M}_n(\mathbb{R})$ une matrice positive. Pour i, j dans J_n , $m \in \mathbb{N}^*$, on note $\mathcal{L}(i, j, m)$ la proposition :

$$\exists (i_0, \dots, i_m) \in J_n^{m+1} \text{ tel que } \begin{cases} i_0 = i, i_m = j \\ \forall k \in \{0, \dots, m-1\}, a_{i_k, i_{k+1}} \neq 0 \end{cases}$$

et $\mathcal{L}(i, j)$ la proposition :

$$\exists m \in \mathbb{N}^* \text{ tel que } \mathcal{L}(i, j, m) \text{ est vraie}$$

1. Montrer que la propriété $\mathcal{L}$ est transitive dans le sens où $\mathcal{L}(i, j)$ et $\mathcal{L}(j, r)$ vraies pour i, j, r dans J_n entraîne $\mathcal{L}(i, r)$ vraie.
2. Montrer que si A est réductible et si (I, J) est une partition non triviale de J_n telle que $a_{i,j} = 0$ pour tout $(i, j) \in I \times J$, alors pour tout couple $(i, j) \in I \times J$, $\mathcal{L}(i, j)$ n'est pas vraie.
3. Soit j fixé dans J_n . On note J la partie de J_n constituée de j et des indices $i \in J_n$ tels que $\mathcal{L}(i, j)$ est vraie. Montrer que le sous-espace $V_J = \text{Vect}\{e_i \mid i \in J\}$ est stable par A .
4. Dédire de ce qui précède l'équivalence :

$$(A \text{ irréductible}) \Leftrightarrow (\text{Pour tout } (i, j) \in J_n^2, \mathcal{L}(i, j) \text{ est vraie})$$

5. On suppose que $\mathcal{L}(i, j)$ est vraie, avec $i \neq j$. Montrer qu'il existe m dans J_{n-1} tel que $\mathcal{L}(i, j, m)$ soit vraie.
6. Pour tout $m \in \mathbb{N}^*$, on pose $A^m = \left(\left(a_{i,j}^{(m)} \right) \right)_{1 \leq i, j \leq n}$. Montrer que pour $i \neq j$ on a l'équivalence :

$$(\mathcal{L}(i, j, m) \text{ est vraie}) \Leftrightarrow \left(a_{i,j}^{(m)} > 0 \right)$$

7. En conclure que les trois assertions suivantes sont équivalentes :

- (a) la matrice A est irréductible ;
- (b) pour tout $(i, j) \in J_n^2$ tel que $i \neq j$, il existe $m \in J_{n-1}$ tel que $a_{i,j}^{(m)} > 0$;
- (c) $(I_n + A)^{n-1} > 0$.

Solution.

1. Si $\mathcal{L}(i, j)$ et $\mathcal{L}(j, r)$ sont vraies, il existe alors des entiers naturels non nuls m et p , des suites d'éléments dans J_n , $(i_0 = i, i_1, \dots, i_m = j)$ et $(j_0 = j, j_1, \dots, j_p = r)$ tels que $a_{i_k, i_{k+1}} \neq 0$ pour $k = 0, \dots, m-1$ et $a_{j_k, j_{k+1}} \neq 0$ pour $k = 0, \dots, p-1$. En posant $q = m + p$ et :

$$(s_0, s_1, \dots, s_q) = (i_0 = i, i_1, \dots, i_m = j_0 = j, j_1, \dots, j_p = r)$$

on a $a_{s_k, s_{k+1}} \neq 0$ pour $k = 0, \dots, q-1$, c'est-à-dire que $\mathcal{L}(i, r)$ est vraie.

2. Supposons la matrice A réductible et soit (I, J) une partition non triviale de J_n telle que $a_{ij} = 0$ pour tout $(i, j) \in I \times J$. Pour $(i, j) \in I \times J$, $m \geq 1$ et $(i_0, \dots, i_m) \in J_n^{m+1}$ avec $i_0 = i$, $i_m = j$, en désignant par k le plus grand indice compris entre 0 et $m-1$ tel que $i_k \in I$, on a $(i_k, i_{k+1}) \in I \times J$ et $a_{i_k, i_{k+1}} = 0$, la proposition $\mathcal{L}(i, j, m)$ est donc fausse pour tout $m \geq 1$ et également la proposition $\mathcal{L}(i, j)$.

3. Supposons V_J non stable par A , il existe alors un indice $i \in J$ tel que $Ae_i \notin V_J$, soit $Ae_i = \sum_{k=1}^n a_{ki}e_k$ avec $a_{ki} \neq 0$ pour un indice $k \notin J$. On a donc $k \neq j$. La propriété $\mathcal{L}(k, i, 1)$ est donc vraie $((i_0, i_1) = (k, i) \in J_n^2$ et $a_{i_0, i_1} = a_{ki} \neq 0)$ ainsi que $\mathcal{L}(k, i)$. Si $i = j$, alors $\mathcal{L}(k, j)$ est vraie avec $k \neq j$, ce qui signifie que $k \in J$ en contradiction avec $k \notin J$. On a donc $i \neq j$ et $i \in J$, ce qui signifie que $\mathcal{L}(i, j)$ est vraie. Mais $\mathcal{L}(k, i)$ et $\mathcal{L}(i, j)$ vraies entraîne $\mathcal{L}(k, j)$ vraie par transitivité et comme $k \neq j$, cela veut dire que $k \in J$ toujours en contradiction avec $k \notin J$. En conclusion V est stable par A .

4. On a vu en 2 que si A est réductible il existe alors un couple (i, j) dans J_n^2 tel que $\mathcal{L}(i, j)$ n'est pas vraie. On déduit donc que si $\mathcal{L}(i, j)$ est vraie pour tout (i, j) dans J_n^2 , A est alors irréductible. Réciproquement supposons A irréductible. Pour tout $j \in J_n$, l'ensemble J de la question précédente n'est pas vide ($j \in J$) et l'espace V_J associé est stable par A , ce qui entraîne $J = J_n$ (théorème 4.5), encore équivalent à dire que $\mathcal{L}(i, j)$ est vraie pour tout $i \in J_n \setminus \{j\}$. Si $n \geq 2$, pour tout $j \in J_n$ il existe alors $i \neq j$ dans J_n et $\mathcal{L}(i, j)$, $\mathcal{L}(j, i)$ vraies entraîne par transitivité que $\mathcal{L}(j, j)$ est vraie. On a donc bien $\mathcal{L}(i, j)$ est vraie pour tout $i \in J_n$.

5. Si on suppose $\mathcal{L}(i, j)$ vraie avec $i \neq j$, il existe alors un entier $m \geq 1$ et un m -uplet d'entiers $(i_0, \dots, i_m)$ avec $i_0 = i$, $i_m = j$ tels que $a_{i_k, i_{k+1}} \neq 0$ pour $k = 0, \dots, m-1$. On désigne par m le plus petit entier naturel non nul vérifiant cette propriété. Pour cet indice m les entiers i_k sont nécessairement deux à deux distincts. En effet si $i_p = i_q$ avec $0 \leq p < q \leq m$, alors la suite $(i_k)_{0 \leq k \leq m}$ privée de i_q convient aussi, ce qui contredit la caractère minimal de m . On a donc $\text{card} \{i_0, \dots, i_m\} = m+1$ avec $\{i_0, \dots, i_m\} \subset J_n$, donc $m+1 \leq n$, soit $m \leq n-1$.

6. Pour $m \geq 2$ et i, j dans J_n , on a $a_{ij}^{(m)} = \sum_{k=1}^n a_{ik}^{(m-1)} a_{kj}$. On montre par récurrence sur $m \geq 1$ que pour $i \neq j$ dans J_n , on a $\mathcal{L}(i, j, m)$ vraie si, et seulement si, $a_{ij}^{(m)}$ est strictement positif. Pour $m = 1$ c'est la définition de $\mathcal{L}(i, j, 1)$. En supposant

le résultat acquis au rang $m - 1$, on a $a_{ij}^{(m)} > 0$ si, et seulement si, il existe un entier $k \in J_n$ tel que $a_{ik}^{(m-1)} > 0$ et $a_{kj} > 0$ et avec l'hypothèse de récurrence cela équivaut à l'existence de $k \in J_n$ tel que $\mathcal{L}(i, k, m - 1)$ et $\mathcal{L}(k, j, 1)$ soient vraies, ce qui équivaut à l'existence d'entiers $i_0 = i, i_2, \dots, i_{m-1} = k$ dans J_n tels que $a_{i_p, i_{p+1}} \neq 0$ pour tout p compris entre 0 et $m - 2$ et $a_{kj} \neq 0$, encore équivalent à $\mathcal{L}(i, j, m)$ vraie.

7. Si la matrice positive A est irréductible, pour $i \neq j$ dans J_n la propriété $\mathcal{L}(i, j)$ est alors vraie, ce qui entraîne l'existence de $m \in J_{n-1}$ tel que $\mathcal{L}(i, j, m)$ soit vraie, encore équivalent à $a_{ij}^{(m)} > 0$. Supposons que pour tous $i \neq j$ dans J_n il existe $m \in J_{n-1}$ (dépendant de i et j) tel que $a_{ij}^{(m)} > 0$. Avec :

$$(I_n + A)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} A^k = I_n + \sum_{k=1}^{n-1} \binom{n-1}{k} A^k$$

on déduit que le coefficient d'indice (i, i) de cette matrice est :

$$c_{ii} = 1 + \sum_{k=1}^{n-1} \binom{n-1}{k} a_{ii}^{(k)} \geq 1 > 0$$

et pour $i \neq j$, du fait qu'il existe un indice m compris entre 1 et $m - 1$ tel que $a_{ij}^{(m)} > 0$, on a pour le coefficient d'indice (i, j) de $(I_n + A)^{n-1}$:

$$c_{ij} = \sum_{k=1}^{n-1} \binom{n-1}{k} a_{ij}^{(k)} \geq C_{n-1}^m a_{ij}^{(m)} > 0$$

En définitive la matrice $(I_n + A)^{n-1}$ est strictement positive. Si la matrice $(I_n + A)^{n-1}$ est strictement positive, on a alors $\sum_{k=1}^{n-1} \binom{n-1}{k} a_{ij}^{(k)} > 0$ pour tous $i \neq j$ dans J_n et il existe $m \in J_{n-1}$ tel que $\mathcal{L}(i, j, m)$ soit vraie, ce qui entraîne que $\mathcal{L}(i, j)$ est vraie. On a donc $\mathcal{L}(i, j)$ vraie pour tous $i \neq j$ dans J_n , ce qui équivaut à $\mathcal{L}(i, j)$ vraie pour tous i, j dans J_n puisque $n \geq 2$, équivalent à dire que A est irréductible.

Exercice 4.2. Montrer que, pour tout entier $n \geq 3$, la matrice de Wei-

$$\text{landt } A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 0 \end{pmatrix} \text{ est primitive.}$$

Solution. On utilise les notations et résultats de l'exercice précédent. Le coefficient $a_{i, i+1}$ pour i compris entre 1 et $n - 1$ et le coefficient $a_{n, 1}$ étant non nul, on déduit que $\mathcal{L}(i, j)$ est vraie pour tous i, j (par transitivité) et donc que A est

irréductible. Le polynôme caractéristique de la matrice positive A est $X^n - X - 1$. Avec le théorème de Cayley-Hamilton, on a $A^n = A + I$ et $A^{n^2-n} = (A_0 + I)^{n-1}$ est strictement positive puisque A est irréductible, la matrice A est donc primitive. On peut en fait vérifier que le coefficient d'indice $(1, 1)$ de A^{n^2-2n+1} est nul et que $A^{n^2-2n+2} > 0$.

Exercice 4.3. On dit qu'une matrice rectangulaire $B \in \mathcal{M}_{m,n}(\mathbb{C})$ est non redondante si aucune de ses lignes ni aucune de ses colonnes n'est nulle. Une matrice non redondante $B \in \mathcal{M}_{m,n}(\mathbb{C})$ est dite décomposable s'il existe des matrices de permutation $P \in \mathcal{M}_n(\mathbb{C})$ et $Q \in \mathcal{M}_m(\mathbb{C})$ telles que PBQ soit de la forme $\begin{pmatrix} B' & 0 \\ 0 & B'' \end{pmatrix}$, où B' et B'' sont des matrices rectangulaires. Une matrice rectangulaire B est dite indécomposable si elle est non redondante et n'est pas décomposable.

1. Soient $B \in \mathcal{M}_{m,n}(\mathbb{C})$ et $C = \begin{pmatrix} 0 & B \\ {}^tB & 0 \end{pmatrix} \in \mathcal{M}_\ell(\mathbb{C})$, avec $\ell = m + n$.
Montrer que B est indécomposable si, et seulement si, C est irréductible.
2. Soit $B \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice à coefficients réels positifs ou nuls. Montrer que si B est indécomposable, $B {}^tB$ et ${}^tBB$ sont alors primitives.

Solution.

1. Soit $B \in \mathcal{M}_{m,n}(\mathbb{C})$ non redondante et décomposable. Il existe alors une partition (I, J) de $\{1, \dots, m\}$ (les numéros de lignes) et une partition (K, L) de $\{1, \dots, n\}$ (les numéros de colonnes) telles que :

$$\forall (i, j) \in I \times K \cup J \times L, b_{ij} = 0$$

On définit alors la partition (U, V) de $\{1, \dots, m + n\}$ par $U = I \cup (m + L)$, $V = J \cup (m + K)$, où on a posé $m + L = \{m + j \mid j \in L\}$ et pour $(i, j) \in U \times V$, on a les quatre possibilités suivantes :

- soit $i \geq m + 1$ et $j \geq m + 1$, dans ce cas on a $c_{ij} = 0$ par définition de la matrice $C = \begin{pmatrix} 0 & B \\ {}^tB & 0 \end{pmatrix}$ avec $B \in \mathcal{M}_{m,n}(\mathbb{C})$ et ${}^tB \in \mathcal{M}_{n,m}(\mathbb{C})$;
- soit $i \geq m + 1$ et $j \leq m$, dans ce cas en écrivant $i = m + i'$, c_{ij} est le coefficient de la ligne i' et de la colonne j de ${}^tB$, soit $c_{ij} = b_{ji'}$ avec $(j, i') \in J \times L$, ce qui donne $c_{ij} = b_{ji'} = 0$;
- soit $i \leq m$ et $j \leq m$, dans ce cas on a $c_{ij} = 0$ par définition de la matrice C ;
- soit $i \leq m$ et $j \geq m + 1$, dans ce cas en écrivant $j = m + j'$, c_{ij} est le coefficient de la ligne i et de la colonne j' de B , soit $c_{ij} = b_{ij'}$ avec $(i, j') \in I \times K$, ce qui donne $c_{ij} = b_{ij'} = 0$.

On a donc $c_{ij} = 0$ pour tout $(i, j) \in U \times V$ et avec la symétrie de la matrice C on a également $c_{ij} = 0$ pour tout $(i, j) \in V \times U$. Cette partition (U, V) de $\{1, \dots, m + n\}$ définit donc une matrice de permutation P telle que

$P^{-1}CP = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$, ce qui signifie que la matrice C est réductible. Réciproquement supposons la matrice C réductible, il existe alors une partition (U, V) de $\{1, \dots, m+n\}$ telle que $c_{ij} = 0$ pour tout $(i, j) \in U \times V \cup V \times U$. On définit alors les partitions (I, J) de $\{1, \dots, m\}$ et (K, L) de $\{1, \dots, n\}$ par :

$$\begin{cases} I = \{1, \dots, m\} \cap U, J = \{1, \dots, m\} \cap V \\ K = \{j \in \{1, \dots, n\} \mid m+j \in V\}, L = \{j \in \{1, \dots, n\} \mid m+j \in U\} \end{cases}$$

Pour $(i, j) \in I \times K$, on a $(i, m+j) \in U \times V$ et $c_{ij} = 0$ (coefficient de la ligne $i \leq m$ et de la colonne $m+j$ de C), c'est-à-dire $b_{ij} = 0$. De même si $(i, j) \in J \times L$, on a $(i, m+j) \in V \times U$ et $c_{ij} = 0$, soit $b_{ij} = 0$. En conclusion B est décomposable. On a donc ainsi montré que B est indécomposable si, et seulement si, C est irréductible.

2. Si $D = B {}^tB$ est réductible, étant symétrique cela équivaut à l'existence d'une partition (I, J) de $\{1, \dots, m\}$ telle que $d_{ij} = 0$ pour tout $(i, j) \in I \times J$, soit :

$$\forall (i, j) \in I \times J, \sum_{k=1}^n b_{ik}b_{jk} = 0$$

La matrice B étant positive, cela équivaut à :

$$\forall (i, j) \in I \times J, \forall k \in \{1, \dots, n\}, b_{ik}b_{jk} = 0$$

On définit alors une partition (K, L) de $\{1, \dots, n\}$ en posant :

$$\begin{cases} K = \{k \in \{1, \dots, n\} \mid \exists j \in J \mid b_{jk} \neq 0\} \\ L = \{k \in \{1, \dots, n\} \mid \exists i \in I \mid b_{ik} \neq 0\} \end{cases}$$

En effet, pour tout $k \in \{1, \dots, n\}$ il existe $i \in \{1, \dots, m\} = I \cup J$ tel que $b_{ik} \neq 0$ puisque B est supposée non redondante, on a donc $k \in K \cup L$. Si $k \in K$ il existe alors $j \in J$ tel que $b_{jk} \neq 0$ et $b_{ik}b_{jk} = 0$ pour tout $i \in I$ entraîne $k \notin L$. On a donc $K \cap L = \emptyset$, c'est-à-dire que (K, L) est bien une partition de $\{1, \dots, n\}$. Pour $(i, k) \in I \times K$, il existe $j \in J$ tel que $b_{jk} \neq 0$ et $b_{ik}b_{jk} = 0$ donne $b_{ik} = 0$. De même $(i, k) \in J \times L$ donne $b_{ik} = 0$. La matrice B est donc décomposable. On a donc ainsi montré que si B est indécomposable alors $B {}^tB$ est irréductible. De manière analogue on montre que ${}^tBB$ est irréductible. La matrice $B {}^tB$ étant symétrique réelle positive (au sens euclidien) est diagonalisable à valeurs propres réelles positives. Cette matrice étant positive et irréductible, son rayon spectral r est valeur propre avec un espace propre de dimension 1, cette valeur propre est donc simple (puisque $B {}^tB$ est diagonalisable) et $|\lambda| < r$ pour toute autre valeur propre de $B {}^tB$. Il en est de même de ${}^tBB$ puisque cette matrice a même polynôme caractéristique que $B {}^tB$.

Exercice 4.4. Montrer que $P = \left\{ x \in (\mathbb{R}^+)^n \mid \|x\|_1 = \sum_{i=1}^n x_i = 1 \right\}$ est un polyèdre convexe.

Solution. L'ensemble P est borné puisque contenu dans la boule unité fermée de $(\mathbb{R}^n, \|\cdot\|_1)$. En désignant par $(e_i^*)_{1 \leq i \leq n}$ la base duale de la base canonique de $\mathbb{R}^n$

($e_i^*(x) = x_i$, pour tout $x = \sum_{i=1}^n x_i e_i \in \mathbb{R}^n$) et par φ la forme linéaire définie par

$$\varphi = \sum_{i=1}^n e_i^*, \text{ on a :}$$

$$(x \in P) \Leftrightarrow \begin{cases} e_i^*(x) \geq 0 \quad (1 \leq i \leq n) \\ \varphi(x) \geq 1, \varphi(x) \leq 1 \end{cases}$$

ce qui prouve que P est un polyèdre convexe de $\mathbb{R}^n$.

Exercice 4.5. Montrer que si C est un convexe compact dans $\mathbb{R}^n$ euclidien, il est alors l'enveloppe convexe de sa frontière.

Solution. L'ensemble C étant fermé, on a $\text{Fr}(C) \subset C$ et $\text{Cv}(\text{Fr}(C))$ est contenu dans C puisque C est convexe. Soit x un élément de C . Pour toute droite D passant par x , $D \cap C$ est convexe comme intersection de convexes, fermé comme intersection de fermés et borné car contenu dans C qui est compact, c'est donc un convexe compact de D , c'est-à-dire un segment de D (qui peut être identifiée à $\mathbb{R}$ par le choix d'une origine). On a donc $x \in D \cap C = [y, z]$ avec y, z dans la frontière de C , ce qui entraîne $x \in \text{Cv}(\text{Fr}(C))$. On a donc bien $C = \text{Cv}(\text{Fr}(C))$.

Exercice 4.6. Montrer que les vecteurs de la base canonique de $\mathbb{R}^n$ sont les points extrémaux du convexe $P = \left\{ x \in (\mathbb{R}^+)^n \mid \|x\|_1 = \sum_{i=1}^n x_i = 1 \right\}$.

Solution. Si $e_i = (1-t)x + ty$ avec x, y dans P et $0 \leq t \leq 1$, on a alors :

$$(1-t)x_j + ty_j = \begin{cases} 0 & \text{si } j \neq i \\ 1 & \text{si } j = i \end{cases}$$

avec x_j, y_j positifs pour $1 \leq j \leq n$. Si $0 < t < 1$, on a alors $x_j = y_j = 0$ pour $j \neq i$

et $x_i = y_i = 1$ puisque $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j = 1$, c'est-à-dire que $x = y = e_i$. Chaque

vecteur e_i est donc extrême dans P . Réciproquement si a est un élément extrême de P et a n'est égal à aucun des e_i , ce vecteur a alors au moins deux composantes a_i et a_j strictement positives avec $1 \leq i < j \leq n$. Si $t = \min(a_i, a_j)$, on a alors $0 < t < 1$ et en posant $x = a + t(e_i - e_j)$ et $y = a + t(-e_i + e_j)$, on a :

$$\begin{cases} x = (a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_{j-1}, a_j - t, a_{j+1}, \dots, a_n) \in P \\ y = (a_1, \dots, a_{i-1}, a_i - t, a_{i+1}, \dots, a_{j-1}, a_j + t, a_{j+1}, \dots, a_n) \in P \end{cases}$$

avec $a = \frac{1}{2}(x + y)$, c'est-à-dire que a est le milieu du segment $[x, y] \subset P$ avec $a \neq x$, $a \neq y$, en contradiction avec a extrême. On a donc ainsi montré que les e_i sont les seuls points extrémaux de P .

Chapitre 5

Systèmes linéaires

$\mathbb{K}$ est le corps des réels ou des complexes, n est un entier supérieur ou égal à 2 et $(e_k)_{1 \leq k \leq n}$ désigne la base canonique de $\mathbb{K}^n$.

Pour i, j entiers compris entre 1 et n , on note E_{ij} la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) qui vaut 1. La famille $(E_{i,j})_{1 \leq i, j \leq n}$ est alors une base de $\mathcal{M}_n(\mathbb{K})$.

Une matrice est notée $A = ((a_{ij}))_{1 \leq i, j \leq n}$, où l'indice i est le numéro de la ligne et j celui de la colonne.

On note I_n la matrice identité d'ordre n .

5.1 Position des problèmes et notations

On s'intéresse dans ce chapitre à la résolution de systèmes linéaires, $Ax = b$, de n équations à n inconnues à coefficients dans $\mathbb{K}$ et au problème de l'inversion d'une matrice de $GL_n(\mathbb{K})$.

On distingue deux catégories de méthodes de résolution d'un système linéaire, les méthodes directes et les méthodes itératives.

Si on suppose une précision infinie, les méthodes directes conduisent à la solution exacte du système en un nombre fini d'étapes alors que les méthodes itératives donnent toujours une approximation de la solution, la solution exacte étant obtenue en un nombre infini d'étapes.

Les méthodes directes ont un caractère exact, mais pour les grands systèmes, la propagation des erreurs d'arrondis en diminue l'efficacité.

Les méthodes itératives sont bien adaptées au cas des matrices creuses (avec beaucoup de termes nuls), car ces méthodes ne transforment pas la matrice de départ, contrairement aux méthodes directes.

Pour l'aspect programmation on pourra consulter [?] ou [24].

Exemple 5.1 Une méthode directe, la méthode de Cramer : Si on note, pour $j = 1, 2, \dots, n$, A_j la matrice déduite de A en remplaçant par b sa colonne numéro j , la solution du système $Ax = b$ est alors donnée par les formules de Cramer :

$$x_j = \frac{\det(A_j)}{\det(A)} \quad (j = 1, 2, \dots, n)$$

En calculant un déterminant par la formule $\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) \prod_{j=1}^n a_{j,\sigma(j)}$, où $\mathfrak{S}_n$ désigne l'ensemble des permutations de $\{1, 2, \dots, n\}$ et $\text{sign}(\sigma)$ la signature de σ , cela nécessite $n!(n-1)$ multiplications et $(n!-1)$ additions, donc environ $nn!$ opérations élémentaires. Comme il y a $n+1$ déterminants à calculer puis n divisions à faire, on aura donc un total d'environ $n^2n!$ opérations élémentaires à effectuer, ce qui peut être beaucoup trop important pour de grandes valeurs de n .

Les méthodes directes sont plutôt adaptées au « petits système » ($n < 50$) et les méthodes itératives aux « grands systèmes ».

Les méthodes décrites dans ce chapitre sont utilisées pour la résolution des systèmes non linéaires par la méthode de Newton-Raphson et pour résoudre des problèmes d'interpolation et d'approximation.

Les méthodes itératives sont utilisées en relation avec les méthodes de différences finies et d'éléments finis pour la résolution de certaines équations aux dérivées partielles linéaires.

Si la matrice A et le vecteur b sont à coefficients complexes, en écrivant que $A = C + iD$, $b = \alpha + i\beta$ et $x = \xi + i\eta$, avec $C, D, \alpha, \beta, \xi, \eta$ à coefficients réels, la résolution du système de n équations à n inconnues à coefficients complexes $Ax = b$, se ramène à un système de $2n$ équations à $2n$ inconnues à coefficients réels $\begin{pmatrix} C & -D \\ D & C \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. On peut donc se contenter de ne considérer que les systèmes à coefficients réels.

Si la matrice A est inversible, les colonnes de A^{-1} sont alors les solutions des n systèmes linéaires :

$$Ax = e_j \quad (j = 1, 2, \dots, n)$$

En effet, en notant pour tout $j = 1, 2, \dots, n$, C_j la colonne numéro j de A^{-1} , on a $AA^{-1} = (AC_1, AC_2, \dots, AC_n) = (e_1, e_2, \dots, e_n) = I_n$, c'est-à-dire que le calcul de l'inverse d'une matrice se ramène à la résolution simultanée de n systèmes linéaires de même matrice. Le coût du calcul de l'inverse d'une matrice d'ordre n est donc n fois supérieur à celui de la résolution d'un système linéaire de n équations à n inconnues. Pour cette raison il est déconseillé, d'un point de vue numérique, d'inverser une matrice si cette opération ne s'avère pas nécessaire.

5.2 Problèmes numériques liés à la résolution des systèmes linéaires

La précision des ordinateurs n'étant pas infinie, un algorithme de résolution numérique d'un système linéaire ne donnera pas, en général, une solution exacte mais seulement une approximation de cette solution.

D'autre part il sera important de savoir évaluer le temps de calcul nécessaire à un tel algorithme.

Les erreurs numériques sont de trois types :

- les erreurs sur les données (pour des données expérimentales) ;
- les erreurs d'arrondis (calculs en virgule flottante) ;

— les erreurs de troncature.

Les erreurs d’arrondis vont s’accumuler au cours des calculs, il est donc important de connaître le nombre d’opérations élémentaires que nécessite un algorithme.

Un système $Ax = b$ est dit (théoriquement) dégénéré si $\det(A) = 0$, mais vérifier cette condition de manière exacte nécessite une précision infinie. Il faut donc définir un concept de dégénérescence numérique d’un système linéaire.

Un système sera dit numériquement dégénéré, si la valeur de son déterminant calculé en machine est non significative (c’est-à-dire trop petite ou trop grande).

Exemple 5.2 La matrice de Hilbert définie par $H_n = \left(\left(\frac{1}{i+j-1} \right) \right)_{1 \leq i, j \leq n}$ est symétrique définie positive, donc non dégénérée. En fait, on a $\det(H_n) = \frac{\Phi_{n-1}^4}{\Phi_{2n-1}}$, où $\Phi_n = \prod_{k=1}^n k!$ (exercice 5.2). Pour n grand, on a $\det(H_n)_{\text{machine}} = 0$ et le système est numériquement dégénéré, mais non dégénéré.

Des exemples de calculs de déterminants de matrices de Hilbert, en utilisant, d’une part la décomposition LR et d’autre part la formule de Hilbert sont donnés dans le tableau qui suit.

n	Déterminant Formule	Déterminant calculé avec LR
1	1.00000000000000E+00	1.00000000000000E+00
2	8.33333333333333E-02	8.33333333333333E-02
3	4.62962962962963E-04	4.62962962962961E-04
4	1.65343915343915E-07	1.65343915343930E-07
5	3.74929513251509E-12	3.74929513251290E-12
6	5.36729988735869E-18	5.36729988682088E-18
7	4.83580262392612E-25	4.83580261566227E-25
8	2.73705011379151E-33	2.73705023258196E-33
9	9.72023431192500E-43	9.72026335654786E-43
10	2.16417922643149E-53	2.16436997064480E-53
11	3.01909533444935E-65	3.02449714771783E-65
12	2.63778065125355E-78	2.70874882840569E-78
13	1.44289651879114E-92	Pivot trop petit dans LR
14	4.94031491459083E-108	Pivot trop petit dans LR
15	1.05854274306972E-124	Pivot trop petit dans LR
16	1.41913921144317E-142	Pivot trop petit dans LR
17	1.19027124270300E-161	Pivot trop petit dans LR
18	6.24486061412311E-182	Pivot trop petit dans LR
19	2.04934373331812E-203	Pivot trop petit dans LR
20	4.20617895662473E-226	Pivot trop petit dans LR
21	5.39898624190004E-250	Pivot trop petit dans LR

Dans la pratique, les coefficients de la matrice et du second membre ne sont connus que de façon approximative. Un problème important est donc de savoir si de petites variations sur ces coefficients peuvent entraîner de grosses variations

sur la solution. La notion de conditionnement d'une matrice (paragraphe 3.5) nous fournit un outil qui permet d'analyser ce problème.

Les matrices de Hilbert sont mal conditionnées (exercice 3.23).

5.3 Cas des matrices triangulaires

On suppose pour ce paragraphe que la matrice A est inversible, triangulaire supérieure et à coefficients réels ou complexes. Le système $Ax = b$ s'écrit alors :

$$a_{ii}x_i + a_{i,i+1}x_{i+1} + \cdots + a_{i,n}x_n = b_i \quad (1 \leq i \leq n)$$

avec $a_{ii} \neq 0$ pour tout $i = 1, \dots, n$ et la résolution de ce système se fait alors « en remontant » :

$$\begin{cases} x_n = \frac{b_n}{a_{nn}} \\ x_i = \frac{\left(b_i - \sum_{j=i+1}^n a_{ij}x_j\right)}{a_{ii}} \quad (i = n-1, \dots, 1) \end{cases}$$

Pour un système triangulaire inférieur, on procède de même « en descendant ».

Pour le calcul de l'inverse d'une matrice triangulaire on utilise le fait que les colonnes de A^{-1} sont les solutions des systèmes $Ax = e_i$ avec i compris entre 1 et n , ce qui donne, pour A triangulaire inférieure, pour les coefficients de $U = A^{-1}$:

$$\begin{cases} u_{ii} = \frac{1}{a_{ii}} \\ u_{ij} = 0 \quad (1 \leq i < j \leq n) \\ u_{ki} = -\frac{\sum_{j=i}^{k-1} a_{kj}u_{ji}}{a_{kk}} \quad (1 \leq i < k \leq n) \end{cases}$$

Si la matrice A est triangulaire supérieure la matrice ${}^tA$ est alors triangulaire inférieure et on peut alors écrire $A^{-1} = ({}^tA)^{-1}$.

5.4 Matrices de dilatation et de transvection. Opérations élémentaires

Définition 5.1. On appelle matrice de transvection toute matrice de la forme $T_{ij}(\lambda) = I_n + \lambda E_{ij}$, avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$.

Définition 5.2. On appelle matrice de dilatation toute matrice de la forme $D_i(\lambda) = I_n + (\lambda - 1)E_{ii}$, avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$.

Définition 5.3. On appelle *matrice élémentaire* une matrice de dilatation ou de transvection.

Lemme 5.1 Une matrice élémentaire est inversible d'inverse $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$ pour une matrice de transvection et $D_i(\lambda)^{-1} = D_i\left(\frac{1}{\lambda}\right)$ pour une matrice de dilatation.

Preuve. Pour $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$ on a $T_{ij}(\lambda)T_{ij}(-\lambda) = I_n - \lambda^2 E_{ij}^2$ avec $E_{ij}^2 = 0$. En effet pour tout entier k compris entre 1 et n on a :

$$E_{ij}e_k = \begin{cases} 0 & \text{si } k \neq j \\ e_i & \text{si } k = j \end{cases} \Rightarrow E_{ij}^2 e_k = \begin{cases} 0 & \text{si } k \neq j \\ E_{ij}e_i = 0 & \text{si } k = j \end{cases}$$

Le deuxième résultat est évident. □

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ on note L_i ($1 \leq i \leq n$) sa ligne numéro i (matrice à 1 ligne et n colonnes) et C_j ($1 \leq j \leq n$) sa colonne numéro j (matrice à n lignes et 1 colonne). On écrira $A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$ ou $A = (C_1 \quad \cdots \quad C_n)$.

On appelle matrice déduite de A par opération élémentaire sur les lignes de A toute matrice de la forme :

$$A_i(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \lambda L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$$

avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$, c'est-à-dire que la matrice $A_i(\lambda)$ est déduite de la matrice A en multipliant sa ligne numéro i par λ ou de la forme :

$$A_{ij}(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + \lambda L_j \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$$

avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$, c'est-à-dire que la matrice $A_{ij}(\lambda)$ est déduite de la matrice A en ajoutant à la ligne numéro i la ligne numéro j multipliée par λ .

On appelle matrice déduite de A par opération élémentaire sur les colonnes de A toute matrice de la forme :

$$A'_j(\lambda) = (C_1 \quad \cdots \quad C_{j-1} \quad \lambda C_j \quad C_{j+1} \quad \cdots \quad C_n)$$

avec $1 \leq j \leq n$ et $\lambda \in \mathbb{K}^*$, c'est-à-dire que la matrice $A'_j(\lambda)$ est déduite de la matrice A en multipliant sa colonne numéro j par λ ou de la forme :

$$A'_{ji}(\lambda) = (C_1 \quad \cdots \quad C_{j-1} \quad C_j + \lambda C_i \quad C_{j+1} \quad \cdots \quad C_n)$$

avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$, c'est-à-dire que la matrice $A'_{ji}(\lambda)$ est déduite de la matrice A en ajoutant à la colonne numéro j la colonne numéro i multipliée par λ .

Lemme 5.2 Avec les notations qui précèdent on a :

$$\begin{aligned} A_i(\lambda) &= D_i(\lambda) A, \quad A_{ij}(\lambda) = T_{ij}(\lambda) A \\ A'_j(\lambda) &= A D_j(\lambda), \quad A'_{ji}(\lambda) = A T_{ji}(\lambda) \end{aligned}$$

Preuve. Le coefficient d'indice (p, q) du produit de matrices $D_i(\lambda) A$ est obtenu en faisant le produit de la ligne p de $D_i(\lambda)$ par la colonne q de A , ce qui donne en notant $\alpha_{p,q}$ ce coefficient :

$$\alpha_{p,q} = \begin{cases} a_{p,q} & \text{si } 1 \leq p \neq i \leq n, 1 \leq q \leq n \\ \lambda a_{i,q} & \text{si } p = i, 1 \leq q \leq n \end{cases}$$

On a donc bien $A_i(\lambda) = D_i(\lambda) A$. Les autres égalités se montrent de façon analogue. $\square$

On sera amené à utiliser deux autres types d'opérations élémentaires que sont les permutations de lignes ou de colonnes. En fait ces opérations se déduisent des précédentes. Par exemple pour permuter les lignes i et j où $1 \leq i < j \leq n$ on effectue les opérations suivantes :

$$\begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} \mapsto \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + L_j \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix} \mapsto \begin{pmatrix} L_1 \\ \vdots \\ L_i + L_j \\ \vdots \\ -L_i \\ \vdots \\ L_n \end{pmatrix} \mapsto \begin{pmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ -L_i \\ \vdots \\ L_n \end{pmatrix} \mapsto \begin{pmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix}$$

ce qui revient à effectuer les produits $D_j(-1) T_{ij}(1) T_{ji}(-1) T_{ij}(1) A$. La matrice $D_j(-1) T_{ij}(1) T_{ji}(-1) T_{ij}(1)$ est la matrice de permutation qui a pour action de permuter les vecteurs e_i et e_j , elle s'écrit $P_{ij} = I_n - (E_{ii} + E_{jj}) + (E_{ij} + E_{ji})$. De même la permutation des colonnes i et j est obtenue en multipliant à droite la matrice A par $T_{ij}(1) T_{ji}(-1) T_{ij}(1) D_i(-1)$. Le produit de matrices $T_{ij}(1) T_{ji}(-1) T_{ij}(1) D_i(-1)$ est la matrice de permutation ${}^t P_{ij} = P_{ji}$.

L'ensemble des matrices de dilatation ou de transvection forme un système générateur du groupe multiplicatif $GL_n(\mathbb{K})$.

Théorème 5.1.

Pour toute matrice $A \in GL_n(\mathbb{K})$ il existe des matrices de transvection $P_1, \dots, P_r$ et $Q_1, \dots, Q_s$ telles que $A = P_1 \cdots P_r D_n(\det(A)) Q_1 \cdots Q_s$.

Preuve. On procède par récurrence sur $n \geq 1$. Pour $n = 1$ le résultat est évident. On le suppose vrai sur $GL_{n-1}(\mathbb{K})$.

On se ramène tout d'abord au cas où $a_{21} \neq 0$. Si $a_{21} = 0$, la matrice A étant inversible, il existe alors un indice $i \in \{1, 3, \dots, n\}$ tel que $a_{i1} \neq 0$ et la matrice $T_{2i}(1)A$ (déduite de A en ajoutant la ligne i à la ligne 2) est telle que son coefficient d'indice $(2, 1)$ est non nul. Une fois ramené à $a_{21} \neq 0$, on se ramène à $a_{11} = 1$ en remplaçant la première ligne L_1 par $L_1 + \lambda L_2$ (multiplication à gauche par $T_{12}(\lambda)$) où le scalaire λ est choisi tel que $a_{11} + \lambda a_{21} = 1$. Ensuite, pour tout $i \in \{2, 3, \dots, n\}$, en remplaçant la ligne L_i par $L_i - a_{i1} L_1$ (multiplication à gauche par $T_{i1}(-a_{i1})$), on annule le coefficient d'indice $(i, 1)$. On peut donc trouver des matrices de transvection $P_1, \dots, P_k$ telles que :

$$P_k \cdots P_1 A = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

De manière analogue, en multipliant à droite par des matrices de transvection, $Q_1, \dots, Q_m$, on obtient :

$$P_k \cdots P_1 A Q_1 \cdots Q_m = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & B & & \\ 0 & & & \end{pmatrix}$$

avec $\det(B) = \det(A)$, puisque $\det(P) = 1$ pour toute matrice de transvection P . On peut alors conclure en appliquant l'hypothèse de récurrence à la matrice $B \in GL_{n-1}(\mathbb{K})$. $\square$

En désignant par $GL_n^+(\mathbb{R})$ [resp. $GL_n^-(\mathbb{R})$] l'ensemble des matrices d'ordre n à coefficients réels de déterminant strictement positif [resp. négatif] on déduit du théorème précédent le résultat suivant.

Corollaire 5.1 : Les ensembles $GL_n^+(\mathbb{R})$ et $GL_n^-(\mathbb{R})$ sont connexes par arcs et ce sont les composantes connexes de $GL_n(\mathbb{R})$.

Preuve. Soit $A \in GL_n^+(\mathbb{R})$. Elle s'écrit $A = P_1 \cdots P_r D_n(\det(A)) Q_1 \cdots Q_s$, les P_k et Q_j étant des matrices de transvections. Pour toute matrice de transvection $T = T_{ij}(\lambda)$ et tout $t \in [0, 1]$, on note $T(t) = T_{ij}(t\lambda)$ et on définit l'application $\gamma : [0, 1] \rightarrow \mathcal{M}_n(\mathbb{R})$ par :

$$\forall t \in [0, 1], \gamma(t) = \prod_{k=1}^r P_k(t) \Delta(t) \prod_{k=1}^s Q_k(t)$$

où $\Delta(t) = D_n(t \det(A) + (1-t))$. Pour tout $t \in [0, 1]$, on a :

$$\det(\gamma(t)) = t \det(A) + (1-t) > 0$$

du fait que $A \in GL_n^+(\mathbb{R})$. De plus γ est continue avec $\gamma(0) = I_n$ et $\gamma(1) = A$. On a donc ainsi prouvé que $GL_n^+(\mathbb{R})$ est connexe par arcs.

Pour toute matrice $A \in GL_n^-(\mathbb{R})$ (par exemple $A = D_n(-1)$) l'application $M \mapsto AM$ réalise un homéomorphisme de $GL_n^+(\mathbb{R})$ sur $GL_n^-(\mathbb{R})$. On en déduit alors que $GL_n^-(\mathbb{R})$ est connexe par arcs.

Avec $GL_n(\mathbb{R}) = GL_n^+(\mathbb{R}) \cup GL_n^-(\mathbb{R})$, les ensembles $GL_n^+(\mathbb{R})$ et $GL_n^-(\mathbb{R})$ étant des ouverts connexes disjoints, on déduit que ce sont les composantes connexes de $GL_n(\mathbb{R})$. $\square$

Ce résultat permet de définir une orientation sur $\mathbb{R}^n$. On dit que deux bases $\mathcal{B}$ et $\mathcal{B}'$ de $\mathbb{R}^n$ définissent la même orientation si la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est dans $GL_n^+(\mathbb{R})$.

5.5 Méthode des pivots de Gauss

Effectuant des opérations élémentaires sur les lignes du système linéaire $Ax = b$ (permutations et combinaisons linéaires de lignes, c'est-à-dire des opérations qui ne vont pas changer l'ensemble de ses solutions), on le transforme en un système triangulaire supérieur $Rx = c$. Du fait qu'une permutation de lignes change le déterminant de signe et que d'ajouter un multiple d'une ligne à une autre ne change pas ce dernier, on a $\det(A) = \det(R) = \pm \prod_{i=1}^n r_{ii}$.

L'intérêt de cette méthode est surtout pédagogique, dans la pratique il est préférable d'utiliser la méthode LR (paragraphe 5.7) quand cette dernière peut s'appliquer, ou la méthode de Gauss-Jordan (paragraphe 5.10).

On note ici L_i la ligne numéro i du système linéaire $Ax = b$.

Etape 0 — On se ramène à un système tel que a_{11} non nul.

Si pour tout $i = 1, 2, \dots, n$, on a $a_{i1} = 0$, on a alors $\det(A) = 0$ et c'est terminé. Sinon, il existe $i > 1$ tel que a_{i1} soit non nul, et en permutant les lignes 1 et i (si $i = 1$, on ne fait rien), on se ramène à un système $A^{(1)}x = b^{(1)}$, avec $a_{11}^{(1)}$ non nul.

Le coefficient $a_{11}^{(1)}$ est le premier pivot.

On a alors $\det(A) = \pm \det(A^{(1)})$, avec le signe moins si, et seulement si, il y a une permutation des lignes 1 et i avec $i > 1$.

Etape 1 — Elimination de x_1 dans les équations $2, \dots, n$.

On effectue pour cela les transformations élémentaires suivantes :

$$L_i^{(1)} \mapsto L_i^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} L_1^{(1)} \quad (i = 2, 3, \dots, n)$$

et après une éventuelle permutation des lignes 2 et $i > 3$, on obtient le système $A^{(2)}x = b^{(2)}$, avec $a_{22}^{(2)}$ non nul où :

$$A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}$$

Le coefficient $a_{22}^{(2)}$ est le deuxième pivot.

Etape k — Elimination de x_k dans les équations $k+1, \dots, n$.

À la fin de l'étape $k-1$, on a obtenu le système $A^{(k)}x = b^{(k)}$, avec :

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}$$

et $a_{kk}^{(k)}$ non nul.

Le coefficient $a_{kk}^{(k)}$ est le pivot numéro k .

On effectue alors les transformations élémentaires suivantes :

$$L_i^{(k)} \mapsto L_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} L_k^{(k)} \quad (i = k+1, \dots, n)$$

puis une éventuelle permutation des lignes $k+1$ et $j > k+1$ pour se ramener à $a_{k+1,k+1}^{(k+1)}$ non nul.

Au bout de $n-1$ étapes, on est donc ramené à un système triangulaire supérieur $A^{(n)}x = b^{(n)}$.

De plus, on a $\det(A) = (-1)^p \det(A^{(n)})$, où p est le nombre de permutations qui ont été nécessaires pour avoir des pivots non nuls et $\det(A^{(n)})$ est le produit des pivots.

Pour éviter de faire une division par un nombre trop petit, dans le choix du pivot, on aura intérêt, à l'étape $k-1$, à permuter la ligne k avec la ligne $j \geq k$ telle que $|a_{jk}| = \max_{k \leq i \leq n} |a_{ik}|$, de manière à avoir le pivot le plus grand possible en valeur absolue. Si ce maximum est trop petit, le système est alors numériquement dégénéré.

On peut donner un ordre de grandeur du nombre d'opérations élémentaires que nécessite la méthode de Gauss.

Les opérations $L_i^{(k)} \mapsto L_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} L_k^{(k)}$ ($i = k+1, \dots, n$), à l'étape numéro k , demandent $(n-k)$ divisions, $(n-k)^2$ multiplications et $(n-k)^2$ additions, ce qui

donne un total de $\sum_{k=1}^{n-1} n - k = \frac{n(n-1)}{2}$ divisions et un nombre d'additions et de multiplications égal à $2 \sum_{k=1}^{n-1} (n-k)^2 = \frac{2n(n-1)(2n-1)}{6}$. Il nous faut donc $\frac{n(n-1)(4n+1)}{6} \sim \frac{2n^3}{3}$ opérations élémentaires pour aboutir à un système triangulaire supérieur. Puis la résolution de ce dernier système nécessite n divisions et $n(n-1)$ additions et multiplications, soit n^2 opérations élémentaires. En résumé, la résolution d'un système de n équations à n inconnues par la méthode de Gauss va nécessiter un nombre d'opérations qui est un $O(n^3)$.

5.6 Résolution des systèmes linéaires à coefficients entiers

Si la matrice A et le vecteur b sont à coefficients entiers relatifs, en modifiant la méthode de Gauss, on peut calculer la solution x de façon exacte comme vecteur à coefficients dans $\mathbb{Q}$.

On procède comme suit.

Etape 0 — Après une éventuelle permutation de lignes, on a obtenu le système $A^{(1)}x = b^{(1)}$ à coefficients dans $\mathbb{Z}$ avec un premier pivot non nul.

Etape 1 — Elimination de x_1 dans les équations $2, \dots, n$.

En effectuant la division de la ligne 1 par le premier pivot, on perd le caractère entier du système. Il est préférable de procéder comme suit.

- garder la ligne 1 ;
- $L_i^{(1)} \mapsto a_{i1}^{(1)} L_i^{(1)} - a_{i1}^{(1)} L_1^{(1)}$, pour $i = 2, \dots, n$.

Ce qui donne le système $A^{(2)}x = b^{(2)}$, avec un deuxième pivot non nul (au prix d'une éventuelle permutation).

Etape 2 — Elimination de x_2 dans les équations $3, \dots, n$.

Si on procède de la même façon pour cette étape, on constate que les coefficients de la matrice $A^{(3)}$ et du vecteur $b^{(3)}$ sont divisibles par le premier pivot. On peut donc s'autoriser une division par ce premier pivot pour cette étape 2, ce qui aura l'avantage de diminuer les coefficients obtenus.

Etape k — Elimination de x_k dans les équations $k+1, \dots, n$.

Cette étape est décrite par les formules $a_{ij}^{(k+1)} = \frac{(a_{kk}^{(k)} a_{ij}^{(k)} - a_{ik}^{(k)} a_{kj}^{(k)})}{a_{k-1,k-1}^{(k-1)}}$ pour

$i = k+1, \dots, n$ et $j = k+1, \dots, n+1$ (en prenant b comme colonne numéro $n+1$), ce qui donne au bout de $n-1$ étapes un système triangulaire supérieur à coefficients entiers qui peut se résoudre de façon exacte dans $\mathbb{Q}^n$.

Calcul du déterminant — Au signe près, on a $\det(A^{(2)}) = \left(a_{11}^{(1)}\right)^{n-1} \det(A)$ et $\det(A^{(3)}) = \left(\frac{a_{22}^{(2)}}{a_{11}^{(1)}}\right)^{n-2} \det(A^{(2)}) = \left(a_{22}^{(2)}\right)^{n-2} a_{11}^{(1)} \det(A^{(1)})$. De proche en

proche, on a $\det(A^{(n)}) = (-1)^p \left(\prod_{k=1}^n a_{kk}^{(k)} \right) \det(A)$, ce qui donne, tenant compte de l'expression de $\det(A^{(n)})$ comme produit des pivots, $\det(A) = (-1)^p a_{nn}^{(n)}$.

5.7 Décomposition LR ou méthode de Crout

Définition 5.4. Les sous-matrices principales d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ sont les matrices :

$$A_k = ((a_{ij}))_{1 \leq i, j \leq k} \quad (k = 1, \dots, n)$$

et les déterminants principaux sont les $\Delta_k = \det(A_k)$.

Dans le cas où tous les déterminants principaux de A sont non nuls, dans la méthode de Gauss, tous les pivots seront non nuls et il ne sera pas nécessaire de faire des permutations de lignes (on ne s'occupe pas de la taille des pivots). Réciproquement, si tous les pivots sont non nuls, tous les déterminants principaux sont alors non nuls.

Exemples 5.1

1. Si A est à diagonale strictement dominante, il en est alors de même de toutes les sous matrices principales de A et ces dernières sont inversibles. On en déduit que, pour A à diagonale strictement dominante, la méthode de Gauss ne nécessite pas de permutations de lignes.
2. Si A est symétrique définie positive, il en est de même de toutes les sous matrices principales de A . Donc, pour A symétrique définie positive, la méthode de Gauss ne nécessite pas de permutations de lignes.

La méthode des pivots de Gauss est basée sur les résultats suivants.

Lemme 5.3 Soit $A \in \mathcal{M}_n(\mathbb{K})$ de coefficient a_{11} non nul. Il existe des matrices de transvection $P_1, \dots, P_r$ telles que :

$$P_r \cdots P_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$$

Preuve. En reprenant les notations du paragraphe 5.5, le passage de la matrice A , de premier terme a_{11} non nul à la matrice unité $A^{(2)}$ se fait en multipliant à gauche la matrice A par les matrices de transvection :

$$P_i = T_{i1}(\lambda_{i1}) \quad (2 \leq i \leq n)$$

où on a noté $\lambda_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}}$, c'est-à-dire que :

$$P_{n-1} \cdots P_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$$

□

On a $\text{rg}(A) = 1 + \text{rg}(B^{(1)})$, en notant $B^{(1)} = \begin{pmatrix} a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \ddots & \vdots \\ a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$.

On a $P_{n-1} \cdots P_1 = F_1$ avec $F_1 = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \lambda_{21} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \lambda_{n-1,1} & 0 & \cdots & 1 & 0 \\ \lambda_{n1} & 0 & \cdots & 0 & 1 \end{pmatrix}$. La matrice

F_1 est une matrice de Frobenius.

Définition 5.5. On appelle matrice de Frobenius une matrice carrée d'ordre n qui ne diffère de l'identité que par une colonne (ou une ligne).

Théorème 5.2.

Toute matrice $A \in GL_n(\mathbb{K})$ peut être réduite à la forme triangulaire supérieur en la multipliant à gauche par des matrices de transvection ou de dilatation de la forme $D_i(-1)$.

Preuve. C'est la transcription matricielle de l'algorithme de Gauss.

On peut aussi raisonner par récurrence sur $n \geq 1$ en utilisant le lemme précédent.

En multipliant au besoin la matrice A par une matrice de permutation (produit d'une matrice de dilatation par trois matrices de transvection) on se ramène au cas où le coefficient a_{11} est non nul. On conclut alors facilement en utilisant le lemme 5.3 et l'hypothèse de récurrence. □

Dans l'hypothèse où tous les déterminants principaux de la matrice A sont non nuls (on n'effectue pas de permutations de lignes dans la méthode de Gauss), le résultat précédent s'exprime en disant qu'il existe des matrices de Frobenius

de la forme $F_k = \prod_{i=k+1}^n T_{ik}(\lambda_{ik})$ ($\lambda_{ik} = -\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$) avec les notations du paragraphe

5.5) telles que $F_{n-1} \cdots F_1 A = R$, où R est une matrice triangulaire supérieure. La matrice produit $F_{n-1} \cdots F_1$ est triangulaire inférieure à diagonale unité. En notant $L = (F_{n-1} \cdots F_1)^{-1}$ on a alors la décomposition $A = LR$ avec L triangulaire inférieure à diagonale unité et R triangulaire supérieure.

Théorème 5.3.

Une matrice inversible A possède une décomposition $A = LR$, avec L triangulaire inférieure à diagonale unité et R triangulaire supérieure si, et seulement si, tous les déterminants principaux de A sont non nuls. Une telle décomposition est unique et les coefficients diagonaux de R sont donnés par :

$$\begin{cases} r_{11} = a_{11} \\ r_{kk} = \frac{\det(A_k)}{\det(A_{k-1})} \quad (k = 2, \dots, n) \end{cases}$$

où les A_k désignent les sous matrices principales de A .

Preuve. Si tous les déterminants principaux de A sont non nuls on a vu que la matrice A admet une décomposition LR .

Pour $k = 1, \dots, n$, on peut décomposer les matrices A , L et R par blocs de la façon suivante :

$$A = \begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix}, \quad L = \begin{pmatrix} L_k & 0 \\ E_k & G_k \end{pmatrix}, \quad R = \begin{pmatrix} R_k & H_k \\ 0 & M_k \end{pmatrix}$$

où A_k , L_k et R_k désignent respectivement les sous matrices principales de A , L et R . Le produit $A = LR$ donne alors $A_k = L_k R_k$ et on a :

$$\det(A_k) = \det(L_k) \det(R_k) = \det(R_k) = \prod_{i=1}^k r_{ii}$$

L'unicité de la décomposition LR provient du fait que si $A = LR$ et $A = L'R'$ avec L, L' triangulaires inférieures de diagonale unité et R, R' triangulaires supérieures, alors $L'^{-1}L = R'R^{-1}$ est à la fois triangulaire inférieure et triangulaire supérieure de diagonale unité et c'est alors nécessairement l'identité.

Inversement si la matrice inversible A admet une décomposition LR , alors R est aussi inversible et la décomposition par bloc faite précédemment nous montre que tous les déterminants principaux de A sont non nuls. $\square$

Pour obtenir pratiquement de la décomposition LR on peut procéder par coefficients indéterminés, c'est-à-dire qu'on écrit $A = LR$, on effectue le produit et on identifie, ce qui donne en effectuant les calculs dans l'ordre indiqué :

$$\begin{cases} r_{1j} = a_{1j} \quad (j = 1, \dots, n) \\ L_{i1} = \frac{a_{i1}}{a_{11}} \quad (i = 2, \dots, n) \end{cases}$$

puis, pour $k = 2, \dots, n$:

$$r_{kj} = a_{kj} - \sum_{i=1}^{k-1} L_{ki} r_{ij} \quad (j = k, \dots, n)$$

et :

$$L_{ik} = \frac{\left(a_{ik} - \sum_{j=1}^{k-1} L_{ij} r_{jk} \right)}{r_{kk}} \quad (i = k+1, \dots, n)$$

avec $L_{ii} = 1$.

En utilisant la décomposition $A = LR$, le système $Ax = b$ équivaut aux deux systèmes triangulaires :

$$\begin{cases} Ly = b & (\text{triangulaire inférieur}) \\ Rx = y & (\text{triangulaire supérieur}) \end{cases}$$

La méthode obtenue est appelée méthode de Crout. Cette méthode est intéressante pour résoudre en parallèle plusieurs systèmes linéaires de même matrice A , car la décomposition LR ne fait pas intervenir le second membre.

En utilisant la décomposition $A = LR$, on a $A^{-1} = R^{-1}L^{-1}$ et il suffit alors d'utiliser les procédures d'inversion des matrices triangulaires du paragraphe 5.3.

5.8 Décomposition LD^tL des matrices symétriques réelles

On suppose dans ce paragraphe que la matrice A est inversible symétrique et qu'il n'y a pas de permutations dans la méthode de Gauss.

Dans la décomposition LR de A , on a $\det(R) = \det(A) \neq 0$, donc tous les termes diagonaux de R sont non nuls et on peut écrire R sous la forme $R = DR'$, où D est diagonale et R' est triangulaire supérieure à diagonale unité (il suffit de diviser chaque ligne de R par son terme diagonal). On a donc $A = LDR'$, puis en écrivant que ${}^tA = A$ et en utilisant le fait que la décomposition LR est unique, on déduit que $R' = {}^tL$. On a donc pour toute matrice symétrique A , dont tous les déterminants principaux sont non nuls, la décomposition unique $A = LD^tL$, la matrice L étant triangulaire inférieure et à diagonale unité et la matrice D diagonale.

Cette décomposition nous donne un moyen de calculer la signature de la matrice symétrique réelle A .

On rappelle qu'une matrice symétrique réelle a toutes ses valeurs propres réelles et que, pour A inversible, le couple d'entiers (p, q) formé du nombre p de valeurs propres strictement positives et du nombre $q = n - p$ de valeurs propres strictement négatives est uniquement déterminé par A . Ce couple d'entiers est la signature de A .

L'égalité $A = LD^tL$ avec L inversible se traduit en disant que les matrices A et D sont congruentes et le théorème de Sylvester nous dit que deux matrices congruentes ont même signature. On a donc $\text{sign}(A) = \text{sign}(D) = (p, n - p)$, où p est le nombre de termes strictement positifs de la diagonale de D .

La matrice A est définie positive si, et seulement si, tous les coefficients de D sont strictement positifs.

Comme pour la décomposition LR , on trouve les coefficients de L et D par identification, ce qui donne, $d_1 = a_{11}$ et pour $i = 2, \dots, n$:

$$\begin{cases} L_{ij} = \frac{\left(a_{ij} - \sum_{k=1}^{j-1} L_{ik}d_kL_{jk} \right)}{d_j} & (j = 1, \dots, i-1) \\ d_i = \left(a_{ii} - \sum_{k=1}^{i-1} L_{ik}^2d_k \right) \end{cases}$$

5.9 Décomposition de Cholesky des matrices symétriques réelles définies positives

L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique.

Dans le cas où la matrice A est symétrique définie positive, la décomposition LD^tL permet de montrer le résultat suivant.

Théorème 5.4.

Une matrice réelle A est symétrique définie positive si, et seulement si, il existe B triangulaire inférieure et inversible telle que $A = B^tB$. De plus une telle décomposition est unique si on impose la positivité des coefficients diagonaux de la matrice B .

Preuve. Si A est symétrique définie positive, dans la décomposition LD^tL , on a alors $d_i > 0$ pour tout $i = 1, \dots, n$. On peut donc écrire $D = D'^2$ et, en posant $B = LD'$, on a $A = B^tB$ avec B triangulaire inférieure, la diagonale de B étant formée des $\pm\sqrt{d_i}$. Réciproquement, si $A = B^tB$, la matrice A est alors symétrique et pour tout vecteur x non nul, on a $\langle Ax | x \rangle = \langle B^tBx | x \rangle = \langle {}^tBx | {}^tBx \rangle > 0$, puisque B est inversible. $\square$

Le calcul effectif des coefficients de B se fait par identification ce qui donne :

$$\begin{cases} b_{ij} = \frac{\left(a_{ij} - \sum_{k=1}^{j-1} b_{ik}b_{jk}\right)}{b_{jj}} & (j = 1, \dots, i-1) \\ b_{ii}^2 = a_{ii} - \sum_{k=1}^{i-1} b_{ik}^2 \end{cases} \quad (i = 1, \dots, n)$$

Le nombre d'opérations, dans la décomposition de Cholesky est un $O(n^3)$. En effet, il y a n racines carrées à calculer, $\sum_{i=1}^n i-1 = \frac{n(n-1)}{2}$ divisions,

$$\sum_{i=1}^n \left(i-1 + \sum_{j=1}^{i-1} j-1\right) = \sum_{i=1}^n \left(i-1 + \frac{(i-1)(i-2)}{2}\right) = \frac{1}{2} \sum_{i=1}^n i(i-1) \text{ additions}$$

et autant de multiplications. Ce qui donne un total $\frac{n(2n^2+3n+1)}{6}$ opérations.

Le déterminant de A se calcule avec $\det(A) = \prod_{i=1}^n b_{ii}^2$.

La résolution du système $Ax = e$ se ramène à la résolution de deux systèmes triangulaires.

Pour calculer l'inverse de A il suffit d'inverser la matrice triangulaire B .

5.10 Méthode d'élimination de Gauss-Jordan

La méthode de Gauss-Jordan est une variante de la méthode de Gauss. Elle consiste à se ramener à un système diagonal, à l'aide de transformations élémentaires.

À l'étape k , on ne se contente pas seulement d'éliminer le coefficient de x_k dans les lignes $k+1$ à n , mais on l'élimine aussi dans les lignes du dessus soit 1 à $k-1$. Ce qui donnera au bout de n étapes un système diagonal.

Description de la $k^{\text{ème}}$ étape du calcul ($1 \leq k \leq n$) À la $k^{\text{ème}}$ étape, on est au départ dans la situation suivante :

$$\left\{ \begin{array}{cccccc} x_1 + & & & +a_{1k}^{(k)} x_k & + \cdots & +a_{1n}^{(k)} x_n & = & b_1^{(k)} \\ & \ddots & & & & & & \ddots \\ & & x_{k-1} & +a_{k-1,k}^{(k)} x_k & + \cdots & +a_{k-1,n}^{(k)} x_n & = & b_{k-1}^{(k)} \\ & & & +a_{kk}^{(k)} x_k & + \cdots & +a_{kn}^{(k)} x_n & = & b_k^{(k)} \\ & & & +a_{k+1,k}^{(k)} x_k & + \cdots & +a_{k+1,n}^{(k)} x_n & = & b_{k+1}^{(k)} \\ & \ddots & & & & & & \ddots \\ & & & +a_{nk}^{(k)} x_k & + \cdots & +a_{nn}^{(k)} x_n & = & b_n^{(k)} \end{array} \right.$$

On peut supposer que $a_{kk}^{(k)}$ est le pivot maximum, sinon, il suffit de permuter avec une des lignes suivantes.

On commence par extraire x_k de la $k^{\text{ème}}$ équation, soit :

$$x_k = \frac{b_k^{(k)} - a_{k,k+1}^{(k)} x_{k+1} - \cdots - a_{kn}^{(k)} x_n}{a_{k,k}^{(k)}}$$

que l'on reporte ensuite dans les autres équations, ce qui donne :

$$\left\{ \begin{array}{cccccc} x_1 + & & & +a_{1k+1}^{(k+1)} x_{k+1} & + \cdots & +a_{1n}^{(k+1)} x_n & = & b_1^{(k+1)} \\ & \ddots & & & & & & \ddots \\ & & x_{k-1} & +a_{k-1,k+1}^{(k+1)} x_{k+1} & + \cdots & +a_{k-1,n}^{(k+1)} x_n & = & b_{k-1}^{(k+1)} \\ & & & x_k & +a_{k,k+1}^{(k+1)} x_{k+1} & + \cdots & +a_{kn}^{(k+1)} x_n & = & b_k^{(k+1)} \\ & & & & +a_{k+1,k+1}^{(k+1)} x_{k+1} & + \cdots & +a_{k+1,n}^{(k+1)} x_n & = & b_{k+1}^{(k+1)} \\ & \ddots & & & & & & \ddots \\ & & & & +a_{nk+1}^{(k+1)} x_{k+1} & + \cdots & +a_{nn}^{(k+1)} x_n & = & b_n^{(k+1)} \end{array} \right.$$

avec, pour $j = k+1, \dots, n$, $a_{kj}^{(k+1)} = \frac{a_{kj}^{(k)}}{a_{kk}^{(k)}}$, $b_k^{(k+1)} = \frac{b_k^{(k)}}{a_{kk}^{(k)}}$ et, pour $i \neq k$ dans $\{1, \dots, n\}$, j dans $\{k+1, \dots, n\}$:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - a_{ik}^{(k)} a_{kj}^{(k+1)}, \quad b_i^{(k+1)} = b_i^{(k)} - a_{ik}^{(k)} b_k^{(k+1)}$$

On obtient alors directement la solution, après la n -ième étape, soit $x_1 = b_1^{(n)}$, $\dots$, $x_n = b_n^{(n)}$.

5.11 Méthodes itératives de résolution des systèmes linéaires

Étant donné un système linéaire $Ax = b$ de n équations à n inconnues, la matrice A étant supposée inversible, on veut construire une suite $(x^{(k)})_{k \in \mathbb{N}}$ d'éléments de $\mathbb{R}^n$ qui va converger vers la solution x de ce système. Le calcul de chaque $x^{(k)}$ étant plus simple que la résolution directe du système.

Pour ce faire, on écrit la matrice A sous la forme $A = M - N$, où M est facilement inversible, et la résolution de $Ax = b$ est ramenée au problème de point fixe qui consiste à trouver x dans $\mathbb{R}^n$ solution de $x = M^{-1}Nx + M^{-1}b$. Pour résoudre ce problème, on utilise la méthode des approximations successives, c'est-à-dire qu'on considère la suite $(x^{(k)})_{k \in \mathbb{N}}$ définie par :

$$\begin{cases} x^{(0)} \in \mathbb{R}^n \\ x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b \end{cases} \quad (5.1)$$

Si cette suite converge, c'est nécessairement vers la solution de $Ax = b$.

Une méthode itérative de résolution du système linéaire $Ax = b$ est donc définie par le choix d'une décomposition $A = M - N$.

Définition 5.6. Si A est une matrice réelle inversible, on dit alors que la méthode itérative associée à la décomposition $A = M - N$ avec M inversible est convergente, si pour tout $x^{(0)} \in \mathbb{R}^n$, la suite $(x^{(k)})_{k \in \mathbb{N}}$ définie par (5.1) est convergente.

Les principaux problèmes posés par les méthodes itératives sont les suivants :

- à quelles conditions sur la matrice M , la suite définie par (5.1) converge-t-elle ?
- comment choisir M facilement inversible ?
- quelle est la rapidité de la convergence ?

Le théorème qui suit nous donne une condition nécessaire et suffisante de convergence de la suite définie par (5.1).

Théorème 5.5.

Soient $A = M - N$ une matrice inversible avec M inversible et $\rho(M^{-1}N)$ le rayon spectral de la matrice $M^{-1}N$. La méthode itérative associée à la décomposition $A = M - N$ converge si, et seulement si, $\rho(M^{-1}N) < 1$.

Preuve. On pose $B = M^{-1}N$ et on suppose que $\rho(B) < 1$. En écrivant que l'équation $Ax = b$ équivaut à $x = Bx + M^{-1}b$, on déduit que :

$$\forall k \in \mathbb{N}, x^{(k+1)} - x = B(x^{(k)} - x) = B^{k+1}(x^{(0)} - x)$$

Avec $\rho(B) < 1$ on déduit que $\lim_{k \rightarrow +\infty} B^k = 0$ et donc que la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers x , quelle que soit la valeur initiale $x^{(0)}$ (corollaire 3.17). Réciproquement, si la

méthode associée à la décomposition $A = M - N$ est convergente, alors pour tout vecteur $y^{(0)} \in \mathbb{R}^n$ la suite $(y^{(k)})_{k \in \mathbb{N}}$ définie par $y^{(k+1)} = By^{(k)}$ pour tout entier k est convergente vers le vecteur nul, ce qui équivaut à $\rho(B) < 1$ (corollaire 3.17). $\square$

S'il existe une norme matricielle telle que $\|B\| < 1$, la méthode itérative correspondante est alors convergente puisque $\rho(B) \leq \|B\|$.

Pour toute norme matricielle induite par une norme vectorielle, on a :

$$\|x^{(k)} - x\| = \|B^k(x^{(0)} - x)\| \leq \|B^k\| \|x^{(0)} - x\|$$

On peut donc utiliser $\|B^k\|$ pour avoir une idée de la vitesse de convergence d'une méthode itérative. De manière plus précise, on peut remarquer que :

$$\forall k \in \mathbb{N}, \|B^k\| = \sup_{y^{(0)} \in \mathbb{R}^n \setminus \{0\}} \frac{\|B^k y^{(0)}\|}{\|y^{(0)}\|} = \sup_{y^{(0)} \in \mathbb{R}^n \setminus \{0\}} \frac{\|y^{(k)}\|}{\|y^{(0)}\|}$$

la suite $(y^{(k)})_{k \in \mathbb{N}}$ étant définie par $y^{(k+1)} = By^{(k)}$, ce qui peut aussi s'écrire, en notant x la solution de $Ax = b$:

$$\forall k \in \mathbb{N}, \|B^k\| = \sup_{x^{(0)} \in \mathbb{R}^n \setminus \{x\}} \frac{\|x^{(k)} - x\|}{\|x^{(0)} - x\|}.$$

D'autre part, on sait que $\rho(B) = \lim_{k \rightarrow +\infty} \left(\|B^k\|^{\frac{1}{k}} \right)$ (corollaire 3.15).

Ces considérations nous amènent à donner les définitions suivantes.

Définition 5.7. Soit $A = M - N$ une matrice inversible avec M inversible telle que $\rho(M^{-1}N) < 1$. La quantité $R_\infty(M^{-1}N) = -\ln(\rho(M^{-1}N))$ est appelée le *taux asymptotique de convergence* de la méthode itérative associée à la décomposition $A = M - N$. Pour toute norme matricielle $\|\cdot\|$ induite par une norme vectorielle et pour tout entier k strictement positif tel que $\|B^k\| < 1$, la quantité $R_k(M^{-1}N) = -\frac{1}{k} \ln \left(\|(M^{-1}N)^k\| \right)$ est appelée le *taux moyen de convergence* pour k itérations et pour la norme $\|\cdot\|$ de la méthode itérative associée à la décomposition $A = M - N$.

On a bien sûr, $R_\infty(M^{-1}N) = \lim_{k \rightarrow +\infty} R_k(M^{-1}N)$.

La méthode itérative associée à la décomposition $A = M - N$ sera d'autant plus performante que $R_\infty(M^{-1}N)$ est grand.

5.12 Méthode de Jacobi

Pour $A = ((a_{ij}))_{1 \leq i, j \leq n}$, on choisit $M = D$ matrice diagonale définie par $d_{ii} = a_{ii}$ pour tout $i = 1, \dots, n$, en supposant que tous les coefficients diagonaux de A sont non nuls. L'algorithme de construction des $x^{(k)}$ est alors le suivant :

$$x_i^{(k+1)} = -\frac{1}{a_{ii}} \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} x_j^{(k)} + b_i \quad (i = 1, \dots, n)$$

Le calcul des composantes de $x^{(k+1)}$ nécessite de garder en mémoire le vecteur $x^{(k)}$. Une itération va donc immobiliser $2n$ cases mémoires.

On peut décider d'arrêter les itérations à un rang *MaxIter* donné ou lorsque $\|x^{(k+1)} - x^{(k)}\| < \varepsilon \|x^{(k+1)}\|$ (ε étant une précision donnée).

Comme valeur initiale, on peut prendre $x^{(0)} = \left(\left(\frac{b_i}{a_{ii}} \right) \right)_{1 \leq i \leq n}$.

Dans le cas particulier des matrices à diagonale strictement dominante, on a le résultat suivant.

Théorème 5.6.

Si la matrice A est à diagonale strictement dominante, la méthode de Jacobi est alors convergente.

Preuve. Une matrice à diagonale strictement dominante a tous ses termes diagonaux non nuls, on peut donc utiliser la méthode de Jacobi. Pour $J = D^{-1}N$, avec $A = D - N$, on a $\|J\|_\infty = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1$, donc la méthode de Jacobi est convergente. $\square$

5.13 Méthode de Gauss-Seidel

On choisit pour matrice M le triangle inférieur de A , c'est-à-dire que M est définie par :

$$\begin{cases} m_{ij} = 0, & \text{pour } 1 \leq i < j \leq n \\ m_{ij} = a_{ij}, & \text{pour } 1 \leq j \leq i \leq n \end{cases}$$

En notant $A = M - N$, le vecteur $x^{(k+1)}$ est solution du système triangulaire inférieur $Mx^{(k+1)} = Nx^{(k)} + b$. D'où l'algorithme de Gauss-Seidel :

$$a_{ii}x_i^{(k+1)} = - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} + b_i \quad (i = 1, \dots, n)$$

Cet algorithme est en fait une amélioration de l'algorithme de Jacobi. En effet, dans le calcul de $x_i^{(k+1)}$, on utilise les composantes 1 à $i-1$ de $x^{(k+1)}$ (alors que dans la méthode de Jacobi ce sont celles de $x^{(k)}$ qui sont utilisées) et les composantes $i+1$ à n de $x^{(k)}$ (comme dans la méthode de Jacobi). Cet algorithme sera donc en général plus performant que celui de Jacobi.

Pour une itération on garde seulement n termes en mémoire.

Les théorèmes qui suivent nous donnent des conditions suffisantes de convergence de la méthode de Gauss-Seidel.

Théorème 5.7.

Si la matrice A est à diagonale strictement dominante, la méthode de Gauss-Seidel est alors convergente.

Preuve. Une matrice à diagonale strictement dominante a tous ses termes diagonaux non nuls, on peut donc utiliser la méthode de Gauss-Seidel. Soient A à diagonale strictement dominante, $G = M^{-1}N$, $\lambda \in \mathbb{C}$ une valeur propre de G et $x \in \mathbb{C}^n$ un vecteur propre non nul associé. On a alors $Nx = \lambda Mx$, c'est-à-dire :

$$\begin{cases} -\sum_{j=i+1}^n a_{ij}x_j = \lambda \sum_{j=1}^i a_{ij}x_j & (i = 1, \dots, n-1) \\ 0 = \lambda \sum_{j=1}^n a_{nj}x_j \end{cases}$$

Si on suppose que $|\lambda| \geq 1$, en prenant i tel que $\|x\|_\infty = |x_i|$, on déduit que $|\lambda| |a_{ii}| \leq \sum_{j \neq i} |a_{ij}| |\lambda|$, ce qui contredit le fait que la matrice A est à diagonale strictement dominante. Le rayon spectral de G est donc strictement inférieur à 1 et la méthode de Gauss-Seidel converge. $\square$

Théorème 5.8.

Si la matrice A est symétrique définie positive, la méthode de Gauss-Seidel est alors convergente.

Preuve. Une matrice symétrique définie positive a tous ses termes diagonaux strictement positifs, on peut donc utiliser la méthode de Gauss-Seidel. Soient A symétrique définie positive, $\lambda \in \mathbb{C}$ une valeur propre de $G = M^{-1}N$ et $x \in \mathbb{C}^n$ un vecteur propre non nul associé. En écrivant $A = D + E + F$, où D est la diagonale de A , E le triangle inférieur strict, F le triangle supérieur strict et en notant $\langle \cdot | \cdot \rangle$ le produit scalaire hermitien canonique de $\mathbb{C}^n$, on a :

$$\langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Ex \rangle + \langle x | Fx \rangle \quad (5.2)$$

Puis avec $Fx = -\lambda(Dx + Ex)$ ($M = D + E$, $N = -F$), on déduit que :

$$\langle x | Ax \rangle = (1 - \bar{\lambda}) (\langle x | Dx \rangle + \langle x | Ex \rangle)$$

donc $\lambda \neq 1$ (puisque $\langle x | Ax \rangle > 0$) et :

$$\frac{1}{1 - \bar{\lambda}} \langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Ex \rangle \quad (5.3)$$

Par conjugaison complexe de (5.2), en considérant que A et D sont hermitiennes et que ${}^t\bar{E} = F$ on déduit que :

$$\frac{1}{1 - \lambda} \langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Fx \rangle \quad (5.4)$$

En faisant (5.3) + (5.4) - (5.2), on en conclut que $\frac{1 - |\lambda|^2}{|1 - \lambda|^2} \langle x | Ax \rangle = \langle x | Dx \rangle$.

Avec $\langle x | Ax \rangle > 0$ et $\langle x | Dx \rangle > 0$ (A et D sont définies positives) on déduit que $|\lambda| < 1$. Le rayon spectral de G est donc strictement inférieur à 1 et la méthode de Gauss-Seidel converge. $\square$

De manière générale, quand les méthodes de Jacobi et de Gauss-Seidel sont convergentes il vaut mieux choisir celle de Gauss-Seidel. Mais il se peut que la méthode de Gauss-Seidel diverge alors que celle de Jacobi converge comme le

montre l'exemple de la matrice $A = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$. Pour la méthode Jacobi

toutes les valeurs propres de J sont nulles et pour la méthode de Gauss-Seidel, on

a $G = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & -2 & 4 \end{pmatrix}$ avec les valeurs propres 0, 0.35 et 5.64.

Dans le cas des matrices tridiagonales les méthodes de Jacobi et Gauss-Seidel convergent ou divergent simultanément et dans le cas de la convergence c'est la méthode de Gauss-Seidel qui est la plus rapide (exercice 5.8).

5.14 Méthode de relaxation

On pose $A = D + E + F$, où D est la partie diagonale de A , E sa partie triangulaire inférieure stricte et F sa partie triangulaire supérieure stricte. On suppose toujours que la matrice A est inversible et que tous ses coefficients diagonaux sont non nuls.

En vue d'accélérer la convergence de la méthode de Gauss-Seidel, on introduit un paramètre dans la matrice $G = -(D + E)^{-1} F$. Précisément, on pose :

$$M_\omega = \frac{1}{\omega} D + E, \quad A = M_\omega - N_\omega, \quad L_\omega = M_\omega^{-1} N_\omega$$

où ω est un paramètre réel non nul à préciser. On a alors :

$$L_\omega = (D + \omega E)^{-1} ((1 - \omega) D - \omega F)$$

Le vecteur $x^{(k+1)}$ est alors la solution du système triangulaire :

$$(D + \omega E) x^{(k+1)} = ((1 - \omega) D - \omega F) x^{(k)} + \omega b$$

Ce qui donne l'algorithme de calcul :

$$x_i^{(k+1)} = (1 - \omega) x_i^{(k)} + \omega \hat{x}_i^{(k+1)} \quad (i = 1, \dots, n)$$

où $\hat{x}_i^{(k+1)}$ est donné par les formules de Gauss-Seidel en fonction des composantes 1 à $i - 1$ de $x^{(k+1)}$ et des composantes $i + 1$ à n de $x^{(k)}$, soit :

$$a_{ii} \hat{x}_i^{(k+1)} = - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i \quad (i = 1, \dots, n)$$

Ce qui peut encore s'écrire :

$$x^{(k+1)} = x^{(k)} - \omega \xi^{(k+1)} \quad (k \geq 0)$$

où $\xi^{(k+1)}$ est le vecteur résidu défini par :

$$\xi_i^{(k+1)} = \frac{1}{a_{ii}} \left(\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i}^n a_{ij} x_j^{(k)} - b_i \right) \quad (i = 1, \dots, n)$$

Comme test de convergence on peut prendre $\|\xi^{(k)}\| < \varepsilon$, où $\varepsilon > 0$ est une précision donnée, une norme matricielle étant choisie.

Une condition nécessaire sur ω pour que la méthode converge est donnée par le résultat suivant.

Théorème 5.9.

Avec les notations de ce paragraphe, la méthode de relaxation correspondante ne peut converger que si $\omega \in]0, 2[$.

Preuve. Si $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de L_ω , on a alors :

$$\prod_{i=1}^n \lambda_i = \det(L_\omega) = \frac{\det((1-\omega)D - \omega F)}{\det(D + \omega E)} = (1-\omega)^n$$

et $\rho(L_\omega) \geq |\omega - 1|$. □

Dans le cas où la matrice A est symétrique définie positive, la condition nécessaire du théorème précédent est aussi suffisante.

Théorème 5.10.

Soit A une matrice symétrique définie positive. Avec les notations de ce paragraphe, la méthode de relaxation correspondante est convergente si, et seulement si, $\omega \in]0, 2[$.

Preuve. Si la matrice A est symétrique définie positive, on a alors $a_{ii} > 0$ pour tout i compris entre 1 et n . On peut donc définir la suite $(x^{(k)})_{k \in \mathbb{N}}$ de la méthode de relaxation. Avec le théorème précédent, il nous suffit de montrer que la méthode converge pour $\omega \in]0, 2[$.

En reprenant les notations du début de ce paragraphe, on a $A = M_\omega - N_\omega$, avec $M_\omega = \frac{1}{\omega}D + E$ et $N_\omega = \left(\frac{1}{\omega} - 1\right)D - F$.

On note $\langle \cdot | \cdot \rangle$ le produit scalaire hermitien canonique de $\mathbb{C}^n$.

Soit $\lambda \in \mathbb{C}$ une valeur propre de $L_\omega = M_\omega^{-1}N_\omega$ et $x \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé. On a $N_\omega x = \lambda M_\omega x$ et :

$$\langle x | Ax \rangle = \langle x | M_\omega x \rangle - \langle x | N_\omega x \rangle = (1 - \bar{\lambda}) \langle x | M_\omega x \rangle$$

La matrice A étant symétrique définie positive, on en déduit que $\lambda \neq 1$ et :

$$\frac{1}{1 - \bar{\lambda}} \langle x | Ax \rangle = \frac{1}{\omega} \langle x | Dx \rangle + \langle x | Ex \rangle \quad (5.5)$$

Par conjugaison complexe, en considérant que A est hermitienne et que ${}^t E = F$ on déduit que :

$$\frac{1}{1 - \lambda} \langle x | Ax \rangle = \frac{1}{\omega} \langle x | Dx \rangle + \langle x | Fx \rangle \quad (5.6)$$

Enfin, avec $A = D + E + F$, on peut aussi écrire :

$$\langle x | Ax \rangle = \langle x | Dx \rangle + \langle x | Ex \rangle + \langle x | Fx \rangle \quad (5.7)$$

En faisant (5.5) + (5.6) - (5.7), on en conclut que :

$$\frac{(1 - |\lambda|^2)}{|1 - \lambda|^2} \langle x | Ax \rangle = \left(\frac{2}{\omega} - 1 \right) \langle x | Dx \rangle$$

En tenant compte de la positivité de A et D et du fait que $\omega \in]0, 2[$, on déduit que $|\lambda| < 1$. On a donc ainsi montré que $\rho(L_\omega) < 1$ et que la méthode de relaxation associée converge. $\square$

Pour $\omega = 1$, on retrouve le fait que la méthode de Gauss-Seidel converge pour les matrices symétriques définies positives.

Dans le cas des matrices tridiagonales symétriques définies positives on peut faire une étude plus précise.

Lemme 5.4 *Pour toute matrice $M = ((m_{ij}))_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$ et pour tout scalaire t la matrice $M(t) = ((m_{ij}(t)))_{1 \leq i, j \leq n}$ définie par :*

$$m_{ij}(t) = t^{i-j} m_{ij} \quad (1 \leq i, j \leq n)$$

est semblable à M .

Preuve. Il suffit de remarquer que $M(t) = P(t)MP(t)^{-1}$, où :

$$P(t) = \begin{pmatrix} t & 0 & \cdots & 0 \\ 0 & t^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t^n \end{pmatrix} \in GL_n(\mathbb{K})$$

pour t non nul. $\square$

On suppose que A est tridiagonale symétrique définie positive. Les coefficients diagonaux de A sont alors tous non nuls et on définit les matrices $J = -D^{-1}(E + F)$ (méthode de Jacobi) et $L_\omega = \left(\frac{1}{\omega} D + E \right)^{-1} \left(\left(\frac{1}{\omega} - 1 \right) D - F \right)$ (méthode de relaxation).

On désigne par P_J le polynôme caractéristique de J et par P_{L_ω} celui de L_ω .

Lemme 5.5 *Soit A tridiagonale symétrique définie positive. Avec les notations de ce paragraphe, pour tout $\omega \in \mathbb{R}^*$ et tout $\lambda \in \mathbb{C}^*$ on a :*

$$P_{L_\omega}(\lambda^2) = \omega^n \lambda^n P_J \left(\frac{\lambda^2 + \omega - 1}{\omega \lambda} \right)$$

Preuve. On a $P_{L_\omega}(\lambda) = \frac{1}{\det \left(\left(\frac{1}{\omega} D + E \right) \right)} \det \left(\left(\frac{1}{\omega} - 1 \right) D - F - \lambda \left(\frac{1}{\omega} D + E \right) \right)$,

soit $P_{L_\omega}(\lambda^2) = \frac{\omega^n}{\det(D)} \det(H_{\omega, \lambda})$, en notant $H_{\omega, \lambda}$ la matrice tridiagonale définie

par $H_{\omega,\lambda} = \frac{1-\omega-\lambda^2}{\omega}D - \lambda^2E - F$. En utilisant les notations du lemme 5.4, on peut écrire, pour $\lambda \in \mathbb{C}^*$, que $\det(H_{\omega,\lambda}) = \det\left(H_{\omega,\lambda}\left(\frac{1}{\lambda}\right)\right)$ avec :

$$H_{\omega,\lambda}\left(\frac{1}{\lambda}\right) = \lambda\left(\frac{1-\omega-\lambda^2}{\omega\lambda}D - E - F\right)$$

ce qui nous donne $P_{L_\omega}(\lambda^2) = \omega^n \lambda^n P_J\left(\frac{\lambda^2 + \omega - 1}{\omega\lambda}\right)$, en tenant compte de $P_J(\mu) = \frac{(-1)^n}{\det(D)} \det(E + F + \mu D)$. $\square$

Lemme 5.6 Soit A tridiagonale symétrique définie positive. Avec les notations de ce paragraphe, on a $P_J(\lambda) = (-1)^n \lambda^q \prod_{k=1}^p (\lambda^2 - \mu_k^2)$ avec $0 < \mu_1 \leq \dots \leq \mu_p < 1$ dans $\mathbb{R}$, $2p + q = n$ et :

$$P_{L_\omega}(\lambda) = (-1)^n (\lambda + \omega - 1)^q \prod_{k=1}^p \left((\lambda + \omega - 1)^2 - \omega^2 \mu_k^2 \lambda \right)$$

C'est-à-dire que les valeurs propres de L_ω sont $1 - \omega$ d'ordre supérieur ou égal à q et les $\lambda_1(\mu_k, \omega)$, $\lambda_2(\mu_k, \omega)$ pour $k = 1, \dots, p$, où les $\lambda_i(\mu_k, \omega)$ ($i = 1, 2$) sont les racines de l'équation :

$$(\lambda + \omega - 1)^2 - \omega^2 \mu_k^2 \lambda = 0 \quad (5.8)$$

Preuve. Pour la forme du polynôme caractéristique de J , voir l'exercice 5.8. Il reste à montrer que toutes les valeurs propres de J sont réelles. Si $\mu \in \mathbb{C}$ est une valeur propre de J et $x \in \mathbb{C}^n \setminus \{0\}$ est un vecteur propre associé, on a alors $(1 - \mu)Dx = Ax$ et $(1 - \mu)\langle Dx | x \rangle = \langle Ax | x \rangle$. Les matrices A et D étant symétriques définies positives, on déduit que μ est réelle et $\mu < 1$. Comme $-\mu$ est aussi valeur propre de J on déduit que $|\mu| < 1$. En utilisant le lemme 5.5, on a pour $\lambda \in \mathbb{C}^*$:

$$\begin{aligned} P_{L_\omega}(\lambda^2) &= \omega^n \lambda^n (-1)^n \left(\frac{\lambda^2 + \omega - 1}{\omega\lambda} \right)^q \prod_{k=1}^p \left(\left(\frac{\lambda^2 + \omega - 1}{\omega\lambda} \right)^2 - \mu_k^2 \right) \\ &= (-1)^n (\lambda^2 + \omega - 1)^q \prod_{k=1}^p \left((\lambda^2 + \omega - 1)^2 - \mu_k^2 \omega^2 \lambda^2 \right) \end{aligned}$$

ce qui équivaut à $P_{L_\omega}(\lambda) = (-1)^n (\lambda + \omega - 1)^q \prod_{k=1}^p \left((\lambda + \omega - 1)^2 - \omega^2 \mu_k^2 \lambda \right)$. $\square$

Lemme 5.7 Pour $\mu \in]0, 1[$, $\omega \in]0, 2[$, on considère l'équation en λ :

$$(\lambda + \omega - 1)^2 - \omega^2 \mu^2 \lambda = 0 \quad (5.9)$$

et on pose $\omega_0(\mu) = \frac{2}{1 + \sqrt{1 - \mu^2}}$. Pour $\omega \in]0, \omega_0(\mu)[$ l'équation (5.9) admet deux racines réelles distinctes :

$$\lambda_1(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) - \omega \mu \sqrt{\omega^2 \mu^2 - 4\omega + 4}}{2}$$

$$\lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) + \omega \mu \sqrt{\omega^2 \mu^2 - 4\omega + 4}}{2} = \frac{(\omega - 1)^2}{\lambda_1(\mu, \omega)}$$

pour $\omega = \omega_0(\mu)$ elle admet une racine double :

$$\lambda_1(\mu, \omega) = \lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2}{2} - (\omega - 1) = \omega_0(\mu) - 1$$

et pour $\omega \in]\omega_0(\mu), 2[$ elle admet deux racines complexes conjuguées :

$$\lambda_1(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) - i\omega \mu \sqrt{-\omega^2 \mu^2 + 4\omega - 4}}{2}$$

$$\lambda_2(\mu, \omega) = \frac{\omega^2 \mu^2 - 2(\omega - 1) + i\omega \mu \sqrt{-\omega^2 \mu^2 + 4\omega - 4}}{2} = \frac{(\omega - 1)^2}{\lambda_1(\mu, \omega)}$$

avec $|\lambda_1(\mu, \omega)| = |\lambda_2(\mu, \omega)| = \omega - 1$.

Preuve. À $\mu \in]0, 1[$ fixé, le discriminant $\Delta(\mu, \omega) = \omega^2 \mu^2 (\omega^2 \mu^2 - 4\omega + 4)$ de (5.9) a une seule racine dans $]0, 2[$, $\omega_0(\mu) = \frac{2}{1 + \sqrt{1 - \mu^2}}$. Le calcul des racines de (5.9) est alors immédiat. $\square$

Théorème 5.11. Young-Varga

Si A est symétrique définie positive et tridiagonale, alors :

1. les méthodes de Jacobi et de relaxation pour $\omega \in]0, 2[$ sont convergentes ;
2. en posant $r = \rho(J)$ et $\omega_0(r) = \frac{2}{1 + \sqrt{1 - r^2}}$, on a :

$$\rho(L_\omega) = \begin{cases} \frac{\omega^2 r^2 - 2(\omega - 1) + \omega r \sqrt{\omega^2 r^2 - 4\omega + 4}}{2} & \text{si } 0 < \omega < \omega_0(r) \\ \omega - 1 & \text{si } \omega_0(r) \leq \omega < 2 \end{cases}$$

C'est-à-dire que la fonction $\omega \mapsto \rho(L_\omega)$ a l'allure indiquée par la figure 5.1 et que la valeur optimale du paramètre de relaxation est $\omega_0(r)$ avec $\rho(L_{\omega_0(r)}) = \omega_0(r) - 1$.

Preuve.

1. La matrice A est symétrique définie positive, on sait alors que $r = \rho(J) \in [0, 1[$ et la méthode de Jacobi converge. Pour $\omega \in]0, 2[$ la méthode de relaxation converge d'après le théorème 5.10.

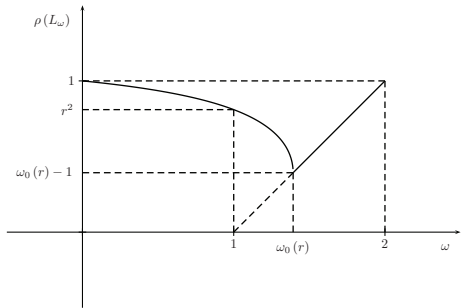


FIGURE 5.1 –

2. Si $r = 0$, toutes les valeurs propres de J sont alors nulles et $1 - \omega$ est la seule valeur propre de L_ω d'après le lemme 5.6. Et dans ce cas on a $\rho(L_\omega) = |1 - \omega|$ (figure 5.2).

Si $r \in]0, 1[$, d'après les lemmes 5.6 et 5.7 on a pour $\omega \in]\omega_0(\mu), 2[\subset]1, 2[$, $\rho(L_\omega) = |1 - \omega| = 1 - \omega$. Si $\omega \in]0, \omega_0(\mu)[$, en remarquant que la fonction $\mu \mapsto \lambda_1(\mu, \omega)$ est décroissante et $\mu \mapsto \lambda_2(\mu, \omega)$ croissante sur $\mathbb{R}^+$, on déduit avec les notations du lemme 5.6 que les valeurs propres de L_ω vérifient :

$$\lambda_1(\mu_p, \omega) \leq \dots \leq \lambda_1(\mu_1, \omega) < 1 - \omega < \lambda_2(\mu_1, \omega) \leq \dots \leq \lambda_2(\mu_p, \omega)$$

et $\rho(L_\omega) = \lambda_2(\mu_p, \omega) = \lambda_2(r, \omega)$.

□

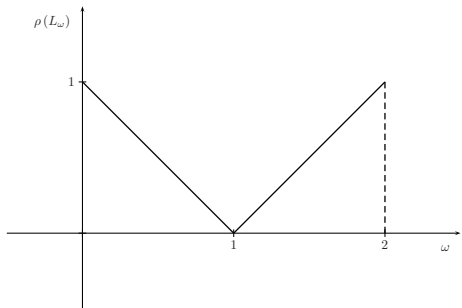


FIGURE 5.2 –

Du point de vue numérique il est préférable de surestimer la valeur optimale de ω du fait que la pente de la demi-tangente à droite en $\omega_0(r)$ est égale à 1 alors qu'elle infinie à gauche.

Si on sait calculer la valeur optimale de ω , ce choix n'est pas judicieux dès la première itération. Le procédé d'accélération de convergence de Tchebychev

consiste à changer de paramètre ω à chaque itération de la manière suivante. Pour première valeur de ω , on prend $\omega_1 = 1$, puis à l'étape suivante on prend $\omega_2 = \frac{1}{1 - \frac{\rho(J)^2}{2}}$ et aux étapes suivantes on prendra $\omega_{k+1} = \frac{1}{1 - \frac{\rho(J)^2 \omega_k}{4}}$. La suite $(\omega_k)_{k \in \mathbb{N}}$ converge en fait vers $\omega_0(r)$ et le procédé de Tchebychev va diminuer, en général, le nombre d'itérations.

Lorsque la dimension du système est très grande, on aura intérêt à partitionner la matrice A en blocs, ce découpage étant adapté à la forme particulière de A . On a donc :

$$A = A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix}$$

où A_{ij} est une matrice à n_i lignes et n_j colonnes, avec $\sum_{i=1}^n n_i = n$. Les matrices A_{ii} étant supposées inversibles. Le second membre b et l'inconnue x sont alors partitionnés de manière analogue, soit :

$$b = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{pmatrix}, \quad x = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{pmatrix}$$

Le système à résoudre peut alors s'écrire par blocs :

$$\sum_{j=1}^m A_{ij} X_j = B_i \quad (i = 1, \dots, m)$$

Et les formules pour la méthode de relaxation s'écrivent :

$$X^{(k+1)} = X^{(k)} - \omega \Xi^{(k+1)} \quad (k \geq 0)$$

où $\Xi^{(k+1)}$ est le vecteur résidu défini par :

$$\Xi^{(k+1)} = A_{ii}^{-1} \left(\sum_{j=1}^{i-1} A_{ij} X_j^{(k+1)} + \sum_{j=i}^n A_{ij} X_j^{(k)} - B_i \right) \quad (i = 1, \dots, m)$$

Cette méthode est la méthode de relaxation par blocs et est intéressante si les A_{ii} sont facilement inversibles.

La méthode de relaxation est utilisée pour la résolution d'une équation aux dérivées partielles de type elliptique par la méthode des différences finies. On aura à résoudre des systèmes tridiagonaux par blocs, les matrices blocs de la diagonale étant tridiagonales et les autres diagonales.

5.15 Méthodes de descente et de gradient

L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique. On note $\langle \cdot | \cdot \rangle$ le produit scalaire euclidien, $\|\cdot\|_2$ la norme associée et l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ est muni de la norme matricielle induite par $\|\cdot\|_2$.

On désigne par $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{R}^n$.

On rappelle qu'une fonction φ définie sur un ouvert non vide $\mathcal{O}$ de $\mathbb{R}^n$ et à valeurs réelles est dite différentiable en $x \in \mathcal{O}$ s'il existe une forme linéaire L sur $\mathbb{R}^n$ telle que $\varphi(x+h) = \varphi(x) + L(h) + o(\|h\|)$ pour tout h dans un voisinage de 0. Le vecteur gradient de φ en x est le vecteur noté $d\varphi(x)$ défini par :

$$\forall h \in \mathbb{R}^n, L(h) = \langle d\varphi(x) | h \rangle$$

Dans la base canonique de $\mathbb{R}^n$ ce vecteur s'écrit $d\varphi(x) = \sum_{k=1}^n \frac{\partial \varphi}{\partial x_k}(x) e_k$.

On dit que la fonction φ est différentiable sur $\mathcal{O}$ si elle est différentiable en tout point de $\mathcal{O}$.

On dit la fonction φ admet un minimum [resp. maximum] local en $x_0 \in \mathcal{O}$ s'il existe un voisinage ouvert $\mathcal{V}$ de x_0 dans $\mathcal{O}$ tel que :

$$\forall x \in \mathcal{V}, \varphi(x) \geq \varphi(x_0) \text{ [resp. } \varphi(x) \leq \varphi(x_0) \text{]}$$

On dit la fonction φ admet un extremum local en $x_0 \in \mathcal{O}$, si elle admet un minimum ou un maximum local en x_0 .

On rappelle enfin un résultat donnant une condition nécessaire pour qu'une fonction différentiable en un point d'un ouvert de $\mathbb{R}^n$ admette un extremum local en ce point.

Théorème 5.12.

Soient $\mathcal{O}$ un ouvert non vide de $\mathbb{R}^n$, φ une fonction définie sur $\mathcal{O}$ à valeurs réelles et différentiable en un point $x_0 \in \mathcal{O}$. Si φ admet un extremum local en x_0 , le vecteur gradient $d\varphi(x_0)$ est alors nul.

Preuve. On suppose que φ admet un maximum local en x_0 . L'ensemble $\mathcal{O}$ étant ouvert, pour tout vecteur $h \in \mathbb{R}^n$ il existe un réel α strictement positif tel que la fonction $f : t \mapsto \varphi(x_0 + th)$ soit définie sur $]-\alpha, \alpha[$. Cette fonction est dérivable en 0 avec $f'(0) = \langle d\varphi(x_0) | h \rangle$. En écrivant que :

$$\begin{cases} f'(0) = \lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{\varphi(x_0 + th) - \varphi(x_0)}{t} \leq 0 \\ f'(0) = \lim_{\substack{t \rightarrow 0 \\ t < 0}} \frac{\varphi(x_0 + th) - \varphi(x_0)}{t} \geq 0 \end{cases}$$

on déduit que $f'(0) = 0$. On a donc ainsi montré que $\langle d\varphi(x_0) | h \rangle = 0$ pour tout $h \in \mathbb{R}^n$, ce qui équivaut à $d\varphi(x_0) = 0$. $\square$

Définition 5.8. On appelle fonctionnelle quadratique sur $\mathbb{R}^n$ toute fonction φ définie sur $\mathbb{R}^n$ à valeurs réelles telle que :

$$\forall x \in \mathbb{R}^n, \varphi(x) = \frac{1}{2} \langle Ax \mid x \rangle - \langle b \mid x \rangle$$

où A est une matrice symétrique réelle d'ordre n et b un vecteur de $\mathbb{R}^n$.

Pour la suite de ce paragraphe, on se donne une matrice symétrique réelle définie positive d'ordre n , un vecteur non nul b dans $\mathbb{R}^n$ et on note φ la fonctionnelle quadratique associée.

Pour tout réel λ on désigne par $\mathcal{H}_\lambda$ l'hyper-quadrrique d'équation $\varphi(x) = \lambda$.

La matrice A étant inversible, le système linéaire $Ax = b$ admet une unique solution que nous noterons u . On a écarté le cas $b = 0$ qui donne la solution triviale $u = 0$.

Du fait que la matrice A est symétrique, on déduit que pour tout vecteur $\delta \in \mathbb{R}^n$ et tout réel t , on a :

$$\varphi(u + t\delta) = \frac{1}{2} \langle A\delta \mid \delta \rangle t^2 + \langle Au - b \mid \delta \rangle t + \varphi(u)$$

et avec $Au = b$, $\varphi(u) = -\frac{1}{2} \langle Au \mid u \rangle$, on obtient :

$$\varphi(u + t\delta) = \frac{1}{2} \langle A\delta \mid \delta \rangle t^2 - \frac{1}{2} \langle Au \mid u \rangle \quad (5.10)$$

Le vecteur u étant non nul (on a pris b non nul), on a $\langle Au \mid u \rangle > 0$ (la matrice A est définie positive) et pour δ non nul le vecteur $u + t\delta$ appartient à l'hyper-quadrrique $\mathcal{H}_\lambda$ si, et seulement si, t est solution de l'équation polynomiale $\langle A\delta \mid \delta \rangle t^2 = 2\lambda + \langle Au \mid u \rangle$. Pour λ positif ou nul on a deux solutions réelles,

$$t = \pm \sqrt{\frac{2\lambda + \langle Au \mid u \rangle}{\langle A\delta \mid \delta \rangle}}.$$

Le vecteur δ étant quelconque dans $\mathbb{R}^n \setminus \{0\}$, on déduit que pour tout réel λ positif ou nul l'ensemble $\mathcal{H}_\lambda$ est un hyper-ellipsoïde de centre u .

Lemme 5.8 La fonctionnelle quadratique φ associée à la matrice A et au vecteur b est différentiable sur $\mathbb{R}^n$ avec :

$$\forall x \in \mathbb{R}^n, d\varphi(x) = Ax - b$$

Preuve. En utilisant la symétrie de la matrice A , on a pour tous vecteurs x, h dans $\mathbb{R}^n$:

$$\varphi(x + h) = \varphi(x) + \langle Ax - b \mid h \rangle + \frac{1}{2} \langle Ah \mid h \rangle$$

et avec $|\langle Ah \mid h \rangle| \leq \|A\|_2 \|h\|_2^2$, on déduit le résultat. $\square$

Le vecteur gradient $d\varphi(x) = Ax - b$ est aussi appelé vecteur résidu et souvent noté $r(x)$.

Pour tout réel λ positif ou nul le vecteur $d\varphi(x)$ est dirigé suivant la normale à l'hyper-ellipsoïde $\mathcal{H}_\lambda$ passant par le centre u .

Théorème 5.13.

La solution u du système linéaire $Ax = b$ est le vecteur qui réalise le minimum global de la fonctionnelle quadratique φ .

Preuve. En utilisant (5.10) avec $t = 1$ et $\delta = x - u$, on a pour tout vecteur x dans $\mathbb{R}^n$, $\varphi(x) = \frac{1}{2} \langle A(x - u) | x - u \rangle - \frac{1}{2} \langle Au | u \rangle$. La matrice A étant définie positive, on en déduit que :

$$\forall x \in \mathbb{R}^n \setminus \{0\}, \varphi(x) > -\frac{1}{2} \langle Au | u \rangle = \varphi(u)$$

ce qui entraîne que φ atteint son minimum en u et que ce minimum est unique sur $\mathbb{R}^n$. Réciproquement si φ admet un extremum global en $x \in \mathbb{R}^n$ alors $d\varphi(x) = 0$, c'est-à-dire que x est solution de $Ax = b$ et nécessairement $x = u$. Cet extremum est un minimum et il est unique. $\square$

La résolution du système linéaire $Ax = b$ est donc équivalente à la résolution du problème d'optimisation sans contraintes qui consiste à déterminer le vecteur u de $\mathbb{R}^n$ solution de $\varphi(u) = \inf_{x \in \mathbb{R}^n} \varphi(x)$. Le principe des méthodes de descente pour résoudre ce type de problème consiste à construire une suite de vecteurs $(x_k)_{k \in \mathbb{N}}$ destinée à converger vers la solution cherchée u telle que le passage de x_k à x_{k+1} se fait en résolvant un problème de minimisation plus simple à une variable. À chaque itération on se donne une direction de descente $\delta_k \in \mathbb{R}^n \setminus \{0\}$ et on cherche à minimiser la restriction de φ à la droite passant par x_k et dirigée par δ_k , c'est-à-dire qu'on veut minimiser la fonction d'une variable réelle $t \mapsto \varphi_k(t) = \varphi(x_k + t\delta_k)$. En remarquant (grâce à la symétrie de la matrice A) que :

$$\varphi_k(t) = \frac{1}{2} \langle A\delta_k | \delta_k \rangle t^2 + \langle d\varphi(x_k) | \delta_k \rangle t + \varphi(x_k)$$

est un polynôme de degré 2 à coefficient dominant strictement positif (la matrice A est définie positive et le vecteur δ_k est non nul), on calcule facilement la valeur de t qui réalise le minimum de la fonction φ_k sur $\mathbb{R}$. Cette valeur est donnée par $t_k = -\frac{\langle d\varphi(x_k) | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle}$. En notant $r_k = d\varphi(x_k) = Ax_k - b = A(x_k - u)$ le vecteur

résidu d'ordre k , le vecteur x_{k+1} s'écrit $x_{k+1} = x_k - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \delta_k$.

Lemme 5.9 *Avec les notations qui précèdent, le vecteur r_{k+1} est orthogonal au vecteur δ_k .*

Preuve. On a $r_{k+1} = Ax_{k+1} - b = Ax_k - b + t_k A\delta_k = r_k + t_k A\delta_k$ et :

$$\langle r_{k+1} | \delta_k \rangle = \langle r_k | \delta_k \rangle - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \langle A\delta_k | \delta_k \rangle = 0$$

$\square$

Pour $n = 2$, le vecteur $r_{k+1} = d\varphi(x_{k+1})$ est orthogonal à l'ellipse $\mathcal{H}_{\varphi(x_{k+1})}$ et δ_k orthogonal à r_{k+1} est donc tangent à cette ellipse.

C'est le choix des vecteurs de descente qui va définir une méthode.

En prenant $\delta_k = e_1$, on obtient, en notant $x_j^{(k)}$ les composantes du vecteur x_k dans la base canonique de $\mathbb{R}^n$, $x_{k+1} = x_k - \frac{1}{a_{11}} \left(\sum_{j=1}^n a_{1j} x_j^{(k)} - b_1 \right) e_1$, ce qui donne les formules :

$$\begin{cases} a_{11}x_1^{(k+1)} = -\sum_{j=2}^n a_{1j}x_j^{(k)} + b_1 \\ x_j^{(k+1)} = x_j^{(k)} \quad (j = 2, \dots, n) \end{cases}$$

En remplaçant le vecteur x_k par le vecteur, encore noté x_k , de composantes $x_1^{(k+1)}, x_2^{(k)}, \dots, x_n^{(k)}$ et en prenant $\delta_k = e_2$, on obtient pour la deuxième composante de x_{k+1} :

$$a_{22}x_2^{(k+1)} = -a_{21}x_1^{(k+1)} - \sum_{j=3}^n a_{2j}x_j^{(k)} + b_2$$

les autres composantes de x_k étant inchangées. En continuant ainsi avec $e_3, \dots, e_n$ on retrouve les formules de Gauss-Seidel (paragraphe 5.13). On sait que pour A symétrique définie positive cette méthode est convergente.

Le théorème qui suit nous indique comment choisir des directions de descente qui assurent la convergence de la méthode.

On note $0 < \lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres de la matrice A rangées dans l'ordre croissant.

Lemme 5.10 Avec les notations qui précèdent, pour tout vecteur x dans $\mathbb{R}^n$, on a $\langle Ax \mid x \rangle \geq \lambda_1 \|x\|_2^2$.

Preuve. La matrice A , symétrique, se diagonalise dans une base orthonormée $(f_k)_{1 \leq k \leq n}$ avec $Af_k = \lambda_k f_k$ pour tout entier k compris entre 1 et n . Dans cette base on a $x = \sum_{k=1}^n \alpha_k f_k$ et $Ax = \sum_{k=1}^n \alpha_k \lambda_k f_k$, de sorte que :

$$\langle Ax \mid x \rangle = \sum_{k=1}^n \lambda_k \alpha_k^2 \geq \lambda_1 \sum_{k=1}^n \alpha_k^2 = \lambda_1 \|x\|_2^2$$

□

Théorème 5.14.

Soit $(x_k)_{k \in \mathbb{N}}$ la suite de vecteurs définie par :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, x_{k+1} = x_k - \frac{\langle r_k \mid \delta_k \rangle}{\langle A\delta_k \mid \delta_k \rangle} \delta_k \end{cases}$$

où $r_k = Ax_k - b$ est le vecteur résidu d'ordre k et δ_k un vecteur unitaire dans $\mathbb{R}^n$. S'il existe un réel α strictement positif tel que :

$$\forall k \in \mathbb{N}, \langle r_k | \delta_k \rangle \geq \alpha \|r_k\|_2 \quad (5.11)$$

la suite $(x_k)_{k \in \mathbb{N}}$ converge alors vers la solution du système linéaire $Ax = b$.

Preuve. Si il existe un entier naturel k_0 tel que r_{k_0} soit nul, la suite $(x_k)_{k \in \mathbb{N}}$ est alors stationnaire sur u à partir du rang k_0 et c'est terminé. On suppose donc r_k non nul pour tout entier naturel k .

On remarque tout d'abord que les inégalités (5.11) entraînent $0 < \alpha \leq 1$. En effet avec l'inégalité de Cauchy-Schwarz on a $\langle r_k | \delta_k \rangle \leq \|\delta_k\|_2 \|r_k\|_2 = \|r_k\|_2$ ce qui entraîne $\alpha \leq 1$ puisque r_k est non nul. Avec les inégalités :

$$\forall k \in \mathbb{N}, \|x_k - u\|_2^2 \leq \frac{1}{\lambda_1} \langle A(x_k - u) | x_k - u \rangle = \frac{1}{\lambda_1} e_k$$

il suffit de montrer que la suite $(e_k)_{k \in \mathbb{N}}$ converge vers 0 sur $\mathbb{R}$.

Pour tout entier naturel k on a $e_k = \langle Ax_k - b | x_k - u \rangle = \langle r_k | x_k - u \rangle$ et avec $\langle r_{k+1} | \delta_k \rangle = 0$, on déduit que :

$$e_{k+1} = \left\langle r_{k+1} | x_k - u - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \delta_k \right\rangle = \langle r_{k+1} | x_k - u \rangle$$

Avec $r_{k+1} = r_k - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} A\delta_k$ (démonstration du lemme 5.9) on déduit que :

$$e_{k+1} = e_k - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \langle A\delta_k | x_k - u \rangle$$

Enfin avec la symétrie de la matrice A et $Au = b$ on peut écrire que :

$$\langle A\delta_k | x_k - u \rangle = \langle \delta_k | Ax_k - b \rangle = \langle \delta_k | r_k \rangle$$

$$\text{et } e_{k+1} = e_k - \frac{\langle r_k | \delta_k \rangle^2}{\langle A\delta_k | \delta_k \rangle}.$$

Avec l'hypothèse r_k non nul pour tout entier naturel k on déduit que e_k est non nul pour tout entier naturel k et l'égalité précédente s'écrit :

$$e_{k+1} = e_k \left(1 - \frac{1}{e_k} \frac{\langle r_k | \delta_k \rangle^2}{\langle A\delta_k | \delta_k \rangle} \right) \quad (5.12)$$

Avec les inégalités :

$$\begin{cases} \langle r_k | \delta_k \rangle \geq \alpha \|r_k\|_2 \\ \langle A\delta_k | \delta_k \rangle \leq \|A\|_2 \|\delta_k\|_2^2 = \|A\|_2 = \lambda_n \\ e_k = \langle r_k | A^{-1}r_k \rangle \leq \|A^{-1}\|_2 \|r_k\|_2^2 = \frac{1}{\lambda_1} \|r_k\|_2^2 \end{cases}$$

on déduit que $e_{k+1} \leq \left(1 - \frac{\lambda_1}{\lambda_n} \alpha^2\right) e_k$, ce qui s'écrit $e_{k+1} \leq (1 - \text{cond}_2(A) \alpha^2) e_k$. Avec α et $\text{cond}_2(A)$ dans l'intervalle $]0, 1]$ on déduit que :

$$\forall k \in \mathbb{N}, 0 < e_k \leq \beta^k e_0$$

avec $\beta = 1 - \text{cond}_2(A) \alpha^2 \in [0, 1[$. Il en résulte que $\lim_{k \rightarrow +\infty} e_k = 0$ et $\lim_{k \rightarrow +\infty} x_k = u$. $\square$

En désignant par θ_k la mesure dans $[0, \pi]$ de l'angle des vecteurs r_k et δ_k , la condition (5.11) s'écrit :

$$\forall k \in \mathbb{N}, \cos(\theta_k) \geq \alpha$$

avec $\alpha \in]0, 1]$. En particulier on a $\theta_k \in \left[0, \frac{\pi}{2}\right[$ et les vecteurs r_k et δ_k ne sont pas orthogonaux.

Le cas $\theta_k = 0$ est réalisé pour $\delta_k = \frac{1}{\|r_k\|_2} r_k$ (et donc $\alpha = 1$) en supposant toujours r_k non nul. La méthode obtenue pour ces choix de vecteurs de descente est la méthode de plus profonde descente ou méthode du gradient ($r_k = d\varphi(x_k)$) à paramètre optimal.

Le théorème précédent nous dit que cette méthode est convergente et nous donne une majoration de l'erreur, ce qui est résumé par le résultat suivant.

Théorème 5.15.

Soient A une matrice réelle symétrique définie positive d'ordre n , b un vecteur non nul dans $\mathbb{R}^n$ et $(x_k)_{k \in \mathbb{N}}$ la suite de vecteurs définie par :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, x_{k+1} = \begin{cases} x_k & \text{si } r_k = Ax_k - b = 0 \\ x_k - \frac{\|r_k\|_2^2}{\langle Ar_k | r_k \rangle} r_k & \text{si } r_k \neq 0 \end{cases} \end{cases}$$

La suite $(x_k)_{k \in \mathbb{N}}$ converge vers la solution u du système linéaire $Ax = b$ et une majoration de l'erreur est donnée par :

$$\forall k \in \mathbb{N}, \|x_k - u\|_2 \leq \gamma (1 - \text{cond}_2(A))^{\frac{k}{2}}$$

où γ est une constante réelle.

La méthode du gradient conjugué de Hestenes et Stiefel consiste à construire la suite $(x_k)_{k \in \mathbb{N}}$ de la manière suivante

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, x_{k+1} = \begin{cases} x_k & \text{si } r_k = Ax_k - b = 0 \\ x_k - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \delta_k & \text{si } r_k \neq 0 \end{cases} \end{cases}$$

avec les vecteurs de descente δ_k , pour $k \geq 1$ (quand ils existent), choisis dans le plan engendré par les vecteurs orthogonaux r_k et δ_{k-1} de la forme $\delta_k = r_k + t_k \delta_{k-1}$,

le coefficient t_k étant déterminé de sorte que le coefficient $1 - \frac{1}{e_k} \frac{\langle r_k | \delta_k \rangle^2}{\langle A\delta_k | \delta_k \rangle}$ qui intervient dans la formule (5.12) soit minimum.

Si à l'étape $k \geq 1$ le vecteur r_k est non nul, le vecteur $e_k = \langle r_k | x_k - u \rangle$ est alors déjà déterminé et avec l'orthogonalité des vecteurs r_k et δ_{k-1} (lemme 5.9) on a $\langle r_k | \delta_k \rangle = \langle r_k | r_k + t_k \delta_{k-1} \rangle = \|r_k\|_2^2$ qui est également déterminé. Il s'agit donc de trouver $t_k \in \mathbb{R}$ qui minimise la fonction polynomiale de degré 2 :

$$P_k(t) = \langle A\delta_k | \delta_k \rangle = \langle A\delta_{k-1} | \delta_{k-1} \rangle t^2 + 2 \langle A\delta_{k-1} | r_k \rangle t + \langle Ar_k | r_k \rangle$$

Cette fonction atteint son minimum en $t_k = -\frac{\langle A\delta_{k-1} | r_k \rangle}{\langle A\delta_{k-1} | \delta_{k-1} \rangle}$.

On préfère écrire que le réel t_k est déterminé par $\langle A\delta_{k-1} | r_k + t_k \delta_{k-1} \rangle = 0$, ce qui équivaut à $\langle A\delta_{k-1} | \delta_k \rangle = 0$.

La matrice A étant symétrique définie positive, l'application $(x, y) \mapsto \langle Ax | y \rangle$ définit un produit scalaire sur $\mathbb{R}^n$ et la condition $\langle A\delta_{k-1} | \delta_k \rangle = 0$ signifie que les vecteurs δ_k et δ_{k-1} sont orthogonaux pour ce produit scalaire.

Définition 5.9. Si A est une matrice réelle d'ordre n symétrique définie positive, on dit alors que deux vecteurs x, y dans $\mathbb{R}^n$ sont A -conjugués si $\langle Ax | y \rangle = 0$.

En définitive le vecteur de descente est défini par $\delta_k = r_k - \frac{\langle A\delta_{k-1} | r_k \rangle}{\langle A\delta_{k-1} | \delta_{k-1} \rangle} \delta_{k-1}$ et on a $\delta_k \neq 0$ avec δ_k et δ_{k-1} qui sont A -conjugués.

Lemme 5.11 Avec les hypothèses et notations qui précèdent on a :

$$\langle r_{k+1} | \delta_{k-1} \rangle = 0, \langle r_{k+1} | r_k \rangle = 0, t_k = \frac{\|r_k\|_2^2}{\|r_{k-1}\|_2^2}$$

Preuve. On a, en utilisant l'expression de r_{k+1} obtenue dans la démonstration du lemme 5.9 :

$$\langle r_{k+1} | \delta_{k-1} \rangle = \langle r_k | \delta_{k-1} \rangle - \frac{\langle r_k | \delta_k \rangle}{\langle A\delta_k | \delta_k \rangle} \langle A\delta_k | \delta_{k-1} \rangle$$

avec $\langle r_k | \delta_{k-1} \rangle = 0$ et $\langle A\delta_k | \delta_{k-1} \rangle = \langle A\delta_{k-1} | \delta_k \rangle = 0$, ce qui donne bien $\langle r_{k+1} | \delta_{k-1} \rangle = 0$. Le vecteur r_{k+1} est orthogonal aux vecteurs δ_{k-1} et δ_k , il est donc orthogonal au vecteur r_k qui appartient au plan engendré par les vecteurs δ_{k-1} et δ_k .

Avec $A\delta_{k-1} = \frac{\langle A\delta_{k-1} | \delta_{k-1} \rangle}{\langle r_{k-1} | \delta_{k-1} \rangle} (r_k - r_{k-1})$, $\langle r_k | r_{k-1} \rangle = 0$ et $\langle r_k | \delta_{k-1} \rangle = 0$, on déduit que :

$$\begin{cases} \langle A\delta_{k-1} | r_k \rangle = \frac{\langle A\delta_{k-1} | \delta_{k-1} \rangle}{\langle r_{k-1} | \delta_{k-1} \rangle} \|r_k\|_2^2 \\ \langle A\delta_{k-1} | \delta_{k-1} \rangle = -\frac{\langle A\delta_{k-1} | \delta_{k-1} \rangle}{\langle r_{k-1} | \delta_{k-1} \rangle} \langle r_{k-1} | \delta_{k-1} \rangle \end{cases}$$

et $t_k = \frac{\|r_k\|_2^2}{\langle r_{k-1} \mid \delta_{k-1} \rangle}$. Enfin avec $\delta_k = r_k + t_k \delta_{k-1}$ et $\langle r_k \mid \delta_{k-1} \rangle = 0$, on déduit que $\langle r_k \mid \delta_k \rangle = \|r_k\|_2^2$ et $t_k = \frac{\|r_k\|_2^2}{\|r_{k-1}\|_2^2}$. $\square$

En résumé, à l'étape $k \geq 1$, en supposant toujours que r_k est non nul, le calcul de x_{k+1} se fait de la manière suivante :

$$\delta_k = r_k + \frac{\|r_k\|_2^2}{\|r_{k-1}\|_2^2} \delta_{k-1}, \quad x_{k+1} = x_k - \frac{\langle r_k \mid \delta_k \rangle}{\langle A\delta_k \mid \delta_k \rangle} \delta_k \quad (5.13)$$

Pour l'initialisation on se donne un vecteur x_0 dans $\mathbb{R}^n$ et on prend $\delta_0 = r_0$.

Lemme 5.12 *Avec les hypothèses et notations qui précèdent, pour j compris entre 0 et k , les vecteurs r_j sont deux à deux orthogonaux et les vecteurs δ_j sont deux à deux A -conjugués.*

Preuve. Si $r_0 = Ax_0 - b$ est non nul, on pose alors $\delta_0 = r_0$ et on peut construire les vecteurs x_1, r_1 et δ_1 , le vecteur r_1 étant orthogonal à $\delta_0 = r_0$ et le vecteur δ_1 A -conjugué avec δ_0 .

Supposons le résultat acquis jusqu'à l'ordre $j - 1$ avec j compris entre 2 et k . Il s'agit alors de montrer que $\langle r_j \mid \delta_i \rangle = 0$ et $\langle A\delta_j \mid \delta_i \rangle = 0$ pour i compris entre 0 et $j - 1$. On sait déjà que ces égalités sont vérifiées pour $i = j - 1$. On a :

$$r_j = r_{j-1} + \frac{\langle r_{j-1} \mid \delta_{j-1} \rangle}{\langle A\delta_{j-1} \mid \delta_{j-1} \rangle} A\delta_{j-1}$$

et avec $\langle r_{j-1} \mid r_{j-2} \rangle = 0$, il reste $\langle r_j \mid r_{j-2} \rangle = \frac{\langle r_{j-1} \mid \delta_{j-1} \rangle}{\langle A\delta_{j-1} \mid \delta_{j-1} \rangle} \langle A\delta_{j-1} \mid r_{j-2} \rangle$.

Pour $j = 2$ on a $r_0 = \delta_0$ et $\langle A\delta_1 \mid r_0 \rangle = \langle A\delta_1 \mid \delta_0 \rangle = 0$.

Pour $j > 2$, on a $r_{j-2} = \delta_{j-2} - \frac{\|r_{j-2}\|_2^2}{\|r_{j-3}\|_2^2} \delta_{j-3}$ et avec l'hypothèse de récurrence

on aboutit à $\langle A\delta_{j-1} \mid r_{j-2} \rangle = \langle A\delta_{j-1} \mid \delta_{j-2} \rangle - \frac{\|r_{j-2}\|_2^2}{\|r_{j-3}\|_2^2} \langle A\delta_{j-1} \mid \delta_{j-3} \rangle = 0$.

On a donc $\langle r_j \mid r_{j-2} \rangle = 0$.

D'autre part, on a :

$$\langle A\delta_j \mid \delta_{j-2} \rangle = \langle \delta_j \mid A\delta_{j-2} \rangle = \left\langle r_j + \frac{\|r_j\|_2^2}{\|r_{j-1}\|_2^2} \delta_{j-1} \mid A\delta_{j-2} \right\rangle = \langle r_j \mid A\delta_{j-2} \rangle$$

et avec $r_{j-1} = r_{j-2} + \frac{\langle r_{j-2} \mid \delta_{j-2} \rangle}{\langle A\delta_{j-2} \mid \delta_{j-2} \rangle} A\delta_{j-2}$, on a :

$$\langle r_j \mid A\delta_{j-2} \rangle = \frac{\langle A\delta_{j-2} \mid \delta_{j-2} \rangle}{\langle r_{j-2} \mid \delta_{j-2} \rangle} (\langle r_j \mid r_{j-1} \rangle - \langle r_j \mid r_{j-2} \rangle) = 0$$

(on a $\langle r_{j-2} \mid \delta_{j-2} \rangle = \|r_{j-2}\|_2^2 \neq 0$). On a donc $\langle A\delta_j \mid \delta_{j-2} \rangle = 0$. De manière analogue on montre que $\langle r_j \mid r_i \rangle = 0$ et $\langle A\delta_j \mid \delta_i \rangle = 0$ pour $i = j - 3, \dots, 0$. $\square$

Les vecteurs r_j étant deux à deux orthogonaux et non nuls pour j compris entre 0 et k , on a nécessairement $k \leq n-1$ (dans $\mathbb{R}^n$ il n'est pas possible d'avoir plus de n vecteurs orthogonaux non nuls, puisqu'un tel système est nécessairement libre), c'est-à-dire que le vecteur r_n est nécessairement nul. Précisément il existe un entier p compris entre 0 et $n-1$ tel que $r_p \neq 0$ et $r_{p+1} = 0$, ce qui équivaut à $Ax_{p+1} = b$ et $Ax_k = b$ pour tout $k \geq p+1$. En d'autres termes on a obtenu le résultat suivant.

Théorème 5.16. Stiefel

La méthode du gradient conjugué converge en n itérations au plus, c'est-à-dire que la suite de vecteurs $(x_k)_{k \in \mathbb{N}}$ définie par (5.13) avec $x_0 \in \mathbb{R}^n$ et $\delta_0 = r_0$ est stationnaire à partir d'un rang $p \leq n$ sur la solution du système linéaire $Ax = b$.

5.16 Exercices

Exercice 5.1. Soient α, β dans $\mathbb{K}$ et $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre n supérieur ou égal à 3 définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, \dots, n\} \setminus \{i\} \end{cases}$$

1. Donner une condition nécessaire et suffisante sur α et β pour que la matrice $A(\alpha, \beta)$ soit inversible.
2. On suppose que $A(\alpha, \beta)$ est inversible. Résoudre le système linéaire $A(\alpha, \beta)x = e$ dans les deux cas suivants :

- (a) Toutes les composantes de e valent 1.
- (b) e est un vecteur quelconque.

Solution.

1. Le déterminant de $A(\alpha, \beta)$ est $\Delta(\alpha, \beta) = (\beta + (n-1)\alpha)(\beta - \alpha)^{n-1}$ (exercice 1.7), donc $A(\alpha, \beta)$ est inversible si, et seulement si, $\beta \neq \alpha$ et $\beta \neq -(n-1)\alpha$.
- 2.

(a) Ajoutant les lignes 2 à n à la première, on obtient $\alpha \sum_{j=1}^n x_j = \frac{n\alpha}{\beta + (n-1)\alpha}$, puis en retranchant cette équation aux équations 1 à n , on obtient :

$$\forall i \in \{1, \dots, n\}, x_i = \frac{1}{\beta + (n-1)\alpha}$$

- (b) Les mêmes opérations qu'en 2a donnent $\alpha \sum_{j=1}^n x_j = \frac{\alpha S(e)}{\beta + (n-1)\alpha}$, où $S(e) = \sum_{i=1}^n e_i$, puis :

$$\forall i \in \{1, \dots, n\}, x_i = \frac{1}{\beta - \alpha} \left(e_i - \frac{\alpha S(e)}{\beta + (n-1)\alpha} \right)$$

Exercice 5.2. Pour tout entier $n \geq 1$, on note H_n la matrice de Hilbert d'ordre n définie par $H_n = \left(\left(\frac{1}{i+j-1} \right) \right)_{1 \leq i, j \leq n}$.

1. Calculer le coefficient de $\frac{1}{X+n}$ dans la décomposition en éléments

$$\text{simples de la fraction rationnelle } R_n(X) = \frac{\prod_{k=1}^n (X-k)}{\prod_{k=0}^n (X+k)}.$$

2. On note $A_n = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre n définie par :

$$a_{ij} = \begin{cases} \frac{1}{i+j-1} & \text{si } j = 1, \dots, n-1, i = 1, \dots, n \\ R_{n-1}(i) & \text{si } j = n, i = 1, \dots, n \end{cases}$$

(a) Montrer que $\det(A_n) = \binom{2n}{n} \det(H_n)$.

(b) Montrer que $\det(A_n) = R_{n-1}(n) \det(H_{n-1})$.

3. En notant, pour $n \in \mathbb{N}^*$, $\Phi_n = \prod_{k=1}^n k!$, montrer que $\det(H_n) = \frac{\Phi_{n-1}^4}{\Phi_{2n-1}}$.

4. Montrer que pour tout $n \in \mathbb{N}^*$, on a $\det(H_n) \leq \frac{1}{n^n}$.

Solution.

1. La décomposition en éléments simples de R_n s'écrit $R_n(X) = \sum_{k=0}^n \frac{\lambda_{k,n}}{(X+k)}$, avec

$$\lambda_{k,n} = \lim_{x \rightarrow -k} ((x+k) R_n(x)). \text{ En particulier le coefficient de } \frac{1}{X+n} \text{ est donné}$$

$$\text{par } \lambda_{n,n} = \frac{(2n)!}{(n!)^2} = \binom{2n}{n}.$$

2. On a, en notant $h_{ij} = \frac{1}{i+j-1}$ pour i, j compris entre 1 et n :

$$A_n = \begin{pmatrix} h_{11} & \cdots & h_{1,n-1} & R_{n-1}(1) \\ h_{21} & \cdots & h_{2,n-1} & R_{n-1}(2) \\ \vdots & \ddots & \vdots & \vdots \\ h_{n1} & \cdots & h_{n,n-1} & R_{n-1}(n) \end{pmatrix},$$

avec $R_{n-1}(i) = \sum_{k=0}^{n-1} \frac{\lambda_{k,n-1}}{(i+k)} = \sum_{k=0}^{n-1} \lambda_{k,n-1} h_{i,k+1} = \sum_{j=1}^n \lambda_{j-1,n-1} h_{i,j}$ pour tout i compris entre 1 et n .

- (a) En notant C_j la colonne numéro j de H_n , la colonne numéro n de A_n est combinaison linéaire des $C_1, \dots, C_n$. Le déterminant étant une forme n -linéaire alternée, on en déduit que :

$$\det(A_n) = \lambda_{n-1,n-1} \det(H_n) = \binom{2(n-1)}{n-1} \det(H_n)$$

- (b) On a $R_{n-1}(i) = 0$ pour $i = 1, \dots, n-1$ et en développant le déterminant suivant la dernière colonne, on en déduit que $\det(A_n) = R_{n-1}(n) \det(H_{n-1})$.

3. On en déduit que $\det(H_n) = \frac{R_{n-1}(n)}{\lambda_{n-1,n-1}} \det(H_{n-1})$ avec $\lambda_{n-1,n-1} = \frac{(2n-2)!}{((n-1)!)^2}$

et $R_{n-1}(n) = \frac{((n-1)!)^2}{(2n-1)!}$, ce qui donne :

$$\det(H_n) = \frac{((n-1)!)^4}{(2n-1)!(2n-2)!} \det(H_{n-1})$$

Par récurrence, on en déduit que :

$$\det(H_n) = \frac{((n-1)!(n-2)!\cdots 2!)^4}{(2n-1)!(2n-2)!\cdots 3!2!} = \frac{\Phi_{n-1}^4}{\Phi_{2n-1}}$$

4. On a $\det(H_n) = \frac{((n-1)!(n-2)!\cdots 2!)^3}{(2n-1)!(2n-2)!\cdots n!}$, soit avec $\binom{2k}{k} = \frac{(2k)!}{(k!)^2}$:

$$\det(H_n) = \frac{(n-1)!}{(2n-1)!} \frac{1}{\binom{2(n-1)}{n-1}} \frac{(n-2)!}{(2n-3)!} \frac{1}{\binom{2(n-2)}{n-2}} \cdots \frac{2!}{5!} \frac{1}{\binom{4}{2}} \frac{1}{3!} \frac{1}{\binom{2}{1}}$$

En posant $\psi_k = \frac{k!}{(2k+1)!} \frac{1}{\binom{2k}{k}}$, on a $\det(H_n) = \prod_{k=1}^{n-1} \psi_k$. Avec $\binom{2k}{k} \geq 1$, on

déduit que $\psi_k \leq 1$ et :

$$\det(H_n) \leq \psi_{n-1} = \frac{(n-1)!}{(2n-1)!} \frac{1}{\binom{2(n-1)}{n-1}} \leq \frac{(n-1)!}{(2n-1)!} = \frac{1}{(2n-1)\cdots n} \leq \frac{1}{n^n}$$

Ce qui nous dit que $\det(H_n)$ tend vers 0 très vite quand n tend vers l'infini.

Exercice 5.3. Soit $A = \begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix}$ une matrice

tridiagonale à coefficients réels ou complexes dont tous les déterminants principaux sont non nuls. Donner un algorithme de calcul du déterminant de la matrice A . Décrire un algorithme de résolution du système linéaire $Ax = e$ qui utilise la méthode de Gauss sans échange de lignes.

Solution. En notant D_k le déterminant principal d'ordre k d'une matrice tridiagonale d'ordre n et en développant D_k suivant la dernière ligne on a la relation de récurrence $D_k = a_k D_{k-1} - b_k c_{k-1} D_{k-2}$, ce qui donne, avec les valeurs initiales $D_0 = 1$ et $D_1 = a_1$, un algorithme de calcul très simple.

Dans le cas où tous les déterminants principaux de la matrice A sont non nuls, la méthode de Gauss peut s'effectuer sans permutation de lignes (une telle permutation ferait perdre le caractère tridiagonal), on n'aura à chaque étape qu'une ligne à traiter et pour chaque ligne seulement deux opérations.

À l'étape k de la méthode, la ligne L_k du système devient $L_k - m_k L_{k-1}$ ($k = 2, \dots, n$) avec $m_k = \frac{b_k}{a_{k-1}}$. Ce qui donne les formules de transformations :

$$b_k = 0, \quad a_k = a_k - m_k c_{k-1}, \quad c_k \text{ inchangé}, \quad e_k = e_k - m_k e_{k-1}$$

Le système triangulaire supérieur obtenu sera bidiagonal et aura pour solution :

$$\begin{cases} x_n = \frac{e_n}{a_n} \\ x_i = \frac{(e_i - c_i x_{i+1})}{a_i} \quad (i = n-1, \dots, 1) \end{cases}$$

Cette méthode de résolution d'un système tridiagonal est la méthode de double balayage de Cholesky. Le produit des a_i donnera, en fin d'opération, le déterminant de la matrice A .

Exercice 5.4. On reprend les notations de l'exercice 5.3. Décrire un algorithme de résolution du système $Ax = e$ qui utilise la décomposition LR . Décrire un algorithme de calcul de l'inverse de A qui utilise la décomposition LR .

Solution. Si tous les déterminants principaux de la matrice A sont non nuls on sait alors que cette matrice admet une décomposition LR avec L triangulaire inférieure à diagonale unité et R triangulaire supérieure.

On cherche les matrices L et R sous la forme bidiagonale, soit :

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ L_2 & 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & L_{n-1} & 1 & 0 \\ 0 & \cdots & 0 & L_n & 1 \end{pmatrix}, R = \begin{pmatrix} d_1 & r_1 & 0 & \cdots & 0 \\ 0 & d_2 & r_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 0 & d_{n-1} & r_{n-1} \\ 0 & \cdots & 0 & 0 & d_n \end{pmatrix}$$

On effectuant le produit $A = LR$ et en identifiant avec les coefficients de A , on obtient alors :

$$\begin{cases} d_1 = a_1 \\ r_j = c_j \quad (j = 1, \dots, n-1) \end{cases}$$

puis :

$$L_{j+1} = \frac{b_{j+1}}{d_j}, \quad d_{j+1} = a_{j+1} - r_j L_{j+1} \quad (j = 1, \dots, n-1)$$

On en déduit alors une méthode de résolution d'un système tridiagonal utilisant la décomposition LR . Soit donc à résoudre le système linéaire $Ax = e$. Tout d'abord, on résout le système linéaire $Ly = e$, ce qui donne :

$$\begin{cases} y_1 = e_1 \\ y_j = e_j - L_j y_{j-1} \quad (j = 2, \dots, n) \end{cases}$$

Puis on résout le système linéaire $Rx = y$, ce qui donne :

$$\begin{cases} x_n = \frac{y_n}{d_n} \\ x_j = \frac{(y_j - r_j x_{j+1})}{d_j} \quad (j = n-1, \dots, 1) \end{cases}$$

On a $A^{-1} = R^{-1}L^{-1}$. En utilisant les résultats du paragraphe 5.3, on sait que les coefficients de $U = R^{-1}$ et $V = L^{-1}$ sont donnés, pour j compris entre 1 et n , par :

$$\begin{cases} u_{jj} = \frac{1}{d_j} \\ u_{ij} = -\frac{r_i u_{i+1,j}}{d_i} \quad (i = j-1, \dots, 1) \\ u_{ij} = 0 \quad (i = j+1, \dots, n) \end{cases}, \begin{cases} v_{ij} = 0 \quad (i = 1, \dots, j-1) \\ u_{jj} = 1 \\ v_{ij} = -L_i u_{i-1,j} \quad (i = j+1, \dots, n) \end{cases}$$

Les coefficients de $W = A^{-1}$ sont alors donnés par $w_{ij} = \sum_{k=\max(i,j)}^n u_{ik} v_{kj}$ pour i, j compris entre 1 et n .

Exercice 5.5. On note A_n la matrice symétrique réelle d'ordre $n \geq 2$ définie par $A_n = ((a_{ij}))_{1 \leq i, j \leq n}$, avec :

$$\begin{cases} a_{ij} = n - j + 1 \text{ pour } 1 \leq i \leq j \leq n \\ a_{ij} = a_{ji} \text{ pour } 1 \leq j < i \leq n \end{cases}$$

1. Calculer $\det(A_n)$.
2. Calculer l'inverse B_n de la matrice A_n .
3. Donner la décomposition LR de A_n .

Solution.

1. On a $A_n = \begin{pmatrix} n & n-1 & \cdots & 2 & 1 \\ n-1 & n-1 & \cdots & 2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 2 & \cdots & 2 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$. En retranchant la deuxième ligne

à la première, on déduit que $\det(A_n) = \det(A_{n-1})$ pour tout $n \geq 2$, puis par récurrence que $\det(A_n) = 1$.

2. La matrice A_n est inversible dans $M_n(\mathbb{Z})$ (son déterminant vaut 1), c'est-à-dire que les coefficients de A_n^{-1} sont entiers. Après quelques expériences numériques on fait l'hypothèse que :

$$A_n^{-1} = B_n = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Pour vérifier cette hypothèse, il suffit de vérifier que la colonne numéro j de B_n est solution du système linéaire $A_n x = e_j$ où e_j est le j -ième vecteur de base canonique de $\mathbb{R}^n$.

3. Si $A_n = LR$, alors $A_n^{-1} = R^{-1}L^{-1}$ avec $L' = L^{-1}$ triangulaire inférieure à diagonale unité et $R' = R^{-1}$ triangulaire supérieure. La matrice B_n étant triangulaire, on cherche les matrices L' et R' sous la forme :

$$L' = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ L'_2 & 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & L'_{n-1} & 1 & 0 \\ 0 & \cdots & 0 & L'_n & 1 \end{pmatrix}, \quad R' = \begin{pmatrix} d'_1 & r'_1 & 0 & \cdots & 0 \\ 0 & d'_2 & r'_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 0 & d'_{n-1} & r'_{n-1} \\ 0 & \cdots & 0 & 0 & d'_n \end{pmatrix}$$

Par identification des coefficients dans le produit $A_n^{-1} = R^{-1}L^{-1}$, on obtient :

$$\begin{cases} r'_i = -1, (i = 1, \dots, n-1) \\ d'_1 = \frac{1}{n} \\ d'_i = -\frac{1}{L'_i} = \frac{n+2-i}{n+1-i}, (i = 2, \dots, n) \end{cases}$$

On vérifie alors que l'inverse de la matrice R' est la matrice :

$$R = \begin{pmatrix} n & n-1 & \cdots & 2 & 1 \\ 0 & \frac{n-1}{n} & \cdots & \frac{2}{n} & \frac{1}{n} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \cdots & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

et que l'inverse de la matrice L' est la matrice :

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{n-1}{n} & 1 & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{2}{n} & \frac{2}{n-1} & \cdots & 1 & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \cdots & \frac{1}{2} & 1 \end{pmatrix}$$

Exercice 5.6. Soient α, β dans $\mathbb{K}$ avec $\beta \neq 0$ et $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre n supérieur ou égal à 3 définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, \dots, n\} \setminus \{i\} \end{cases}$$

1. Calculer le rayon spectral de la matrice J intervenant dans la méthode de Jacobi (on peut utiliser les résultats de l'exercice 3.20).
2. Donner une condition nécessaire et suffisante sur la matrice $A(\alpha, \beta)$ pour que la méthode de Jacobi converge.
3. Montrer que si la méthode de Jacobi converge, celle de Gauss-Seidel converge aussi.

Solution. On pose $D = \beta I_n$ et $N = D - A(\alpha, \beta)$. La matrice intervenant dans la méthode de Jacobi est alors donnée par $J = D^{-1}N = A\left(-\frac{\alpha}{\beta}, 0\right)$.

1. Le rayon spectral de J est donné par $\rho(J) = \rho\left(A\left(-\frac{\alpha}{\beta}, 0\right)\right) = (n-1) \frac{|\alpha|}{|\beta|}$ (exercice 3.20).
2. La méthode de Jacobi est convergente si, et seulement si, le rayon spectral de J est strictement inférieur à 1, ce qui équivaut à $|\beta| > (n-1)|\alpha|$, ce qui encore équivalent à dire que la matrice $A(\alpha, \beta)$ est à diagonale strictement dominante.

3. Si la méthode de Jacobi converge, $A(\alpha, \beta)$ est alors à diagonale strictement dominante et la méthode de Gauss-Seidel converge aussi.

Exercice 5.7. On garde les notations de l'exercice 5.6 et en utilisant les notations de l'exercice 1.8 (question 2), on pose $M = M(\alpha, \beta, 0)$, $N = M - A(\alpha, \beta)$ et $G = M^{-1}N$ désigne la matrice qui intervient dans la méthode de Gauss-Seidel.

1. Calculer le polynôme caractéristique de G .
2. Montrer que 0 est valeur propre de G et calculer le produit des valeurs propres non nulles dans le cas où α est non nul.
3. Montrer que pour $|\alpha| \geq |\beta|$, la méthode de Gauss-Seidel ne converge pas.

Solution.

1. Le polynôme caractéristique de G est :

$$\chi_G(X) = \det(XI_n - M^{-1}N) = \det(M^{-1}) \det(MX - N)$$

Soit en utilisant les notations et les résultats de l'exercice 1.8 :

$$\chi_G(X) = \frac{\det(M(\alpha X, \beta X, \alpha))}{\beta^n} = \frac{(\beta - \alpha)^n X^n - X(\beta X - \alpha)^n}{\beta^n (1 - X)}$$

Ce qui peut aussi s'écrire :

$$\begin{aligned} \chi_G(X) &= \frac{(X-1) \left(X - \frac{\alpha}{\beta}\right)^n + \left(X - \frac{\alpha}{\beta}\right)^n - \left(1 - \frac{\alpha}{\beta}\right)^n X^n}{X-1} \\ &= \left(X - \frac{\alpha}{\beta}\right)^n + \frac{\alpha}{\beta} \sum_{k=1}^n \left(X - \frac{\alpha}{\beta}\right)^{n-k} \left(1 - \frac{\alpha}{\beta}\right)^{k-1} X^{k-1} \end{aligned}$$

2. On a $\chi_G(0) = 0$, $\chi'_G(0) = \lim_{x \rightarrow 0} \frac{\chi_G(x)}{x} = (-1)^n \frac{\alpha^n}{\beta^n}$. Pour α non nul, 0 est valeur

propre simple de G et en écrivant χ_G sous la forme $\chi_G(X) = X \prod_{i=2}^n (X - \lambda_i)$, on en déduit que le produit des valeurs propres non nulles de G est donné par $\prod_{i=2}^n \lambda_i = \frac{\alpha^n}{\beta^n}$.

3. Pour $|\alpha| \geq |\beta|$, on a $\left| \prod_{i=2}^n \lambda_i \right| \geq 1$, donc $\rho(G) \geq 1$ et la méthode de Gauss-Seidel ne converge pas.

Exercice 5.8. Soit $A = \begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix}$ une matrice

tridiagonale à coefficients réels ou complexes d'ordre $n \geq 3$, les coefficients a_i étant tous non nuls.

1. Montrer que le polynôme caractéristique de la matrice J qui intervient dans la méthode de Jacobi s'écrit $\chi_J(X) = P(X^2)X^q$, où P est un polynôme tel que $P(0)$ soit non nul.
2. Montrer que le polynôme caractéristique de la matrice G qui intervient dans la méthode de Gauss-Seidel est tel que $\chi_G(X^2) = X^n \chi_J(X)$.
3. Montrer que $\rho(G) = (\rho(J))^2$ et $R_\infty(G) = 2R_\infty(J)$ et conclure.

Solution. On désigne par D la partie diagonale de A , E sa partie triangulaire inférieure strict et par F sa partie triangulaire supérieure strict, de sorte que la matrice A s'écrit $A = E + D + F$. La matrice qui intervient dans la méthode de Jacobi est alors donnée par $J = -D^{-1}(E + F)$ et celle qui intervient dans la méthode de Gauss-Seidel par $G = -(D + E)^{-1}F$.

1. Le polynôme caractéristique de J est défini par :

$$\chi_J(X) = \det(XI_n + D^{-1}(E + F)) = \det(D^{-1}) \det(E + F + DX)$$

la matrice $M = E + F + DX$ étant tridiagonale. Avec les notations du lemme 5.4, on a $M(-1) = -E - F + DX$ et :

$$\det(M) = \det(M(-1)) = (-1)^n \det(E + F - DX)$$

On en déduit que $\chi_J(-X) = (-1)^n \chi_J(X)$, donc il existe un polynôme P tel que $P(0) \neq 0$ et $\chi_J(X) = P(X^2)X^q$.

2. Le polynôme caractéristique de G est défini par :

$$\chi_G(X) = \det(XI_n + (D + E)^{-1}F) = \det((D + E)^{-1}) \det(F + (D + E)X)$$

avec $M = F + (D + E)X$ tridiagonale. Pour $\lambda > 0$, avec les notations du lemme 5.4, on a $M\left(\frac{1}{\sqrt{\lambda}}\right) = \sqrt{\lambda}(E + F) + \lambda D = \sqrt{\lambda}(E + F + \sqrt{\lambda}D)$ et en conséquence, $\det(M) = \det\left(M\left(\frac{1}{\sqrt{\lambda}}\right)\right) = (\sqrt{\lambda})^n \det(E + F + \sqrt{\lambda}D)$, donc :

$$\begin{aligned} \chi_G(X^2) &= \frac{1}{\det(D + E)} X^n \det(E + F + DX) \\ &= \frac{X^n}{\det(D + E)} \chi_J(X) \det(D) = X^n \chi_J(X) \end{aligned}$$

3. C'est une conséquence immédiate de ce qui précède. La méthode de Jacobi converge si, et seulement si, $\rho(J) < 1$, ce qui équivaut à $\rho(G) < 1$ donc à la convergence de la méthode de Gauss-Seidel. En conclusion les méthodes de Jacobi et Gauss-Seidel convergent ou divergent simultanément et dans le cas de la convergence c'est la méthode de Gauss-Seidel qui est la plus rapide.

Exercice 5.9. Soit A une matrice réelle d'ordre $n \geq 2$ telle qu'il existe un réel α strictement positif avec :

$$\forall x \in \mathbb{R}^n, \langle Ax \mid x \rangle \geq \alpha \|x\|_2^2 \quad (5.14)$$

1. Montrer que pour tout vecteur $b \in \mathbb{R}^n$, le système linéaire $Ax = b$ admet une unique solution $u \in \mathbb{R}^n$.
2. Soient θ un réel et $(x_k)_{k \in \mathbb{N}}$ la suite de vecteurs définie par :

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, x_{k+1} = x_k - \theta (Ax_k - b) \end{cases} \quad (5.15)$$

Montrer que pour $\theta \in \left] 0, \frac{2\alpha}{\|A\|_2^2} \right[$ la suite $(x_k)_{k \in \mathbb{N}}$ converge vers u .

Solution.

1. Avec l'inégalité (5.14), on déduit que l'équation $Ax = 0$ équivaut à $x = 0$, ce qui équivaut à dire que la matrice A est inversible. En conséquence, pour tout vecteur $b \in \mathbb{R}^n$ le système linéaire $Ax = b$ admet une unique solution $u \in \mathbb{R}^n$.
2. On note $r_k = Ax_k - b$ le vecteur résidu d'ordre k . La suite $(x_k)_{k \in \mathbb{N}}$ converge vers u si, et seulement si, la suite $(r_k)_{k \in \mathbb{N}}$ converge vers le vecteur nul. En remarquant que $r_{k+1} = r_k - \theta Ar_k = (I_n - \theta A) r_k$, on déduit que $r_k = (I_n - \theta A)^k r_0$ pour tout entier naturel k et la suite $(r_k)_{k \in \mathbb{N}}$ converge vers le vecteur nul quelle que soit la valeur initiale r_0 si, et seulement si, $\rho(I_n - \theta A)$ est strictement inférieur à 1. Avec $\rho(I_n - \theta A) \leq \|I_n - \theta A\|_2$, on déduit que si $\|I_n - \theta A\|_2 < 1$, la suite $(x_k)_{k \in \mathbb{N}}$ converge alors vers u . Pour $x \in \mathbb{R}^n$ tel que $\|x\|_2 = 1$, on a :

$$\begin{aligned} \|(I_n - \theta A)x\|_2^2 &= \|Ax\|_2^2 \theta^2 - 2 \langle Ax \mid x \rangle \theta + \|x\|_2^2 \\ &\leq \|A\|_2^2 \|x\|_2^2 \theta^2 - 2\alpha \|x\|_2^2 \theta + \|x\|_2^2 = \|A\|_2^2 \theta^2 - 2\alpha \theta + 1 \end{aligned}$$

et pour $\|A\|_2^2 \theta^2 - 2\alpha \theta + 1 < 1$, ce qui équivaut à $\theta < \frac{2\alpha}{\|A\|_2^2}$ avec $\theta > 0$, on a $\|I_n - \theta A\|_2 < 1$. En conséquence $\rho(I_n - \theta A) < 1$ et la suite $(x_k)_{k \in \mathbb{N}}$ converge vers u .

Exercice 5.10. Soient A une matrice réelle d'ordre $n \geq 2$ symétrique définie positive de valeurs propres $0 < \lambda_1 \leq \dots \leq \lambda_n$ et b un vecteur dans $\mathbb{R}^n$.

1. Montrer que la matrice A vérifie la condition (5.14) de l'exercice précédent.
2. Montrer que la suite de vecteurs $(x_k)_{k \in \mathbb{N}}$ définie par (5.15) dans l'exercice précédent converge vers la solution u du système linéaire $Ax = b$ quelle que soit la valeur initiale x_0 si, et seulement si, $\theta \in \left]0, \frac{2}{\lambda_n}\right[$.
3. Montrer que, pour x_0 donné, le meilleur choix de $\theta \in \left]0, \frac{2}{\lambda_n}\right[$ (i. e. celui qui assure la convergence la plus rapide) est $\theta_0 = \frac{2}{\lambda_1 + \lambda_n}$ et calculer le rayon spectral ρ_0 de la matrice $I_n - \theta_0 A$.
4. Calculer les valeurs de θ_0 et ρ_0 pour la matrice :

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

Solution.

1. Voir le lemme 5.10.
2. Dans l'exercice précédent on a vu que la suite $(x_k)_{k \in \mathbb{N}}$ converge vers u quelle que soit la valeur initiale x_0 si, et seulement si, $\rho(I_n - \theta A)$ est strictement inférieur à 1, ce qui équivaut à dire que pour tout entier j compris entre 1 et n on a $-1 < 1 - \theta \lambda_j < 1$, ce qui équivaut à $\theta > 0$ et $\theta < \frac{2}{\lambda_n}$, soit à $\theta \in \left]0, \frac{2}{\lambda_n}\right[$.
3. Le meilleur choix de θ est celui qui minimise le rayon spectral :

$$\rho(I_n - \theta A) = \max_{1 \leq j \leq n} |1 - \theta \lambda_j| = \max(|1 - \theta \lambda_1|, |1 - \theta \lambda_n|)$$

En dessinant le graphe de la fonction $\theta \mapsto \max(|1 - \theta \lambda_1|, |1 - \theta \lambda_n|)$, on voit facilement que cette fonction atteint son minimum en $\theta_0 = \frac{2}{\lambda_1 + \lambda_n}$ avec :

$$\begin{aligned} \rho_0 &= \rho(I_n - \theta_0 A) = \max_{1 \leq j \leq n} \left| 1 - \frac{2}{\lambda_1 + \lambda_n} \lambda_j \right| \\ &= \max_{1 \leq j \leq n} \frac{(\lambda_n - \lambda_j) - (\lambda_j - \lambda_1)}{\lambda_1 + \lambda_n} = \frac{\lambda_n - \lambda_1}{\lambda_1 + \lambda_n} \end{aligned}$$

Ce qui s'écrit aussi $\rho_0 = \frac{1 - \frac{\lambda_1}{\lambda_n}}{1 + \frac{\lambda_1}{\lambda_n}} = \frac{1 - \text{cond}_2(A)}{1 + \text{cond}_2(A)}$. La méthode (méthode de descente à paramètre constant de Richardson) sera donc rapidement convergente pour une matrice bien conditionnée, c'est-à-dire telle $\text{cond}_2(A)$ est voisin de 1.

4. Les valeurs propres de la matrice A sont données par :

$$\lambda_k(a, b) = 2 - 2 \cos \left(\frac{k\pi}{n+1} \right) = 4 \sin^2 \left(\frac{k\pi}{2(n+1)} \right) \quad (1 \leq k \leq n)$$

(exercice 1.9). On a donc :

$$\theta_0 = \frac{1}{2 \left(\sin^2 \left(\frac{\pi}{2(n+1)} \right) + \sin^2 \left(\frac{n\pi}{2(n+1)} \right) \right)}, \quad \rho_0 = \frac{\sin^2 \left(\frac{n\pi}{2(n+1)} \right) - \sin^2 \left(\frac{\pi}{2(n+1)} \right)}{\sin^2 \left(\frac{n\pi}{2(n+1)} \right) + \sin^2 \left(\frac{\pi}{2(n+1)} \right)}$$

Chapitre 6

Calcul approché des valeurs et vecteurs propres

6.1 Introduction

Le problème de la détermination des valeurs propres d'une matrice est en général beaucoup plus difficile que celui de la résolution des systèmes linéaires.

Tout d'abord nous allons étudier la méthode de la puissance itérée qui permet de calculer la valeur propre de plus grand module d'une matrice réelle sous certaines hypothèses. Puis en itérant ce procédé on peut en déduire les autres quand elles sont toutes distinctes en modules. Cette méthode, peu performante, est à utiliser pour les matrices ayant toutes les valeurs propres distinctes ou si on en cherche seulement à calculer quelques-unes d'entre elles.

D'autres méthodes sont les méthodes de Rutishauser, Givens et Householder et de Jacobi qui reposent sur le principe suivant : on construit une suite $(A_k)_{k \in \mathbb{N}}$ de matrices toutes semblables à la matrice A , donc ayant les mêmes valeurs propres, qui converge vers une matrice plus simple (diagonale pour Jacobi, tridiagonale pour Givens et Householder, triangulaire pour Rutishauser) encore semblable à la matrice A . Les valeurs propres cherchées sont alors les termes diagonaux de la matrice limite, dans le cas des méthodes de Jacobi et de Rutishauser.

Les problèmes posés par ces méthodes sont :

- savoir décrire le passage de A_k à A_{k+1} aussi simplement que possible c'est-à-dire avec un coût de calcul peu élevé ;
- les valeurs propres doivent être bien conservées (stabilité numérique de la méthode).

Comme pour le chapitre 5, on pourra consulter [?] ou [24] pour la programmation des méthodes décrites dans ce chapitre.

6.2 Méthode de la puissance itérée

La méthode de la puissance itérée permet le calcul du rayon spectral de la valeur propre de plus grand module d'une matrice réelle quand cette valeur propre

dominante est unique. Par exemple, les théorèmes de Perron-Frobenius (théorèmes 4.7 et 4.3) nous dit qu'une matrice réelle à coefficients strictement positifs a une valeur propre dominante unique.

Théorème 6.1.

Soit A une matrice d'ordre $n \geq 2$ à coefficients réels telle que la valeur propre λ_1 de module maximum soit unique. Cette valeur propre est alors réelle et simple, l'espace propre associé est une droite vectorielle et on a $\mathbb{R}^n = \ker(A - \lambda_1 I_n) \oplus \text{Im}(A - \lambda_1 I_n)$, le sous-espace vectoriel $\text{Im}(A - \lambda_1 I_n)$ étant stable par A .

Preuve. On note $\lambda_1, \dots, \lambda_n$ les valeurs propres de A dans $\mathbb{C}$, avec $\rho(A) = |\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$. La valeur propre λ_1 est nécessairement réelle et simple (la matrice A étant réelle son spectre est stable par conjugaison complexe et λ_1 est l'unique valeur propre de module maximum). On note $E_1 = \ker(A - \lambda_1 I_n)$ et $F_1 = \text{Im}(A - \lambda_1 I_n)$. La valeur propre λ_1 étant simple, l'espace propre associé E_1 est nécessairement de dimension 1. De plus, il est facile de vérifier que les sous-espaces vectoriels E_1 et F_1 sont stables par A . Pour montrer que E_1 et F_1 sont en somme directe, il suffit de montrer que leur intersection se réduit au vecteur nul (théorème du rang). Si $y \in E_1 \cap F_1$, il existe alors un vecteur $x \in \mathbb{R}^n$ tel que $y = (A - \lambda_1 I_n)x$ et $(A - \lambda_1 I_n)y = 0$. Le vecteur x est alors dans le noyau de $(A - \lambda_1 I_n)^2$. Les valeurs propres de $A_1 = A - \lambda_1 I_n$ sont $0, \lambda_2 - \lambda_1, \dots, \lambda_n - \lambda_1$, la valeur propre nulle étant simple. La matrice A_1^2 a donc pour valeurs propres $0, (\lambda_2 - \lambda_1)^2, \dots, (\lambda_n - \lambda_1)^2$, la valeur propre nulle étant simple. On en déduit alors que $\ker(A_1^2)$ est de dimension 1 et, tenant compte du fait que $\{0\} \neq \ker(A_1) \subset \ker(A_1^2)$, que $\ker(A_1) = \ker(A_1^2)$. Il en résulte $x \in \ker(A_1)$ et $y = 0$. On a donc $E_1 \cap F_1 = \{0\}$ et $\mathbb{R}^n = E_1 \oplus F_1$. $\square$

De manière plus générale, on peut montrer que si λ est une valeur propre de A d'ordre $p \geq 1$, les sous-espaces vectoriels $\ker(A - \lambda I_n)^p$ et $\text{Im}(A - \lambda I_n)^p$ sont alors stables par A , supplémentaires dans $\mathbb{R}^n$ et $\dim(E_p) = p$ (exercice 6.1).

Dans ce qui suit, on munit $\mathbb{R}^n$ d'une norme quelconque et A désigne une matrice d'ordre $n \geq 2$ à coefficients réels telle que la valeur propre λ_1 de module maximum soit unique. En particulier cette valeur propre est unique. On note $E_1 = \ker(A - \lambda_1 I_n)$, $F_1 = \text{Im}(A - \lambda_1 I_n)$ et on définit la suite $(x^{(k)})_{k \in \mathbb{N}}$ de vecteurs de $\mathbb{R}^n$ par :

$$\begin{cases} x^{(0)} = e_1 + f_1 \text{ avec } e_1 \in E_1 \setminus \{0\}, & f_1 \in F_1 \\ \forall k \in \mathbb{N}, x^{(k+1)} = \frac{1}{\|Ax^{(k)}\|} Ax^{(k)} \end{cases}$$

On verra au cours de la démonstration du théorème 6.2 que cette suite est bien définie.

On note, dans la base canonique de $\mathbb{R}^n$, $e_{1,j}$ les composantes du vecteur e_1 , $x_j^{(k)}$ celles de $x^{(k)}$ et $(Ax^{(k)})_j$ celles de $Ax^{(k)}$.

On a alors le résultat suivant qui permet de calculer une valeur approchée de la valeur propre de plus grand module et d'un vecteur propre associé.

Théorème 6.2.

Avec les notations ci-dessus, on a :

- (i) $\lim_{k \rightarrow +\infty} \|Ax^{(k)}\| = |\lambda_1| = \rho(A).$
- (ii) $\lim_{k \rightarrow +\infty} x^{(2k)} = v_1, \lim_{k \rightarrow +\infty} x^{(2k+1)} = v_2 = \text{signe}(\lambda_1) v_1$ où v_1 est un vecteur propre non nul associé à λ_1 .
- (iii) Pour tout $j \in \{1, \dots, n\}$ tel que $e_{1,j} \neq 0$, on a $\lim_{k \rightarrow +\infty} \frac{(Ax^{(k)})_j}{x_j^{(k)}} = \lambda_1.$

Preuve. On identifie une matrice à l'endomorphisme de $\mathbb{R}^n$ qu'elle définit et on note B la restriction de A à F_1 . D'après le théorème 6.1, B est un endomorphisme de F_1 de valeurs propres $\lambda_2, \dots, \lambda_n$.

On a $Ax^{(0)} = \lambda_1 e_1 + Bf_1 \neq 0$, donc $x^{(1)}$ est bien défini. Par récurrence, on voit que $x^{(k)}$ est bien défini avec une projection non nulle sur E_1 . De manière plus précise, on a :

$$\forall k \in \mathbb{N}, \begin{cases} A^k x^{(0)} = \lambda_1^k e_1 + B^k f \neq 0 \\ x^{(k)} = \frac{1}{\|A^k x^{(0)}\|} A^k x^{(0)} \end{cases}$$

En effet le résultat est vrai pour $k = 1$ et en le supposant vrai pour $k \geq 1$, on a $A^{k+1}x^{(0)} = \lambda_1^{k+1}e_1 + B^{k+1}f \neq 0$ et $Ax^{(k)} = \frac{1}{\|A^k x^{(0)}\|} A^{k+1}x^{(0)} \neq 0$, de sorte que :

$$x^{(k+1)} = \frac{\|A^k x^{(0)}\|}{\|A^{k+1}x^{(0)}\|} \frac{1}{\|A^k x^{(0)}\|} A^{k+1}x^{(0)} = \frac{1}{\|A^{k+1}x^{(0)}\|} A^{k+1}x^{(0)}$$

Pour tout entier $k \geq 1$, on a $A^k x^{(0)} = \lambda_1^k e_1 + B^k f_1 = \lambda_1^k (e_1 + f_k)$ avec $\lim_{k \rightarrow +\infty} f_k = \lim_{k \rightarrow +\infty} \left(\left(\frac{1}{\lambda_1} B \right)^k f_1 \right) = 0$ du fait que $\rho\left(\frac{1}{\lambda_1} B\right) < 1$.

En écrivant que $x^{(k)} = \left(\frac{\lambda_1}{|\lambda_1|} \right)^k \frac{1}{\|e_1 + f_k\|} (e_1 + f_k)$, on déduit que :

$$\|Ax^{(k)}\| = \frac{1}{\|e_1 + f_k\|} \|\lambda_1 e_1 + A f_k\|$$

et $\lim_{k \rightarrow +\infty} \|Ax^{(k)}\| = |\lambda_1|$. On déduit aussi que :

$$\lim_{k \rightarrow +\infty} x^{(2k)} = v_1 = \frac{1}{\|e_1\|} e_1, \lim_{k \rightarrow +\infty} x^{(2k+1)} = \frac{\text{signe}(\lambda_1)}{\|e_1\|} e_1$$

Enfin avec $Ax^{(k)} - \lambda_1 x^{(k)} = \left(\frac{\lambda_1}{|\lambda_1|} \right)^k \frac{1}{\|e_1 + f_k\|} (A f_k - \lambda_1 f_k)$, on déduit que $\lim_{k \rightarrow +\infty} (Ax^{(k)} - \lambda_1 x^{(k)}) = 0$.

Comme $e_1 \neq 0$, il existe $j \in \{1, \dots, n\}$ tel que $e_{1,j} \neq 0$ et pour k assez grand on a $x_j^{(k)} \neq 0$. Avec $\lim_{k \rightarrow +\infty} \left((Ax^{(k)})_j - \lambda_1 x_j^{(k)} \right) = 0$, on déduit alors que

$$\lim_{k \rightarrow +\infty} \frac{(Ax^{(k)})_j}{x_j^{(k)}} = \lambda_1. \quad \square$$

En général on n'a pas d'informations sur le sous-espace propre E_1 de sorte qu'en prenant $x^{(0)}$ quelconque dans $\mathbb{R}^n$ sa projection sur E_1 peut être nulle et théoriquement la méthode décrite ci-dessus ne donnera pas le résultat espéré. En fait F_1 est un hyperplan de $\mathbb{R}^n$, donc de mesure nulle, et en prenant $x^{(0)}$ quelconque dans $\mathbb{R}^n$ on a très peu de chances de tomber sur F_1 . Si $x^{(0)}$ est dans F_1 , à cause des erreurs d'arrondis, au bout de quelques itérations on sort de cet hyperplan, mais la convergence peut être lente. Une autre façon d'éviter ce type de problème est de prendre deux valeurs différentes de $x^{(0)}$ et de garder celle qui donne la convergence la plus rapide.

La méthode de la puissance itérée est en fait une adaptation de la méthode de Bernoulli pour calculer la racine de plus grand module d'un polynôme, quand cette dernière est unique.

Si la matrice A est inversible, en appliquant la méthode de la puissance itérée à la matrice A^{-1} , on a alors un moyen de calculer la valeur propre de plus petit module de A (quand cette dernière est unique). La méthode obtenue est aussi appelée méthode de la puissance inverse. Dans la pratique on évite d'inverser A et on utilise la suite définie par :

$$\begin{cases} x^{(0)} = e_n + f_n \text{ avec } e_n \in \ker(A - \lambda_n I_n) \setminus \{0\}, f_n \in \text{Im}(A - \lambda_n I_n) \\ \forall k \in \mathbb{N}, x^{(k+1)} = \frac{1}{\|u^{(k+1)}\|} u^{(k+1)}, \text{ avec } Au^{(k+1)} = x^{(k)} \end{cases}$$

Le calcul approché des autres valeurs propres peut se faire en utilisant la méthode de déflation dans l'hypothèse où les valeurs propres de A sont telles que $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$. La matrice A est alors diagonalisable avec des valeurs propres réelles et simples.

Soit e_1 un vecteur propre associé à λ_1 de norme euclidienne égale à 1. On va vérifier que les valeurs propres de la matrice $B = A - \lambda_1 e_1 {}^t e_1 = ((a_{ij} - \lambda_1 e_{1,i} e_{1,j}))$ sont $0, \lambda_2, \dots, \lambda_n$. On appliquera alors la méthode de la puissance itérée à cette nouvelle matrice pour obtenir une valeur approchée de λ_2 , puis on continue ainsi pour obtenir des valeurs approchées des autres valeurs propres.

Lemme 6.1 Avec les notations qui précèdent, en notant e_1 un vecteur propre de A associé à λ_1 de norme euclidienne égale à 1, les valeurs propres de la matrice $B = A - \lambda_1 e_1 {}^t e_1$ sont $0, \lambda_2, \dots, \lambda_n$.

Preuve. Pour tout $j = 1, 2, \dots, n$, λ_j est aussi valeur propre de ${}^t A$. On désigne par f_j un vecteur propre de ${}^t A$ associé à λ_j . En écrivant que :

$$\langle e_1 | f_j \rangle = \frac{1}{\lambda_1} \langle Ae_1 | f_j \rangle = \frac{1}{\lambda_1} \langle e_1 | {}^t A f_j \rangle = \frac{\lambda_j}{\lambda_1} \langle e_1 | f_j \rangle$$

on déduit que $\langle e_1 | f_j \rangle = 0$ pour tout j compris entre 2 et n . On a alors pour tout j compris entre 2 et n :

$${}^t B f_j = {}^t A f_j - \lambda_1 e_1 {}^t e_1 f_j = \lambda_j f_j - \lambda_1 \langle e_1 | f_j \rangle e_1 = \lambda_j f_j$$

c'est-à-dire que λ_j est valeur propre de ${}^tB$, c'est donc aussi une valeur propre de B . Enfin, on a $Be_1 = Ae_1 - \lambda_1 e_1$ ${}^te_1 e_1 = \lambda_1 e_1 - \lambda_1 \langle e_1 | e_1 \rangle e_1 = 0$, c'est à dire que 0 est valeur propre de B . $\square$

6.3 Méthode de Jacobi pour les matrices symétriques

On se donne, pour ce paragraphe, une matrice A symétrique réelle. Elle est donc diagonalisable avec ses valeurs propres toutes réelles.

Le principe de la méthode de Jacobi consiste à construire une suite de matrices de rotations planes $(R(\theta_k))_{k \in \mathbb{N}}$ telle que pour k tendant vers l'infini, la suite de matrices $(A_k)_{k \in \mathbb{N}}$ définie par :

$$\begin{cases} A_0 = A \\ A_{k+1} = R(\theta_k)^{-1} A_k R(\theta_k) \quad (k \geq 0) \end{cases}$$

converge vers une matrice diagonale. Comme chacune des matrices A_k est semblable à la matrice A , on en déduit que les valeurs propres de A sont les termes diagonaux de la matrice limite. Cette méthode est bien adaptée aux matrices symétriques réelles de petite taille.

Pour toute matrice $M = ((m_{ij}))_{1 \leq i, j \leq n}$ d'ordre $n \geq 2$ et à coefficients réels, on désigne par $\|M\|_s = \text{Tr}({}^tMM)^{\frac{1}{2}}$ la norme de Schur de M .

On vérifie facilement qu'on a $\|M\|_2 \leq \|M\|_s$, où $\|\cdot\|_2$ désigne la norme matricielle induite par la norme euclidienne et que $\|MR\|_s = \|RM\|_s = \|M\|_s$ pour toute matrice orthogonale R .

Pour tout $\theta \in [-\pi, \pi[$ et tout couple (p, q) d'entiers tel que $1 \leq p < q \leq n$, on note $R_{p,q}(\theta)$ la matrice de rotation d'angle θ dans le plan défini par les vecteurs e_p et e_q de la base canonique $(e_k)_{1 \leq k \leq n}$ de $\mathbb{R}^n$. Une telle matrice est de la forme

$$R_{p,q}(\theta) = \begin{pmatrix} I_{p-1} & 0 & 0 \\ 0 & \rho_{p,q}(\theta) & 0 \\ 0 & 0 & I_{n-q} \end{pmatrix}, \text{ où on a posé } \rho_{p,q}(\theta) = \begin{pmatrix} c & 0 & -s \\ 0 & I_{q-p+1} & 0 \\ s & 0 & c \end{pmatrix}$$

avec $c = \cos(\theta)$ et $s = \sin(\theta)$.

Pour toute matrice symétrique réelle d'ordre $n \geq 2$, $M = ((m_{ij}))_{1 \leq i, j \leq n}$, on note $M_{p,q}(\theta) = R_{p,q}(\theta)^{-1} M R_{p,q}(\theta)$.

Pour une matrice symétrique M , le produit $M_{p,q}(\theta)$ est facile à calculer. La multiplication à gauche par $R_{p,q}(\theta)^{-1} = {}^tR_{p,q}(\theta)$ modifie seulement les lignes p et q de M et la multiplication à droite par $R_{p,q}(\theta)$ change seulement les colonnes p et q de M . Ce qui donne pour les coefficients de $M_{p,q}(\theta) = ((m'_{ij}))_{1 \leq i, j \leq n}$:

$$\begin{cases} m'_{ij} = m_{ij} & (i \neq p, i \neq q, j \neq p, j \neq q) \\ m'_{ip} = cm_{ip} + sm_{iq} & (i \neq p, i \neq q) \\ m'_{pp} = c^2 m_{pp} + s^2 m_{qq} + 2scm_{pq} \\ m'_{iq} = cm_{iq} - sm_{ip} & (i \neq p, i \neq q) \\ m'_{qq} = s^2 m_{pp} + c^2 m_{qq} - 2scm_{pq} \\ m'_{pq} = (c^2 - s^2) m_{pq} - sc(m_{pp} - m_{qq}) \end{cases}$$

la matrice $M_{p,q}(\theta)$ étant symétrique. De plus on a $\|M_{p,q}(\theta)\|_s = \|M\|_s$.

L'idée de la méthode Jacobi est de déterminer θ de manière à annuler les coefficients m'_{pq} et m'_{qp} de $M_{p,q}(\theta)$.

Lemme 6.2 Soit $(p, q) \in \mathbb{N}^2$ tels que $1 \leq p < q \leq n$ et $m_{pq} \neq 0$.

1. Il existe un unique réel $\theta \in \left]-\frac{\pi}{4}, \frac{\pi}{4}\right] \setminus \{0\}$ tel que $m'_{pq} = 0$.
2. $t = \tan(\theta)$ est l'unique solution dans $] -1, 1] \setminus \{0\}$ de l'équation du second degré $t^2 + 2b_{pq}t - 1 = 0$, où on a posé $b_{pq} = \frac{m_{pp} - m_{qq}}{2m_{pq}}$.
3. $c = \cos(\theta)$ et $s = \sin(\theta)$ peuvent se calculer de manière algébrique avec les égalités $c = \frac{1}{\sqrt{1+t^2}}$ et $s = tc$.

Preuve.

1. La condition $m'_{pq} = 0$ équivaut à :

$$\left(\cos(\theta)^2 - \sin(\theta)^2\right) m_{pq} - \sin(\theta) \cos(\theta) (m_{pp} - m_{qq}) = 0$$

Pour $m_{pq} \neq 0$, on a $\theta \neq 0$. Si $m_{pp} = m_{qq}$, on prend alors $\theta = \frac{\pi}{4}$ et pour $m_{pp} \neq m_{qq}$ la condition ci-dessus s'écrit $\tan(2\theta) = \frac{2m_{pq}}{m_{pp} - m_{qq}}$ et cette équation admet une unique solution $\theta \in \left]-\frac{\pi}{4}, \frac{\pi}{4}\right[\setminus \{0\}$.

2. On déduit facilement du point précédent que $t = \tan(\theta) \in] -1, 1] \setminus \{0\}$ et qu'il est solution de l'équation du second degré $t^2 + 2b_{pq}t - 1 = 0$. Cette équation admet deux racines réelles :

$$\begin{cases} t_1 = -b_{pq} + \sqrt{1 + b_{pq}^2} = \frac{1}{b_{pq} + \sqrt{1 + b_{pq}^2}} \\ t_2 = -\frac{1}{t_1} = \frac{-1}{\sqrt{1 + b_{pq}^2} - b_{pq}} \end{cases}$$

La racine de valeur absolue inférieure à 1 est $t = \frac{\text{signe}(b_{pq})}{\sqrt{b_{pq}^2 + 1} + |b_{pq}|}$, avec la convention $\text{signe}(0) = 1$.

3. Des égalités $\cos(\theta) = \frac{1}{\sqrt{1 + \tan^2(\theta)}}$, $\sin(\theta) = \cos(\theta) \tan(\theta)$, on déduit que $c = \frac{1}{\sqrt{1 + t^2}}$ et $s = tc$.

□

En remarquant que $\tau = \tan\left(\frac{\theta}{2}\right) = \frac{s}{1+c}$, les formules donnant les coefficients de $M_{p,q}(\theta)$ se simplifient pour donner finalement :

$$\begin{cases} m'_{ij} = m_{ij} & (i \neq p, i \neq q, j \neq p, j \neq q) \\ m'_{ip} = m_{ip} + s(m_{iq} - \tau m_{ip}) & (i \neq p, i \neq q) \\ m'_{pp} = m_{pp} + tm_{pq} \\ m'_{iq} = m_{iq} - s(m_{ip} + \tau m_{iq}) & (i \neq p, i \neq q) \\ m'_{qq} = m_{qq} - tm_{pq} \\ m'_{pq} = 0 \end{cases} \quad (6.1)$$

Dans ce qui suit, A désigne une matrice symétrique réelle d'ordre $n \geq 3$ de valeurs propres (réelles) $\lambda_1, \dots, \lambda_n$. On lui associe la suite de matrices symétriques $(A_k)_{k \in \mathbb{N}}$ définie par $A_k = \left(\left(a_{ij}^{(k)} \right) \right)_{1 \leq i, j \leq n}$ et :

$$\begin{cases} A_0 = A \\ A_{k+1} = R(\theta_k)^{-1} A_k R(\theta_k) \quad (k \geq 0) \end{cases}$$

où $R(\theta_k)$ est la matrice unité si $a_{pq}^{(k)} = 0$ et une matrice de rotation $R_{p,q}(\theta_k)$ avec $1 \leq p < q \leq n$ tels que $\left| a_{pq}^{(k)} \right| = \max_{1 \leq i < j \leq n} \left| a_{ij}^{(k)} \right|$ et $\theta_k \in \left] -\frac{\pi}{4}, \frac{\pi}{4} \right] \setminus \{0\}$ tel que $a_{pq}^{(k+1)} = 0$.

On décompose chaque matrice A_k sous la forme $A_k = D_k + E_k$ où D_k est la partie diagonale de A_k .

Lemme 6.3 Avec les notations qui précèdent, on a $\lim_{k \rightarrow +\infty} E_k = 0$.

Preuve. En utilisant les formules simplifiées (6.1), on a :

$$\|D_{k+1}\|_s^2 = \sum_{i=1}^n \left| a_{ii}^{(k+1)} \right|^2 = \sum_{\substack{i=1 \\ i \neq p, i \neq q}}^n \left| a_{ii}^{(k)} \right|^2 + \left| a_{pp}^{(k+1)} \right|^2 + \left| a_{qq}^{(k+1)} \right|^2$$

avec :

$$\begin{aligned} \left| a_{pp}^{(k+1)} \right|^2 + \left| a_{qq}^{(k+1)} \right|^2 &= \left| a_{pp}^{(k)} + t_k a_{pq}^{(k)} \right|^2 + \left| a_{qq}^{(k)} - t_k a_{pq}^{(k)} \right|^2 \\ &= \left| a_{pp}^{(k)} \right|^2 + \left| a_{qq}^{(k)} \right|^2 + 2t_k a_{pq}^{(k)} \left(a_{pp}^{(k)} - a_{qq}^{(k)} + t_k a_{pq}^{(k)} \right) \end{aligned}$$

le réel t_k étant l'unique solution dans $] -1, 1[\setminus \{0\}$ de l'équation $t_k^2 + 2b_{pq}t_k - 1 = 0$.

On a alors $a_{pp}^{(k)} - a_{qq}^{(k)} = \frac{1 - t_k^2}{t_k} a_{pq}^{(k)}$ et :

$$\left| a_{pp}^{(k+1)} \right|^2 + \left| a_{qq}^{(k+1)} \right|^2 = \left| a_{pp}^{(k)} \right|^2 + \left| a_{qq}^{(k)} \right|^2 + 2 \left| a_{pq}^{(k)} \right|^2$$

ce qui donne $\|D_{k+1}\|_s^2 = \|D_k\|_s^2 + 2 \left| a_{pq}^{(k)} \right|^2$.

On peut aussi constater que :

$$\begin{pmatrix} a_{pp}^{(k+1)} & a_{pq}^{(k+1)} \\ a_{qp}^{(k+1)} & a_{qq}^{(k+1)} \end{pmatrix} = \begin{pmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{pmatrix} \begin{pmatrix} a_{pp}^{(k)} & a_{pq}^{(k)} \\ a_{qp}^{(k)} & a_{qq}^{(k)} \end{pmatrix} \begin{pmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{pmatrix}$$

et par conservation des normes on a immédiatement :

$$\left| a_{pp}^{(k+1)} \right|^2 + \left| a_{qq}^{(k+1)} \right|^2 + 2 \left| a_{pq}^{(k+1)} \right|^2 = \left| a_{pp}^{(k)} \right|^2 + \left| a_{qq}^{(k)} \right|^2 + 2 \left| a_{pq}^{(k)} \right|^2$$

On conclut alors avec $a_{pq}^{(k+1)} = 0$.

D'autre part, on a $\|A_k\|_s^2 = \|D_k\|_s^2 + \|E_k\|_s^2$ avec $\|A_k\|_s = \|A\|_s$ pour tout $k \geq 0$.

On en déduit alors que $\|E_{k+1}\|_s^2 = \|E_k\|_s^2 - 2 \left| a_{pq}^{(k)} \right|^2$.

En utilisant le fait que $a_{pq}^{(k)}$ est de module maximal dans E_k , on déduit que $E_{k,s}^2 \leq (n^2 - n) \left| a_{pq}^{(k)} \right|^2$ et $\|E_{k+1}\|_s^2 \leq \left(1 - \frac{2}{n^2 - n} \right) \|E_k\|_s^2$. Par récurrence on a alors $\|E_k\|_s^2 \leq \left(1 - \frac{2}{n^2 - n} \right)^k \|E_0\|_s^2$. Il en résulte que $\lim_{k \rightarrow +\infty} \|E_k\|_s = 0$ pour $n \geq 3$ et $\lim_{k \rightarrow +\infty} E_k = 0$. $\square$

Pour montrer que la suite des matrices diagonales $(D_k)_{k \in \mathbb{N}}$ est convergente on a besoin du résultat suivant qui sera aussi utilisé pour calculer les vecteurs propres de la matrice A .

Lemme 6.4 Soit $(x_k)_{k \in \mathbb{N}}$ une suite dans un espace vectoriel normé de dimension finie $(E, \|\cdot\|)$. Cette suite converge dans $(E, \|\cdot\|)$ si et seulement si elle est bornée, n'admet qu'un nombre fini de valeurs d'adhérence et $\lim_{n \rightarrow +\infty} \|x_{k+1} - x_k\| = 0$.

Preuve. Voir [18]. $\square$

Théorème 6.3.

Avec les notations qui précèdent, la suite $(A_k)_{k \in \mathbb{N}}$ converge vers une matrice diagonale D_σ de termes diagonaux $d_i^\sigma = \lambda_{\sigma(i)}$ pour $i = 1, \dots, n$, où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de A et σ une permutation de $\{1, \dots, n\}$.

Preuve. D'après le lemme 6.3, il est équivalent de montrer que la suite $(D_k)_{k \in \mathbb{N}}$ converge vers la matrice D_σ .

Avec $\|D_k\|_s \leq \|A_k\|_s = \|A\|_s$, on déduit que la suite $(D_k)_{k \in \mathbb{N}}$ est bornée.

Avec

$$a_{ii}^{(k+1)} - a_{ii}^{(k)} = \begin{cases} 0 & \text{si } i \neq p, i \neq q \\ \tan(\theta_k) a_{pq}^{(k)}, & \text{si } i = p \\ -\tan(\theta_k) a_{pq}^{(k)}, & \text{si } i = q \end{cases}$$

$|\tan(\theta_k)| \leq 1$ et $\left| a_{pq}^{(k)} \right| \leq \|E_k\|_s \xrightarrow{k \rightarrow +\infty} 0$, on déduit que $\lim_{k \rightarrow +\infty} (D_{k+1} - D_k) = 0$.

Pour montrer que la suite $(D_k)_{k \in \mathbb{N}}$ est convergente il nous reste donc à montrer qu'elle n'a qu'un nombre fini de valeurs d'adhérence.

Soit $(D_{\varphi(k)})_{k \in \mathbb{N}}$ une suite extraite de $(D_k)_{k \in \mathbb{N}}$ qui converge vers une matrice D diagonale. Avec $\lim_{k \rightarrow +\infty} E_{\varphi(k)} = 0$, on déduit que $\lim_{k \rightarrow +\infty} A_{\varphi(k)} = D$. Si $P_k(X)$ désigne le polynôme caractéristique de $A_{\varphi(k)}$, on a $\lim_{k \rightarrow +\infty} P_k(X) = \det(XI_n - D)$ avec $P_k(X) = \det(XI_n - A)$ puisque A et $A_{\varphi(k)}$ sont semblables. On a donc montré que D et A ont le même polynôme caractéristique, donc les mêmes valeurs propres. La matrice D étant diagonale les termes diagonaux de D sont donnés par $d_i^\sigma = \lambda_{\sigma(i)}$ pour tout $i = 1, 2, \dots, n$, où σ est une permutation de $\{1, \dots, n\}$. Le nombre de ces permutations étant fini, on déduit qu'il n'y a qu'un nombre fini de valeurs d'adhérence. $\square$

Pour calculer une famille de vecteurs propres de A , on désigne par $(P_k)_{k \in \mathbb{N}}$ la suite de matrices orthogonales définie par $P_k = R(\theta_0) R(\theta_1) \cdots R(\theta_k)$ pour tout $k \geq 0$.

On suppose que toutes les valeurs propres de A sont distinctes. On désigne par D la matrice diagonale de termes diagonaux $d_i = \lambda_i$ pour $i = 1, \dots, n$ et par $P = (C_1, \dots, C_n)$ une matrice orthogonale telle que $D = P^{-1}AP$, où C_j désigne la colonne numéro j de P .

Théorème 6.4.

Avec les notations ci-dessus, on a :

1. $\lim_{k \rightarrow +\infty} \theta_k = 0$ et $\lim_{k \rightarrow +\infty} R_k = I_n$.
2. La suite $(P_k)_{k \in \mathbb{N}}$ converge vers une matrice $P_\sigma = (\pm C_{\sigma(1)}, \dots, \pm C_{\sigma(n)})$ où σ est une permutation de $\{1, \dots, n\}$. Le vecteur $C_{\sigma(i)}$ étant un vecteur propre associé à la valeur propre $\lambda_{\sigma(i)}$.

Preuve.

1. Avec $\lim_{k \rightarrow +\infty} D_k = D_\sigma$ et l'hypothèse que les valeurs propres de A sont deux à deux distinctes, on déduit qu'il existe un entier k_0 tel que :

$$\forall k \geq k_0, \left| a_{pq}^{(k)} - a_{qq}^{(k)} \right| > \frac{1}{2} \min_{1 \leq i \neq j \leq n} |\lambda_i - \lambda_j| > 0$$

Des inégalités $\left| a_{pq}^{(k)} \right| \leq \|E_k\|_s$, avec $\lim_{k \rightarrow +\infty} \|E_k\|_s = 0$ (lemme 6.3), on dé-

duit que l'on a $\lim_{k \rightarrow +\infty} |b_k| = \lim_{k \rightarrow +\infty} \left| \frac{a_{pp}^{(k)} - a_{qq}^{(k)}}{2a_{pq}^{(k)}} \right| = +\infty$, ce qui implique que

$\lim_{k \rightarrow +\infty} \tan(\theta_k) = \lim_{k \rightarrow +\infty} \frac{\text{signe}(b_k)}{\sqrt{b_k^2 + 1} + |b_k|} = 0$. En conclusion, on a $\lim_{k \rightarrow +\infty} \theta_k = 0$ et $\lim_{k \rightarrow +\infty} R_k = I_n$.

2. On a $\|P_k\|_2 = 1$ (P_k est orthogonale) et, avec $P_{k+1} - P_k = P_k(R(\theta_{k+1}) - I_n)$ on déduit que $\lim_{k \rightarrow +\infty} (P_{k+1} - P_k) = 0$. Il nous reste donc à montrer que la suite $(P_k)_{k \in \mathbb{N}}$ n'a qu'un nombre fini de valeurs d'adhérence. Soit $(P_{\varphi(k)})_{k \in \mathbb{N}}$ une suite

extraite de $(P_k)_{k \in \mathbb{N}}$ qui converge vers une matrice orthogonale Q . Avec l'égalité $A_{\varphi(k)} = {}^t P_{\varphi(k)} A P_{\varphi(k)}$, on déduit que ${}^t P_{\sigma} A P_{\sigma} = D_{\sigma} = \lim_{k \rightarrow +\infty} A_{\varphi(k)} = {}^t Q A Q$. Les valeurs propres de A étant deux à deux distinctes, on a nécessairement $Q = (\pm C_{\sigma(1)}, \dots, \pm C_{\sigma(n)})$. D'où le résultat. $\square$

D'un point de vue numérique, le calcul du maximum des termes non diagonaux de A_k à chaque étape n'est pas intéressant car il augmente le temps de calcul. On préfère procéder de la manière suivante : à l'étape k du calcul, la matrice A_{k-1} étant construite on construit la matrice A_k en effectuant $\frac{n(n-1)}{2}$ transformations de Jacobi en prenant pour valeurs successives de (p, q) les valeurs $(1, 2), (1, 3), \dots, (1, n); (2, 3), \dots, (2, n); \dots; (n-1, n)$. Un tel calcul est appelé un balayage. En notant $S_k = \sum_{1 \leq i < j \leq n} |a_{ij}^{(k)}|$, on arrête les itérations quand $S_k < \varepsilon$, où ε est une précision donnée. On peut alors montrer que la convergence devient très rapidement quadratique.

En pratique on n'effectuera pas de transformation de Jacobi si $|a_{pq}^{(k)}| < \sigma$, où σ est un seuil donné. Usuellement, on prend la valeur $\sigma = \frac{S_k}{5n^2}$ seulement pour les trois premiers balayages, puis à partir du quatrième, on décide que $a_{pq}^{(k)} = 0$ si $|a_{pq}^{(k)}| < \varepsilon' |a_{pp}^{(k)}|$ et $|a_{pq}^{(k)}| < \varepsilon' |a_{qq}^{(k)}|$ avec ε' assez petit.

Le calcul de la matrice P_{σ} , donnant les vecteurs propres de A , se fait avec les relations :

$$\begin{cases} p_{ij}^{(k+1)} = p_{ij}^{(k)} & \text{si } j \neq p, j \neq q \\ p_{ip}^{(k+1)} = p_{ip}^{(k)} + s \left(p_{iq}^{(k)} - \tau p_{ip}^{(k)} \right) \\ p_{iq}^{(k+1)} = p_{iq}^{(k)} - s \left(p_{ip}^{(k)} + \tau p_{iq}^{(k)} \right) \end{cases} \quad (i = 1, 2, \dots, n)$$

où $\tau = \tan\left(\frac{\theta_k}{2}\right) = \frac{s}{1+c}$.

6.4 La méthode de Givens et Householder

Soit A une matrice symétrique réelle d'ordre $n \geq 3$. Au paragraphe 2.6, on a décrit l'algorithme de Householder qui permet de calculer effectivement une matrice T tridiagonale semblable à la matrice A . Le calcul des valeurs propres de T nous donne alors celles de A . La méthode de Givens (ou de bisection) nous fournit, sous certaines conditions, un algorithme dichotomique de calcul des valeurs propres d'une matrice tridiagonale symétrique. La méthode de la puissance inverse permet ensuite d'obtenir des vecteurs propres associés.

Pour décrire cette méthode, il sera plus commode de définir le polynôme caractéristique par $\det(A - X I_n)$.

Dans ce paragraphe on se donne une matrice :

$$A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & b_{n-1} & a_n \end{pmatrix}$$

à coefficients réels, tridiagonale symétrique et supposée irréductible, c'est-à-dire telle que $b_i \neq 0$ pour tout i compris entre 1 et $n-1$. Une telle matrice est de type Hessenberg irréductible et on sait alors qu'elle admet n valeurs propres réelles simples. (exercice 6.2). De plus avec $\rho(A) \leq \|A\|_\infty$ où $\rho(A)$ désigne le rayon spectral de A , on déduit que toutes les valeurs propres de la matrice A sont dans l'intervalle $[a_0, b_0] = [-\|A\|_\infty, \|A\|_\infty]$.

Pour tout $k = 1, \dots, n$, on note $A_k = ((a_{ij}))_{1 \leq i, j \leq k}$ la k -ième sous-matrice principale de A et $P_k(X) = \det(A_k - XI_k)$ son polynôme caractéristique de coefficient dominant $(-1)^k$. Chaque polynôme P_k admet donc k racines réelles distinctes $\lambda_{1,k} < \dots < \lambda_{k,k}$ (la matrice A_k vérifie les mêmes propriétés que la matrice A). Les valeurs propres de A sont les $\lambda_j = \lambda_{j,n}$ pour $j = 1, 2, \dots, n$.

En développant $P_k(X)$ suivant la dernière ligne on obtient la relation de récurrence :

$$\begin{cases} P_0(X) = 1, & P_1(X) = a_1 - X \\ P_k(X) = (a_k - X)P_{k-1}(X) - b_{k-1}^2 P_{k-2}(X) & (k = 2, \dots, n) \end{cases}$$

Lemme 6.5 Pour $k = 2, \dots, n$, les racines de P_{k-1} séparent celles de P_k dans le sens où :

$$\begin{cases} \lambda_{1,k} < \lambda_{1,k-1} \\ \lambda_{j,k-1} < \lambda_{j+1,k} < \lambda_{j+1,k-1} & (1 \leq j \leq k-2) \\ \lambda_{k-1,k-1} < \lambda_{k,k} \end{cases}$$

Preuve. Pour $k = 1$, on a $P_1(X) = a_1 - X$ qui admet une racine réelle $\lambda_{1,1} = a_1$. Pour $k = 2$, on a $P_2(X) = (a_1 - X)(a_2 - X) - b_1^2$, avec $\lim_{\lambda \rightarrow -\infty} P_2(\lambda) = +\infty$, $P_2(\lambda_{1,1}) = P_2(a_1) = -b_1^2 < 0$ et $\lim_{\lambda \rightarrow +\infty} P_2(\lambda) = +\infty$. Avec le théorème des valeurs intermédiaires, on déduit donc que P_2 admet une racine réelle $\lambda_{1,2} < \lambda_{1,1}$ et une racine réelle $\lambda_{2,2} > \lambda_{1,1}$. On a donc montré que la racine $\lambda_{1,1}$ de P_1 sépare les deux racines de P_2 , avec $P_0(\lambda_{1,1}) = 1 > 0$. Ce qui est schématisé par la figure 6.1.

Supposons le résultat acquis pour $p = 2, \dots, k-1$, avec $P_{k-2}(\lambda_{j,k-1})$ du signe de $(-1)^{j-1}$ pour tout $j = 1, \dots, k-1$. Avec $\lim_{\lambda \rightarrow -\infty} P_k(\lambda) = \lim_{\lambda \rightarrow -\infty} (-1)^k \lambda^k = +\infty$ et $P_k(\lambda_{1,k-1}) = -b_{k-1}^2 P_{k-2}(\lambda_{1,k-1}) < 0$, on déduit qu'il existe une racine de P_k , $\lambda_{1,k} < \lambda_{1,k-1}$ telle que $P_{k-1}(\lambda_{1,k}) > 0$, puisque $P_{k-1}(\lambda) > 0$ sur $]-\infty, \lambda_{1,k-1}[$.

Pour $j = 1, \dots, k-2$, on a $P_k(\lambda_{j,k-1}) = -b_{k-1}^2 P_{k-2}(\lambda_{j,k-1})$ du signe de $(-1)^j$ et $P_k(\lambda_{j+1,k-1}) = -b_{k-1}^2 P_{k-2}(\lambda_{j+1,k-1})$ du signe de $(-1)^{j+1}$. On en déduit alors que P_k admet une racine $\lambda_{j+1,k} \in]\lambda_{j,k-1}, \lambda_{j+1,k-1}[$ (figure 6.2). Pour j impair,

on a $P_{k-1}(\lambda_{j+1,k}) < 0$ et pour j pair, on a $P_{k-1}(\lambda_{j+1,k}) > 0$, c'est-à-dire que $P_{k-1}(\lambda_{j+1,k})$ est du signe de $(-1)^j$.

Enfin $P_k(\lambda_{k-1,k-1}) = -b_{k-1}^2 P_{k-2}(\lambda_{k-1,k-1})$ est du signe de $(-1)^{k-1}$ et on a $\lim_{\lambda \rightarrow +\infty} P_k(\lambda) = (-1)^k \infty$, donc on a une racine de P_k , $\lambda_{k,k} \in]\lambda_{k-1,k-1}, +\infty[$ telle que $P_{k-1}(\lambda_{k,k})$ soit du signe de $(-1)^{k-1}$. D'où le résultat. $\square$

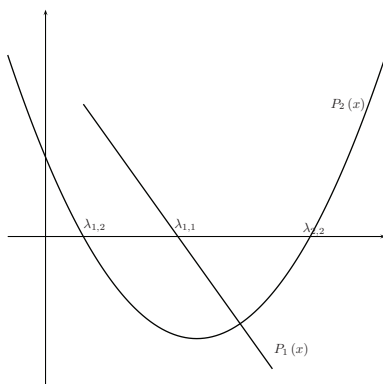


FIGURE 6.1 –

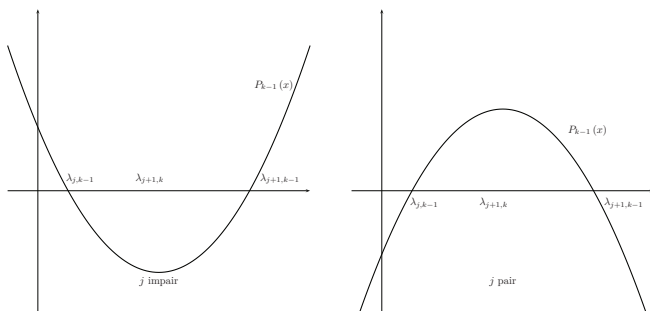


FIGURE 6.2 –

On note, pour tout réel λ et tout entier k compris entre 0 et n :

$$s_k(\lambda) = \begin{cases} \text{signe}(P_k(\lambda)) & \text{si } P_k(\lambda) \neq 0 \\ s_{k-1}(\lambda) & \text{si } P_k(\lambda) = 0 \end{cases}$$

et $N_k(\lambda)$ désigne le nombre de changements de signes entre deux termes consécutifs de la suite $(s_0(\lambda), s_1(\lambda), \dots, s_k(\lambda))$.

En remarquant que $P_0(\lambda) \neq 0$ et que pour $k = 1, \dots, n$, on a $P_{k-1}(\lambda) \neq 0$ si $P_k(\lambda) = 0$, on déduit que la fonction s_k est bien définie et à valeurs dans $\{-1, 1\}$ avec $s_0(\lambda) = 1$ pour tout réel λ .

Lemme 6.6 Avec les notations qui précèdent, on a

$$s_1(\lambda) = \begin{cases} 1 & \text{si } \lambda \in]-\infty, \lambda_{1,1}] \\ -1 & \text{si } \lambda \in]\lambda_{1,1}, +\infty[\end{cases}$$

et pour tout $k = 2, \dots, n$:

$$s_k(\lambda) = \begin{cases} 1 & \text{si } \lambda \in]-\infty, \lambda_{1,k}] \\ (-1)^j & \text{si } \lambda \in]\lambda_{j,k}, \lambda_{j+1,k}] \quad (1 \leq j \leq k-1) \\ (-1)^k & \text{si } \lambda \in]\lambda_{k,k}, +\infty[\end{cases}$$

Preuve. Le résultat sur s_1 est évident. Avec $\lim_{\lambda \rightarrow -\infty} P_k(\lambda) = +\infty$ on déduit que $P_k(\lambda) > 0$ sur $]-\infty, \lambda_{1,k}[$ pour tout entier k compris entre 1 et n . Et avec $\lambda_{1,k} < \lambda_{1,k-1}$, on conclut que $s_k(\lambda) = 1$ sur $]-\infty, \lambda_{1,k}]$. Pour tout entier k compris entre 2 et n et tout entier j compris entre 1 et $k-1$, P_k est du signe de $(-1)^j$ sur $]\lambda_{j,k}, \lambda_{j+1,k}[$. En considérant que $\lambda_{j+1,k} \in]\lambda_{j,k-1}, \lambda_{j+1,k-1}[$, on déduit que $s_k(\lambda) = (-1)^j$ sur $]\lambda_{j,k}, \lambda_{j+1,k}]$ pour tout $k = 3, \dots, n$ et tout $j = 1, \dots, k-1$. Pour $k = 2$, le résultat est encore valable puisque $s_2(\lambda_{2,2}) = s_1(\lambda_{2,2}) = -1$. Enfin, avec $\lim_{\lambda \rightarrow +\infty} P_k(\lambda) = (-1)^k \infty$, on déduit que $s_k(\lambda) = (-1)^k$ sur $]\lambda_{k,k}, +\infty[$. $\square$

Lemme 6.7 Pour tout réel λ et tout entier k compris entre 1 et n , $N_k(\lambda)$ est égal au nombre de racines de P_k qui sont strictement inférieures à λ .

Preuve. On montre le résultat par récurrence sur $k \geq 1$. Pour $k = 1$, on a :

$$(s_0(\lambda), s_1(\lambda)) = \begin{cases} (1, 1) & \text{si } \lambda \leq a_1 = \lambda_{1,1} \\ (1, -1) & \text{si } \lambda > a_1 = \lambda_{1,1} \end{cases}, \quad N_1(\lambda) = \begin{cases} 0 & \text{si } \lambda \leq \lambda_{1,1} \\ 1 & \text{si } \lambda > \lambda_{1,1} \end{cases}$$

Supposons le résultat acquis pour $k-1$ compris entre 1 et $n-1$ et soit $\lambda \in \mathbb{R}$. Si $\lambda \leq \lambda_{1,k} < \lambda_{1,j}$ ($j = 1, \dots, k-1$), on a alors $P_j(\lambda) > 0$ pour tout entier j compris entre 1 et k et $(s_0(\lambda), s_1(\lambda), \dots, s_k(\lambda)) = (1, 1, \dots, 1)$, c'est-à-dire que $N_k(\lambda) = 0$ et c'est bien le nombre de racines de P_k strictement inférieures à λ .

Si $\lambda \in]\lambda_{j,k}, \lambda_{j+1,k}]$ pour un entier $j \in \{1, \dots, k-1\}$, on distingue alors deux cas en fonction de la position de λ par rapport à $\lambda_{j,k-1} \in]\lambda_{j,k}, \lambda_{j+1,k}]$.

Si $\lambda \in]\lambda_{j,k}, \lambda_{j,k-1}]$, on a alors $j-1$ racines de P_{k-1} strictement inférieures à λ , c'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j-1$. En tenant compte de $s_{k-1}(\lambda) = (-1)^{j-1}$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ .

Si $\lambda \in]\lambda_{j,k-1}, \lambda_{j+1,k}]$, on a alors j racines de P_{k-1} strictement inférieures à λ , c'est-à-dire avec l'hypothèse de récurrence que $N_{k-1}(\lambda) = j$. En tenant compte de $s_{k-1}(\lambda) = (-1)^j$ et $s_k(\lambda) = (-1)^j$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) = j$ et c'est bien le nombre de racines de P_k strictement inférieures à λ .

Enfin si $\lambda > \lambda_{k,k} > \lambda_{k-1,k-1}$, avec l'hypothèse de récurrence on déduit alors que $N_{k-1}(\lambda) = k-1$. Avec $s_{k-1}(\lambda) = (-1)^{k-1}$ et $s_k(\lambda) = (-1)^k$, on déduit que $N_k(\lambda) = N_{k-1}(\lambda) + 1 = k$ et c'est bien le nombre de racines de P_k strictement inférieures à λ . $\square$

Pour tout réel λ , on désigne par $N(\lambda)$ le nombre de racines de P_n qui sont strictement inférieures à λ et $c_0 = \frac{a_0 + b_0}{2}$ est le milieu de $[a_0, b_0] = [-\|A\|_\infty, \|A\|_\infty]$.

Lemme 6.8 Soit i un entier compris entre 1 et n . Si $N(c_0) \geq i$, on a alors $\lambda_i \in [a_0, c_0]$, sinon $\lambda_i \in [c_0, b_0]$.

Preuve. Du lemme 6.7, on déduit que si $\lambda_i \in [c_0, b_0]$, alors $N(c_0) \leq i - 1$ et si $\lambda_i \in [a_0, c_0]$, alors $N(c_0) \geq i$. D'où le résultat. $\square$

Pour i fixé entre 1 et $n - 1$, on note $[a_1, b_1]$ la moitié de l'intervalle $[a_0, b_0]$ qui contient la valeur propre λ_i et par récurrence on construit une suite $([a_k, b_k])_{k \geq 1}$ d'intervalles emboîtés telle que :

$$\forall k \geq 1, \begin{cases} b_k - a_k = \frac{b_0 - a_0}{2^k} \\ \lambda_i \in [a_k, b_k] \end{cases}$$

Il suffit en effet de reprendre le raisonnement du lemme 6.8 avec chaque intervalle $[a_k, b_k]$.

On a alors $\lambda_i = \lim_{k \rightarrow +\infty} a_k = \lim_{k \rightarrow +\infty} b_k = \lim_{k \rightarrow +\infty} c_k$, où on a noté c_k le milieu de l'intervalle $[a_k, b_k]$ (théorème des suites adjacentes).

Le calcul des vecteurs propres associés peut se faire ensuite en utilisant la méthode de la puissance inverse.

Soient $i \in \{1, \dots, n\}$ fixé et λ un réel tel que $0 < |\lambda - \lambda_i| < \inf_{1 \leq j \neq i \leq n} |\lambda - \lambda_j|$.

On se donne un vecteur x_0 non orthogonal au sous espace propre associé à λ_i et on définit la suite de vecteurs $(x_k)_{k \in \mathbb{N}}$ par $(A - \lambda I_n)x_{k+1} = x_k$ pour tout $k \geq 0$.

Lemme 6.9 Avec les notations qui précèdent, la suite $\left(\frac{\text{signe}(\lambda_i - \lambda)^k}{\|x_k\|} x_k \right)_{k \in \mathbb{N}}$ converge vers un vecteur propre non nul associé à la valeur propre λ_i .

Preuve. On remarque que si $x_{k+1} = 0$, on a alors $x_k = 0$. On déduit donc par récurrence que si $x_0 \neq 0$, on a alors $x_k \neq 0$ pour tout entier k .

Soit $(f_k)_{1 \leq k \leq n}$ une base orthonormée de vecteurs propres avec $Af_j = \lambda_j f_j$ pour tout j compris entre 1 et n . Pour tout entier k , on note $x_k = \sum_{j=1}^n x_{j,k} f_j$.

Avec $(A - \lambda I_n)x_{k+1} = x_k$, on déduit que $x_{j,k+1} = \frac{1}{\lambda_j - \lambda} x_{j,k}$ et par récurrence $x_{j,k} = \frac{1}{(\lambda_j - \lambda)^k} x_{j,0}$. On peut alors écrire que $x_k = \frac{1}{(\lambda_i - \lambda)^k} y_k$, où y_k est le vecteur de composantes $y_{j,k} = \left(\frac{\lambda_i - \lambda}{\lambda_j - \lambda} \right)^k x_{j,0}$. On a alors :

$$\lim_{k \rightarrow +\infty} y_{j,k} = \begin{cases} 0 & \text{si } j \neq i \\ x_{i,0} & \text{si } j = i \end{cases}$$

avec $x_{i,0} \neq 0$. Donc $\lim_{k \rightarrow +\infty} y_k = x_{i,0} f_i$ est vecteur propre non nul associé à λ_i . On

en déduit que $\lim_{k \rightarrow +\infty} \frac{\text{signe}(\lambda_i - \lambda)^k}{\|x_k\|} x_k = \lim_{k \rightarrow +\infty} \frac{1}{\|y_k\|} y_k = \text{signe}(x_{i,0}) f_i$. $\square$

Si A n'est pas irréductible, on peut alors la découper en blocs de matrices irréductibles et l'algorithme de Givens permet encore de calculer les valeurs propres.

6.5 Exercices

Exercice 6.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ et λ une valeur propre de A d'ordre $p \geq 1$. Montrer que les sous-espaces vectoriels $N_\lambda = \ker(A - \lambda I_n)^p$ et $M_\lambda = \operatorname{Im}(A - \lambda I_n)^p$ sont stables par A , supplémentaires dans $\mathbb{K}^n$ et que N_λ est de dimension p .

Solution. Avec la commutativité de $\mathbb{K}[A]$ on vérifie facilement que les sous-espaces vectoriels N_λ et M_λ sont stables par A . On a vu (théorème 1.12) que N_λ est de dimension p et avec le théorème du rang, on déduit que M_λ est de dimension $n - p$. Enfin en notant $\chi_A(X) = (X - \lambda)^p Q(X)$ le polynôme caractéristique de A , le polynôme Q étant premier avec $(X - \lambda)^p$, on peut trouver des polynômes U et V tels que $U(X)(X - \lambda)^p + V(X)Q(X) = 1$ (théorème de Bézout) et tout vecteur x de $\mathbb{K}^n$ s'écrit $x = U(A)(A - \lambda I_n)^p(x) + V(A)Q(A)(x)$ avec $U(A)(A - \lambda I_n)^p(x)$ dans M_λ et $V(A)Q(A)(x)$ dans N_λ . On a donc $\mathbb{K}^n = N_\lambda + M_\lambda$ et la somme est directe en considérant les dimensions.

Exercice 6.2. On appelle matrice de Hessenberg une matrice A à coefficients complexes qui vérifie $a_{ij} = 0$ pour $j < i - 1$. On dit que A est irréductible si $a_{i,i-1} \neq 0$ pour tout $i = 2, \dots, n$.

1. Montrer que si A est une matrice de Hessenberg irréductible, alors pour toute valeur propre de A , l'espace propre associé est de dimension 1.
2. En déduire que les valeurs propres d'une matrice de Hessenberg irréductible sont simples si, et seulement si, la matrice est diagonalisable.

$$3. \text{ Soit } A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ b_1 & a_2 & b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & b_{n-1} & a_n \end{pmatrix} \quad \text{une matrice à coefficients}$$

réels, tridiagonale symétrique et irréductible.

- (a) Montrer que les valeurs propres de A sont simples.
- (b) Décrire un algorithme de calcul de l'espace propre associé à une valeur propre de A .

$$4. \text{ Soit } A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ c_2 & a_2 & b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & c_{n-1} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & c_n & a_n \end{pmatrix} \quad \text{une matrice tridiagonale à}$$

coefficients complexes.

- (a) Donner un algorithme de calcul du polynôme caractéristique de A .

(b) Montrer que A admet les mêmes valeurs propres que la matrice :

$$B = \begin{pmatrix} a_1 & c_2 b_1 & 0 & \cdots & 0 \\ 1 & a_2 & c_3 b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & a_{n-1} & c_n b_{n-1} \\ 0 & \cdots & 0 & 1 & a_n \end{pmatrix}$$

(c) Montrer que si $a_i \in \mathbb{R}$ pour tout $i = 1, \dots, n$ et $c_{i+1} b_i \in \mathbb{R}_+^*$ pour tout $i = 1, \dots, n-1$, A admet alors n valeurs propres réelles simples et est diagonalisable.

(d) Que peut-on dire si $a_i \in \mathbb{R}$ pour tout $i = 1, \dots, n$ et $c_{i+1} b_i \in \mathbb{R}_-^*$ pour tout $i = 1, \dots, n-1$?

Solution.

1. Soient A une matrice de Hessenberg irréductible, λ un complexe et :

$$A_\lambda = A - \lambda I_n = \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & a_{n-2,n} \\ \vdots & \ddots & a_{n-1,n-2} & a_{n-1,n-1} - \lambda & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n-1} & a_{n,n} - \lambda \end{pmatrix}$$

Si B_λ est la matrice extraite de A_λ en supprimant la première ligne et la dernière colonne, soit :

$$B_\lambda = \begin{pmatrix} a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2,n-1} \\ 0 & a_{32} & a_{33} - \lambda & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & a_{n-2,n} \\ \vdots & \ddots & 0 & a_{n-1,n-2} & a_{n-1,n-1} - \lambda \\ 0 & \cdots & 0 & 0 & a_{n,n-1} \end{pmatrix}$$

on a $\det(B_\lambda) = \prod_{i=2}^n a_{i,i-1} \neq 0$. On a donc ainsi montré que :

$$\forall \lambda \in \mathbb{C}, \text{rang}(A_\lambda) \geq n - 1$$

ou encore :

$$\forall \lambda \in \mathbb{C}, \dim(\ker(A - \lambda I_n)) \leq 1$$

En particulier, pour λ valeur propre de A on a $\dim(\ker(A - \lambda I_n)) = 1$.

2. On sait que la matrice A est diagonalisable sur $\mathbb{C}$ si et seulement pour toute valeur propre λ de A , la dimension de $\ker(A - \lambda I_n)$ est égale à la multiplicité

de λ comme racine du polynôme caractéristique de A . Avec ce qui précède on déduit alors que la matrice A est diagonalisable si et seulement toutes ses valeurs propres sont simples.

3.

- (a) Une matrice symétrique réelle a toutes ses valeurs propres réelles et est diagonalisable. Si de plus elle tridiagonale et irréductible alors toutes ses valeurs propres sont simples.
- (b) Soit λ une valeur propre de A et $x \in \mathbb{R}^n \setminus \{0\}$ un vecteur propre associé. On a alors :

$$\begin{cases} a_1 x_1 + b_1 x_2 = \lambda x_1 \\ b_{k-1} x_{k-1} + a_k x_k + b_k x_{k+1} = \lambda x_k \quad (2 \leq k \leq n-1) \\ b_{n-1} x_{n-1} + a_n x_n = \lambda x_n \end{cases}$$

Si $x_n = 0$, avec l'hypothèse $b_i \neq 0$ pour tout $i = 1, \dots, n-1$, on déduit alors que tous les x_i sont nuls. On peut donc prendre $x_n = 1$ et $(x_1, \dots, x_{n-1})$ est solution du système triangulaire supérieur :

$$\begin{cases} b_{k-1} x_{k-1} + (a_k - \lambda) x_k + b_k x_{k+1} = 0 \quad (2 \leq k \leq n-2) \\ b_{n-2} x_{n-2} + (a_{n-1} - \lambda) x_{n-1} = -b_{n-1} \\ b_{n-1} x_{n-1} = \lambda - a_n \end{cases}$$

La solution de ce système peut se calculer avec l'algorithme :

$$\begin{cases} x_{n-1} = \frac{\lambda - a_n}{b_{n-1}} \\ x_{k-1} = -\frac{b_k x_{k+1} + (a_k - \lambda) x_k}{b_{k-1}} \quad (k = n-1, \dots, 2) \end{cases}$$

On retrouve le fait que l'espace propre associé à la valeur propre λ est de dimension 1.

4.

- (a) Si, pour tout $k = 1, \dots, n$, on désigne par A_k la matrice principale d'ordre k de A , alors la suite $(P_k)_{1 \leq k \leq n}$ des polynômes caractéristiques des A_k vérifie la récurrence :

$$\begin{cases} P_0(X) = 1, \quad P_1(\lambda) = a_1 - X \\ P_k(X) = (a_k - X) P_{k-1}(X) - b_{k-1} c_k P_{k-2}(X) \quad (2 \leq k \leq n) \end{cases}$$

- (b) Le polynôme caractéristique de la matrice B s'obtient avec la même récurrence que celui de A en utilisant les mêmes conditions initiales, ces deux polynômes sont donc identiques.
- (c) En utilisant deux fois le résultat précédent, on voit que A admet les mêmes valeurs propres que la matrice :

$$C = \begin{pmatrix} a_1 & \sqrt{c_2 b_1} & 0 & \cdots & 0 \\ \sqrt{c_2 b_1} & a_2 & \sqrt{c_3 b_2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \sqrt{c_{n-1} b_{n-2}} & a_{n-1} & \sqrt{c_n b_{n-1}} \\ 0 & \cdots & 0 & \sqrt{c_n b_{n-1}} & a_n \end{pmatrix}$$

La matrice C est symétrique tridiagonale et irréductible, elle est donc diagonalisable avec n valeurs propres réelles simples. La matrice A admet donc n valeurs propres réelles simples et est diagonalisable.

(d) Soit $A = \begin{pmatrix} a & -b_1 & 0 \\ 1 & 0 & -b_2 \\ 0 & 1 & a \end{pmatrix}$, avec $a \in \mathbb{R}$, $b_i > 0$ pour $i = 1, 2$. Son

polynôme caractéristique est $P(X) = (a - X)(X^2 - aX + b_2 + b_1)$. Pour $a^2 - 4(b_1 + b_2) = 0$ la matrice A admet une valeur propre double et donc n'est pas diagonalisable.

Exercice 6.3. Soit $A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ c_2 & a_2 & b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & c_{n-1} & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & c_n & a_n \end{pmatrix}$ une matrice

tridiagonale à coefficients complexes telle que $a_i \in \mathbb{R}$ pour tout $i = 1, \dots, n$ et $c_{i+1}b_i \in \mathbb{R}_+^*$ pour tout $i = 1, 2, \dots, n-1$.

1. Montrer que A admet les mêmes valeurs propres que la matrice :

$$B = \begin{pmatrix} a_1 & c_2b_1 & 0 & \cdots & 0 \\ 1 & a_2 & c_3b_2 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & 1 & a_{n-1} & c_nb_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix}$$

Pour ce qui suit, on suppose que $A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ 1 & a_2 & b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & a_{n-1} & b_{n-1} \\ 0 & \cdots & 0 & 1 & a_n \end{pmatrix}$

avec $a_i, b_i \in \mathbb{R}$, $b_i > 0$. On note $D = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$. On

définit alors une homotopie de D vers A en posant :

$$\forall t \in [0, 1], A(t) = (1 - t)D + tA$$

2. Calculer les valeurs propres de D .

3. Montrer que, pour tout t dans $[0, 1]$, $A(t)$ admet n valeurs propres réelles simples.
4. Montrer qu'il existe une fonction $f = (f_1, \dots, f_n) : [0, 1] \rightarrow \mathbb{R}^n$ continûment dérivable telle que le polynôme caractéristique de $A(t)$ s'écrive :

$$\forall t \in [0, 1], P(X, t) = \prod_{j=1}^n (X - f_j(t))$$

Pour tout t dans $[0, 1]$ et $j = 1, \dots, n$, on pose $\varphi_j(t) = P(f_j(t), t)$ et on a alors $\varphi_j(t) = 0$ sur $[0, 1]$.

5. Montrer que pour tout $j = 1, \dots, n$, f_j est solution sur $[0, 1]$ d'une équation différentielle avec condition initiale.
6. Donner un algorithme de calcul de toutes les valeurs propres de la matrice A .

Solution.

1. Voir l'exercice 6.2.
2. En utilisant les résultats de l'exercice 1.9, on déduit que les valeurs propres de la matrice D sont données par $\lambda_k = 2 \cos(k\theta)$ ($1 \leq k \leq n$) avec $\theta = \frac{\pi}{n+1}$.
3. Pour tout t dans $[0, 1]$, on a :

$$A(t) = \begin{pmatrix} ta_1 & 1-t+tb_1 & 0 & \cdots & 0 \\ 1 & ta_2 & 1-t+tb_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & ta_{n-1} & 1-t+tb_{n-1} \\ 0 & \cdots & 0 & 1 & ta_n \end{pmatrix}$$

et cette matrice admet n valeurs propres réelles simples puisqu'elle est irréductible avec $1-t+tb_i > 0$ pour tout $t \in [0, 1]$ et tout $i = 1, 2, \dots, n$ (exercice 6.2).

4. Soit $t_0 \in [0, 1]$ et λ_0 une valeur propre de $A(t_0)$. On a $P(\lambda_0, t_0) = 0$ et $\frac{\partial P}{\partial \lambda}(\lambda_0, t_0) \neq 0$ puisque λ_0 est racine simple de $P(\cdot, t_0)$. Le théorème des fonctions implicites nous dit alors qu'il existe un voisinage ouvert $\mathcal{V}_0$ de t_0 dans $[0, 1]$ et une fonction $f = (f_1, \dots, f_n) : \mathcal{V}_0 \rightarrow \mathbb{R}^n$ continûment dérivable tels que :

$$\forall t \in \mathcal{V}_0, P(X, t) = \prod_{k=1}^n (X - f_k(t))$$

Les valeurs propres de $A(t)$ étant simples, on peut supposer que $f_1(t) < \dots < f_n(t)$. Avec cette condition la fonction f est unique. L'intervalle $[0, 1]$ étant compact, on peut trouver une partition du type $[0, 1] = [a_0, b_0[\cup [a_1, b_1[\cup \dots \cup [a_p, b_p]$, où $0 = a_0 < b_0 = a_1 < b_1 = a_2 < \dots < b_{p-1} = a_p < b_p = 1$, telle

pour tout $k = 0, \dots, p$, il existe une unique fonction continûment dérivable $f_k = (f_{1,k}, \dots, f_{n,k}) : [a_k, b_k] \rightarrow \mathbb{R}^n$ avec :

$$\forall t \in [a_k, b_k], \begin{cases} P(X, t) = \prod_{j=1}^n (X - f_{j,k}(t)) \\ f_{1,k}(t) < f_{2,k}(t) < \dots < f_{n,k}(t) \end{cases}$$

La fonction f définie sur $[0, 1]$ par $f(t) = (f_{1,k}(t), \dots, f_{n,k}(t))$ pour $t \in [a_k, b_k]$ est alors continûment dérivable et on a $P(X, t) = \prod_{j=1}^n (X - f_j(t))$ pour tout $t \in [0, 1]$.

5. En dérivant la relation $\varphi_j(t) = 0$ par rapport à t , pour tout $j = 1, \dots, n$, on déduit que f_j est solution du problème de Cauchy :

$$\begin{cases} f_j(0) = a_j \\ f'_j(t) = -\frac{\frac{\partial P}{\partial t}(f_j(t), t)}{\frac{\partial P}{\partial \lambda}(f_j(t), t)} \quad (t \in [0, 1]) \end{cases}$$

La simplicité des valeurs propres de $A(t)$ nous garantit que $\frac{\partial P}{\partial \lambda}(f_j(t), t) \neq 0$ sur $[0, 1]$.

6. Le calcul du polynôme caractéristique $P(X, t)$ se fait en utilisant la récurrence :

$$\begin{cases} P_0(X, t) = 1, \quad P_1(X, t) = ta_1 - X \\ P_k(X, t) = (ta_k - X)P_{k-1}(X, t) - (1 - t + tb_{k-1})P_{k-2}(X, t) \quad (2 \leq k \leq n) \end{cases}$$

En dérivant par rapport à X et à t , on obtient l'algorithme suivant de calcul des dérivées partielles de P :

$$\begin{cases} \frac{\partial P_0}{\partial X}(X, t) = 0, \quad \frac{\partial P_0}{\partial t}(X, t) = 0 \\ \frac{\partial P_1}{\partial X}(X, t) = -1, \quad \frac{\partial P_1}{\partial t}(X, t) = a_1 \\ \frac{\partial P_k}{\partial X}(X, t) = (ta_k - X) \frac{\partial P_{k-1}}{\partial X}(X, t) - P_{k-1}(X, t) \\ \quad - (1 - t + tb_{k-1}) \frac{\partial P_{k-2}}{\partial X}(X, t) \\ \frac{\partial P_k}{\partial t}(X, t) = (ta_k - X) \frac{\partial P_{k-1}}{\partial t}(X, t) + a_k P_{k-1}(X, t) \\ \quad - (1 - t + tb_{k-1}) \frac{\partial P_{k-2}}{\partial t}(X, t) - (b_{k-1} - 1) P_{k-2}(X, t) \end{cases}$$

En utilisant la méthode Runge-Kutta d'ordre 4, on peut alors calculer des valeurs approchées des valeurs propres $f_j(1)$ de la matrice A .

Chapitre 7

Systèmes différentiels linéaires et exponentielle d'une matrice

Pour $n \in \mathbb{N}^*$, l'espace vectoriel $\mathbb{C}^n$ est muni d'une norme quelconque notée $\|\cdot\|$.

Pour tout intervalle réel $[a, b]$ on désigne par $\mathcal{C}^0([a, b], \mathbb{C}^n)$ l'espace vectoriel des fonctions continues définies sur $[a, b]$ et à valeurs dans $\mathbb{C}^n$. On munit cet espace vectoriel de la norme de la convergence uniforme définie par :

$$\forall y \in \mathcal{C}^0([a, b], \mathbb{C}^n), \|y\|_\infty = \sup_{t \in [a, b]} \|y(t)\|$$

On rappelle que l'espace vectoriel E muni de cette norme est complet (voir [18]).

Étant donné un ouvert Ω de $\mathbb{R} \times \mathbb{C}^n$ et une fonction continue :

$$\begin{aligned} f : \quad \Omega &\rightarrow \mathbb{C}^n \\ (t, y) &\mapsto f(t, y) \end{aligned}$$

on dit qu'une fonction $y : I \rightarrow \mathbb{C}^n$ définie sur un intervalle réel I est solution du système différentiel $y' = f(t, y)$, si cette fonction est dérivable sur I avec :

$$\forall t \in I, (t, y(t)) \in \Omega \text{ et } y'(t) = f(t, y(t))$$

Le problème de Cauchy associé à ce système et aux conditions initiales (t_0, y_0) consiste à rechercher les solutions qui vérifient $y(t_0) = y_0$.

7.1 Systèmes différentiels linéaires à coefficients constants

Pour tout réel strictement positif α , on note $E_\alpha = \mathcal{C}^0([-\alpha, \alpha], \mathbb{C}^n)$ et on munit cet espace vectoriel de la norme de la convergence uniforme $\|\cdot\|_\infty$.

On se donne une matrice non nulle $A \in \mathcal{M}_n(\mathbb{C})$, un vecteur $y_0 \in \mathbb{C}^n$ et on définit l'application $\varphi : E_\alpha \rightarrow E_\alpha$ par :

$$\forall y \in E_\alpha, \forall t \in [-\alpha, \alpha], \varphi(y)(t) = A \int_0^t y(u) du + y_0$$

Lemme 7.1 *Il existe un entier p strictement positif tel que l'application itérée $\varphi^p = \underbrace{\varphi \circ \dots \circ \varphi}_{p \text{ fois}}$ soit strictement contractante.*

Preuve. Pour y, z dans E_α et t dans $[-\alpha, \alpha]$ on a :

$$\begin{aligned} \|\varphi(y)(t) - \varphi(z)(t)\| &= \left\| A \int_0^t (y(u) - z(u)) du \right\| \\ &\leq \|A\| \left| \int_0^t \|y(u) - z(u)\| du \right| \leq \|A\| \|y - z\|_\infty |t| \end{aligned}$$

où $A \mapsto \|A\|$ est la norme matricielle induite par la norme vectorielle $x \mapsto \|x\|$. Par récurrence sur $k \geq 1$, on déduit que :

$$\forall k \geq 1, \forall t \in [-\alpha, \alpha], \|\varphi^k(y)(t) - \varphi^k(z)(t)\| \leq \frac{(\|A\| |t|)^k}{k!} \|y - z\|_\infty$$

En effet, on vient de voir que le résultat est vrai pour $k = 1$ et en le supposant acquis pour $k \geq 1$, on a pour tout $t \in [-\alpha, \alpha]$:

$$\begin{aligned} \|\varphi^{k+1}(y)(t) - \varphi^{k+1}(z)(t)\| &\leq \|A\| \left| \int_0^t \|\varphi^k(y)(u) - \varphi^k(z)(u)\| du \right| \\ &\leq \|A\| \frac{\|A\|^k}{k!} \|y - z\|_\infty \left| \int_0^t |u|^k du \right| \end{aligned}$$

avec $\left| \int_0^t |u|^k du \right| = \int_0^{|t|} u^k du = \frac{|t|^{k+1}}{k+1}$, ce qui donne :

$$\|\varphi^{k+1}(y)(t) - \varphi^{k+1}(z)(t)\| \leq \frac{(\|A\| |t|)^{k+1}}{(k+1)!} \|y - z\|_\infty$$

On en déduit alors que :

$$\forall k \geq 1, \|\varphi^k(y) - \varphi^k(z)\|_\infty \leq \frac{(\|A\| \alpha)^k}{k!} \|y - z\|_\infty$$

et avec $\lim_{k \rightarrow +\infty} \frac{(\|A\| \alpha)^k}{k!} = 0$, il en résulte que, pour $p \geq 1$ assez grand, l'application φ^p est contractante. $\square$

On considère le problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) & (t \in [-\alpha, \alpha]) \\ y(0) = y_0 \end{cases} \quad (7.1)$$

Lemme 7.2 *Si une fonction y définie sur $[-\alpha, \alpha]$ est solution de (7.1), elle est alors de classe C^∞ sur $[-\alpha, \alpha]$.*

Preuve. Soit $y : [-\alpha, \alpha] \rightarrow \mathbb{C}^n$ une solution de (7.1). On vérifie, par récurrence sur $k \geq 1$, que y est de classe $\mathcal{C}^k$ pour tout $k \geq 1$. Le résultat est vrai pour $k = 1$. En le supposant acquis pour $k \geq 1$, on déduit de $y' = Ay$ que y est de classe $\mathcal{C}^{k+1}$ avec $y^{(k+1)} = Ay^{(k)}$. $\square$

Lemme 7.3 Une fonction y définie sur $[-\alpha, \alpha]$ est solution de (7.1) si, et seulement si, c'est un point fixe de l'application φ dans E_α .

Preuve. Soit $y : [-\alpha, \alpha] \rightarrow \mathbb{C}^n$ une solution du système différentiel (7.1), pour tout $t \in [-\alpha, \alpha]$ on a :

$$y(t) = \int_0^t y'(u) du + y(0) = A \int_0^t y(u) du + y_0 = \varphi(y)(t)$$

c'est-à-dire que $y = \varphi(y)$. Réciproquement, si $y \in E_\alpha$ est point fixe de φ , on a alors $y(0) = y_0$ et y est de classe $\mathcal{C}^1$ sur $[-\alpha, \alpha]$ avec $y' = Ay$. $\square$

Le problème de Cauchy (7.1) est donc équivalent au problème de point fixe $y = \varphi(y)$ sur E_α .

Lemme 7.4 Pour tout réel strictement positif α , le problème (7.1) admet une unique solution y définie par $y = \lim_{k \rightarrow +\infty} (y_k)$ où la suite de fonctions $(y_k)_{k \in \mathbb{N}}$ est définie par :

$$\forall k \in \mathbb{N}, \forall t \in [-\alpha, \alpha], y_k(t) = \left(\sum_{j=0}^k \frac{t^j}{j!} A^j \right) y_0$$

la convergence de la suite de fonctions $(y_k)_{k \in \mathbb{N}}$ vers y étant uniforme sur $[-\alpha, \alpha]$.

Preuve. Pour p assez grand, l'itérée φ^p de φ est contractante, donc le théorème du point fixe itéré nous dit que φ admet un unique point fixe donné par $y = \lim_{k \rightarrow +\infty} (y_k)$ dans $(E_\alpha, \|\cdot\|_\infty)$ (c'est-à-dire que la suite de fonctions $(y_k)_{k \in \mathbb{N}}$ converge uniformément vers y sur $[-\alpha, \alpha]$), où $(y_k)_{k \in \mathbb{N}}$ est une suite d'approximations successives de y définie par :

$$\begin{cases} \forall t \in [-\alpha, \alpha], y_0(t) = y_0 \\ \forall k \in \mathbb{N}, y_{k+1} = \varphi(y_k) \end{cases}$$

Par récurrence, on a facilement :

$$\forall k \in \mathbb{N}^*, \forall t \in [-\alpha, \alpha], y_k(t) = \left(\sum_{j=0}^k \frac{t^j}{j!} A^j \right) y_0$$

En effet, pour $k = 1$, on a $y_1(t) = A \int_0^t y_0 du + y_0 = (tA + I_n) y_0$ et en supposant le résultat acquis pour $k \geq 1$:

$$y_{k+1}(t) = A \int_0^t \left(\sum_{j=0}^k \frac{u^j}{j!} A^j \right) y_0 du + y_0 = \left(\sum_{j=0}^{k+1} \frac{t^j}{j!} A^j \right) y_0$$

$\square$

Théorème 7.1.

Le problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(0) = y_0 \end{cases} \quad (7.2)$$

admet une unique solution définie par :

$$\forall t \in \mathbb{R}, y(t) = \lim_{k \rightarrow +\infty} \left(\sum_{j=0}^k \frac{t^j}{j!} A^j \right) y_0$$

Preuve. Pour tout réel $\alpha > 0$, on note y_α la solution de (7.1) sur $[-\alpha, \alpha]$. Pour $0 < \alpha < \beta$, y_α et la restriction de y_β à $[-\alpha, \alpha]$ sont solutions du même problème de Cauchy. Une telle solution étant unique on déduit que les fonctions y_α et y_β coïncident sur $[-\alpha, \alpha]$. On peut donc définir sur $\mathbb{R}$ la fonction y par :

$$\forall t \in \mathbb{R}, y(t) = y_\alpha(t) = \lim_{k \rightarrow +\infty} \left(\sum_{j=0}^k \frac{t^j}{j!} A^j \right) y_0$$

où $\alpha > 0$ est tel que $t \in [-\alpha, \alpha]$. Cette fonction est bien solution de (7.2) sur $\mathbb{R}$. $\square$

Corollaire 7.1 : *L'ensemble S_A des solutions de $y' = Ay$ est un sous espace vectoriel de dimension n de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C}^n)$. Si $(e_k)_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$ et si, pour tout entier k compris entre 1 et n , la fonction y_k est la solution du problème de Cauchy :*

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(0) = e_k \end{cases}$$

la famille $(y_k)_{1 \leq k \leq n}$ est alors une base de S_A .

Preuve. Il est clair que l'ensemble S_A des solutions sur $\mathbb{R}$ du système différentiel $y' = Ay$ est un sous espace vectoriel de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C}^n)$. Le théorème 7.1 se traduit en disant que l'application $y \mapsto y(0)$ réalise un isomorphisme de S_A sur $\mathbb{C}^n$, donc S_A est de dimension n . Pour montrer que la famille $(y_k)_{1 \leq k \leq n}$ est une base de S_A , il

suffit de montrer qu'elle est libre. L'égalité $\sum_{k=1}^n \lambda_k y_k = 0$ équivaut à $\sum_{k=1}^n \lambda_k y_k(t) = 0$

pour tout réel t et $t = 0$ donne $\sum_{k=1}^n \lambda_k e_k = 0$, ce qui entraîne $\lambda_k = 0$ pour tout

$k \in \{1, \dots, n\}$ puisque $(e_k)_{1 \leq k \leq n}$ est libre. La famille $(y_k)_{1 \leq k \leq n}$ est donc une base de S_A . $\square$

7.2 L'exponentielle d'une matrice

Théorème 7.2.

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on peut définir une unique fonction de classe $\mathcal{C}^\infty$, $E_A : \mathbb{R} \rightarrow \mathcal{M}_n(\mathbb{C})$ telle que :

$$\begin{cases} E_A(0) = I_n \\ \forall t \in \mathbb{R}, E'_A(t) = AE_A(t) \end{cases} \quad (7.3)$$

Cette fonction est définie par :

$$\forall t \in \mathbb{R}, E_A(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} A^k$$

Preuve. On note $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ et on définit, la fonction matricielle $E_A = (C_1, \dots, C_n)$, en disant que la colonne numéro j de E_A , notée C_j , est la solution sur $\mathbb{R}$ du problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) \quad (t \in \mathbb{R}) \\ y(0) = e_j \end{cases} \quad (7.4)$$

ce qui se traduit matriciellement par (7.3). Réciproquement si E est solution de (7.3), ses colonnes vérifient alors le système (7.4) et elles sont uniquement déterminées. Pour $j \in \{1, \dots, n\}$, la solution de (7.4) est de classe $\mathcal{C}^\infty$ et donnée

par $C_j(t) = \lim_{p \rightarrow +\infty} \left(\sum_{k=0}^p \frac{t^k}{k!} A^k \right) e_j$, donc la fonction E_A est de classe $\mathcal{C}^\infty$ avec, pour

tout réel t $E_A(t) = \lim_{p \rightarrow +\infty} \left(\sum_{k=0}^p \frac{t^k}{k!} A^k \right) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} A^k$, la convergence étant uniforme sur tout compact de $\mathbb{R}$. $\square$

On note, pour tout réel t , $E_A(t) = e^{tA}$ et la matrice $e^A = \sum_{k=0}^{+\infty} \frac{1}{k!} A^k$ est appelée l'exponentielle de la matrice A . La convergence de cette série est uniforme sur tout compact de $\mathcal{M}_n(\mathbb{C})$. En effet, si on note $\|\cdot\|$ une norme sous-multiplicative sur $\mathcal{M}_n(\mathbb{C})$, pour tout compact K de $\mathcal{M}_n(\mathbb{C})$ il existe une constante $M > 0$ telle que $\|A\| \leq M$ pour toute matrice A dans K et, avec :

$$\sum_{k=0}^{+\infty} \frac{1}{k!} \|A^k\| \leq \sum_{k=0}^{+\infty} \frac{1}{k!} \|A\|^k \leq \sum_{k=0}^{+\infty} \frac{1}{k!} M^k = e^M$$

on déduit que la série de terme général $\frac{1}{k!} A^k$ est uniformément convergente sur K . On en déduit alors que l'application $A \mapsto e^A$ est continue sur $\mathcal{M}_n(\mathbb{C})$.

Le calcul de l'exponentielle d'une matrice est simplifié dans le cas d'une matrice diagonalisable. En effet si la matrice $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable, il existe alors une matrice inversible P et une matrice diagonale $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ telles que

$A = PDP^{-1}$. Avec la continuité du produit matriciel, on déduit que $e^A = Pe^DP^{-1}$ est également diagonalisable avec $e^D = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.

Ce calcul se simplifie également dans le cas des matrices nilpotentes. En effet si la matrice A est nilpotente d'ordre $p \geq 1$, on a alors $e^A = \sum_{k=0}^{p-1} \frac{1}{k!} A^k$.

Pour toute matrice A dans $\mathcal{M}_n(\mathbb{C})$, la matrice e^A est limite d'une suite de polynômes en A , c'est donc un élément de l'adhérence de $\mathbb{C}[A]$ dans $\mathcal{M}_n(\mathbb{C})$. Le sous-espace vectoriel $\mathbb{C}[A]$ étant de dimension finie dans $\mathcal{M}_n(\mathbb{C})$ est fermé et donc confondu avec son adhérence. On déduit donc que e^A est un polynôme en A , c'est-à-dire qu'il existe un polynôme R_A dépendant de A tel que $e^A = R_A(A)$. En fait il n'est pas possible de trouver un polynôme R tel que $e^A = R(A)$ pour toute matrice A . En effet si un tel polynôme existe en notant p son degré on a alors pour toute matrice A et tout réel t , $e^{tA} = \sum_{k=0}^{p+1} \frac{t^k}{k!} A^k + o(t^{p+1}) = \sum_{k=0}^p r_k t^k A^k$ et avec l'unicité du développement limité d'ordre $p+1$ on déduit que nécessairement $A^{p+1} = 0$ pour toute matrice A , ce qui est impossible (prendre par exemple $A = I_n$).

Avec la continuité de l'application $M \mapsto My_0$ et le théorème 7.1, on déduit que la solution du problème (7.1) est la fonction y définie par :

$$\forall t \in \mathbb{R}, y(t) = e^{tA} y_0$$

De manière plus générale, pour tout $t_0 \in \mathbb{R}$, le problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(t_0) = y_0 \end{cases} \quad (7.5)$$

admet pour une unique solution la fonction y définie par :

$$\forall t \in \mathbb{R}, y(t) = e^{(t-t_0)A} y_0$$

En effet, la fonction y est solution de (7.5) si, et seulement si, la fonction z définie par :

$$\forall t \in \mathbb{R}, z(t) = y(t + t_0)$$

est solution de (7.2) et on sait que ce problème admet pour unique solution la fonction z définie par :

$$\forall t \in \mathbb{R}, z(t) = e^{tA} y_0$$

c'est-à-dire que la solution du problème (7.5) est donnée par :

$$\forall t \in \mathbb{R}, y(t) = z(t - t_0) = e^{(t-t_0)A} y_0$$

Lemme 7.5 Si A, B sont deux matrices qui commutent dans $\mathcal{M}_n(\mathbb{C})$, on a alors $e^{A+B} = e^A e^B = e^B e^A$.

Preuve. On définit les fonctions Y et Z par :

$$\forall t \in \mathbb{R}, Y(t) = e^{t(A+B)}, Z(t) = e^{tA} e^{tB}$$

Ces fonctions sont dérivables sur $\mathbb{R}$ avec :

$$\forall t \in \mathbb{R}, Y'(t) = (A + B)Y(t), Z'(t) = Ae^{tA}e^{tB} + e^{tA}Be^{tB}$$

Du fait que A et B commutent et que l'application $X \mapsto XB$ est continue, on déduit que $e^{tA}B = Be^{tA}$. On a donc $Z'(t) = (A + B)Z(t)$. Finalement avec $Y(0) = Z(0) = I_n$ on déduit que $Y = Z$ (unicité de la solution du problème de Cauchy), c'est-à-dire que $e^{t(A+B)} = e^{tA}e^{tB}$ pour tout réel t . On démontre de manière analogue que $e^{t(A+B)} = e^{tB}e^{tA}$ (on peut aussi utiliser la commutativité de A et B et la continuité du produit matriciel pour montrer que e^A et e^B commutent).

□

Si les matrices A et B ne commutent pas, en général il n'y pas d'égalité entre e^{A+B} , e^Ae^B et e^Be^A (exercices 7.1, 7.2 et 7.3).

En utilisant la décomposition de Dunford $A = D + V$ avec D diagonalisable, V nilpotente et $DV = VD$ on a en notant p l'indice de nilpotence de la matrice V , $e^A = e^De^V = e^D \sum_{k=0}^{p-1} \frac{1}{k!} V^k$, la matrice e^D étant facilement diagonalisable si on sait diagonaliser D .

En fait la décomposition de Dunford de la matrice A permet d'obtenir celle de e^A .

Lemme 7.6 Si $A = D + V$ est la décomposition de Dunford de la matrice A , celle de e^A est alors donnée par $e^A = e^D + e^D(e^V - I_n)$, avec e^D diagonalisable et $e^D(e^V - I_n)$ nilpotente, ces deux matrices commutent.

Preuve. Comme D et V commutent, on a $e^A = e^De^V = e^D + e^D(e^V - I_n)$. On a vu que la matrice e^D est diagonalisable. D'autre part, avec la continuité du produit matriciel on déduit que les matrices V et e^D commutent et on peut écrire que :

$$e^D(e^V - I_n) = e^D V \sum_{k=1}^{p-1} \frac{1}{k!} V^{k-1} = V \left(e^D \sum_{k=1}^{p-1} \frac{1}{k!} V^{k-1} \right)$$

V commutant avec $e^D \sum_{k=1}^{p-1} \frac{1}{k!} V^{k-1}$. Il en résulte que la matrice $e^D(e^V - I_n)$ est nilpotente. Les matrices e^D et $e^D(e^V - I_n)$ commutent, on a obtenu ainsi la décomposition de Dunford de e^A (cette décomposition est unique). □

Lemme 7.7 L'Application $\exp : A \mapsto e^A$ est continue de $\mathcal{M}_n(\mathbb{C})$ dans $GL_n(\mathbb{C})$. Plus précisément, pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, la matrice e^A est inversible d'inverse e^{-A} .

Preuve. Résulte de $I_n = e^{A-A} = e^A e^{-A}$. □

Le calcul du déterminant de e^A pour toute matrice A permet de retrouver le caractère inversible de e^A .

Lemme 7.8 Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ on a $\det(e^A) = e^{\text{Tr}(A)}$.

Preuve. Pour $A \in \mathcal{M}_n(\mathbb{C})$ diagonalisable on a $A = PDP^{-1}$ avec P inversible et D diagonale de valeurs propres $\lambda_1, \dots, \lambda_n$. On a alors $e^A = Pe^DP^{-1}$ et :

$$\det(e^A) = \det(e^D) = \prod_{k=1}^n e^{\lambda_k} = e^{\sum_{k=1}^n \lambda_k} = e^{\text{Tr}(A)}$$

Si A est matrice quelconque dans $\mathcal{M}_n(\mathbb{C})$, on peut alors l'écrire $A = \lim_{k \rightarrow +\infty} A_k$, où $(A_k)_{k \in \mathbb{N}}$ est une suite de matrices diagonalisables. Avec la continuité des applications exponentielles complexe et matricielle et de l'application trace, on déduit que :

$$\det(e^A) = \lim_{k \rightarrow +\infty} \det(e^{A_k}) = \lim_{k \rightarrow +\infty} e^{\text{Tr}(A_k)} = e^{\text{Tr}(A)}$$

On peut aussi utiliser une trigonalisation de la matrice A . □

Lemme 7.9 Si $\lambda \in \mathbb{C}$ est une valeur propre de la matrice $A \in \mathcal{M}_n(\mathbb{C})$, alors pour tout réel t , $e^{\lambda t}$ est valeur propre de e^{tA} et pour tout vecteur propre associé $x \in \mathbb{C}^n$, la solution du problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(0) = x \end{cases} \quad (7.6)$$

est la fonction y définie par :

$$\forall t \in \mathbb{R}, y(t) = e^{\lambda t}x$$

Preuve. Soit $x \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé à la valeur propre λ de A . Pour tout $k \in \mathbb{N}$, on a $A^k x = \lambda^k x$ et, du fait de la continuité de $M \mapsto Mx$, $e^{tA}x = \sum_{k=0}^{+\infty} \frac{t^k}{k!} A^k x = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \lambda^k x = e^{t\lambda}x$, c'est-à-dire que x est un vecteur propre de e^{tA} associé à la valeur propre $e^{t\lambda}$. La solution du problème (7.6) est donc la fonction $t \mapsto e^{tA}x = e^{t\lambda}x$. □

Théorème 7.3.

Si $A \in \mathcal{M}_n(\mathbb{C})$ est diagonalisable de valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_p$ ($1 \leq p \leq n$), les solutions du système différentiel $y' = Ay$ sont alors les fonctions définies par :

$$\forall t \in \mathbb{R}, y(t) = \sum_{k=1}^p e^{\lambda_k t} C_k$$

où les C_k sont des vecteurs de $\mathbb{C}^n$.

Preuve. On désigne par $\{\mu_1, \dots, \mu_n\}$ l'ensemble des valeurs propres de A et par $(e_k)_{1 \leq k \leq n}$ une base de vecteurs propres associée. Avec le corollaire 7.1 et le lemme 7.9, on déduit que $(e^{t\mu_k} e_k)_{1 \leq k \leq n}$ est une base de l'espace des solutions de

$y' = Ay$, c'est à dire que toute solution de ce système s'écrit $y(t) = \sum_{k=1}^n \alpha_k e^{\mu_k t} e_k$,

ou encore en regroupant les valeurs propres identiques $y(t) = \sum_{k=1}^p e^{\lambda_k t} C_k$ avec C_k dans $\ker(A - \lambda_k I_n)$ pour tout $k \in \{1, \dots, p\}$. $\square$

Lemme 7.10 Soit $A \in \mathcal{M}_n(\mathbb{C})$ admettant une seule valeur propre $\lambda \in \mathbb{C}$ et $\pi_A(X) = (X - \lambda)^\alpha$ le polynôme minimal de A avec $1 \leq \alpha \leq n$. Les solutions du système différentiel $y' = Ay$ sont les fonctions définies par :

$$\forall t \in \mathbb{R}, y(t) = e^{\lambda t} \sum_{j=0}^{\alpha-1} \frac{t^j}{j!} (A - \lambda I_n)^j y_0$$

où $y_0 \in \mathbb{C}^n$, ce qui revient à dire que $e^{tA} = e^{\lambda t} \sum_{k=0}^{\alpha-1} \frac{t^k}{k!} (A - \lambda I_n)^k$.

Preuve. Toute solution de $z' = (A - \lambda I_n)z$ s'écrit :

$$z(t) = e^{t(A - \lambda I_n)} z_0 = \sum_{k=0}^{+\infty} \frac{t^k}{k!} (A - \lambda I_n)^k z_0 = \sum_{k=0}^{\alpha-1} \frac{t^k}{k!} (A - \lambda I_n)^k z_0$$

du fait que $(A - \lambda I_n)$ est nilpotente d'ordre α . On en déduit que toute solution de $y' = Ay$ s'écrit :

$$y(t) = e^{tA} y_0 = e^{t(A - \lambda I_n)} e^{t\lambda I_n} y_0 = e^{\lambda t} \sum_{k=0}^{\alpha-1} \frac{t^k}{k!} (A - \lambda I_n)^k y_0$$

$\square$

Théorème 7.4.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ de polynôme minimal $\pi_A(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$, les valeurs propres $\lambda_1, \dots, \lambda_p$ étant deux à deux distincts dans $\mathbb{C}$. Les solutions du système différentiel $y' = Ay$ sont les fonctions définies par :

$$\forall t \in \mathbb{R}, y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t)$$

où, pour tout $k \in \{1, \dots, p\}$, P_k désigne une fonction polynomiale de degré inférieur ou égal à $\alpha_k - 1$ et à valeurs dans $\mathbb{C}^n$.

Preuve. Pour tout $k \in \{1, \dots, p\}$ on note $E_k = \ker(A - \lambda_k I_n)^{\alpha_k}$ le sous espace caractéristique associé à la valeur propre λ_k . D'après le théorème de décomposition des noyaux, on a $\mathbb{C}^n = \bigoplus_{k=1}^p E_k$. Soit y solution de $y' = Ay$. On peut écrire, de manière unique, $y(0) = \sum_{k=1}^p x_k$ avec $x_k \in E_k$ pour tout $k \in \{1, \dots, p\}$. Si, pour

$k \in \{1, \dots, p\}$, on désigne par y_k la solution du problème de Cauchy :

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(0) = x_k \end{cases}$$

$z = \sum_{k=1}^p y_k$ est alors solution du même problème de Cauchy que y . On a donc, du

fait de l'unicité de cette solution, $y = \sum_{k=1}^p y_k$. Mais, d'autre part, on a :

$$y_k(t) = e^{tA} x_k = e^{\lambda_k t} e^{t(A - \lambda_k I_n)} x_k = e^{\lambda_k t} \sum_{j=0}^{\alpha_k - 1} \frac{t^j}{j!} (A - \lambda_k I_n)^j x_k$$

du fait que $(A - \lambda_k I_n)^j x_k = 0$ pour tout $j \geq \alpha_k$. Ce qui donne $y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t)$,

en notant $P_k(t) = \sum_{j=0}^{\alpha_k - 1} \frac{t^j}{j!} (A - \lambda_k I_n)^j x_k$. □

Dans la démonstration précédente, le vecteur x_k est la projection de $y(0)$ sur $\ker(A - \lambda_k I_n)^{\alpha_k}$. En notant π_k cette projection (projecteur spectral), on en déduit que :

$$e^{tA} = \sum_{k=1}^p e^{\lambda_k t} \left(\sum_{j=0}^{\alpha_k - 1} \frac{t^j}{j!} (A - \lambda_k I_n)^j \right) \pi_k$$

On peut retrouver ce résultat en utilisant la décomposition de Dunford (paragraphe 1.6). En effet la matrice A s'écrit $A = D + V$ avec D diagonalisable, V nilpotente et $DV = VD$. On a alors $e^A = e^D e^V$. Puis en écrivant que :

$$\forall r \geq 0, \begin{cases} D^r = \sum_{k=1}^p \lambda_k^r \pi_k \\ V^r = \sum_{k=1}^p (A - \lambda_k I_n)^r \pi_k \end{cases}$$

on obtient :

$$e^D = \sum_{r=0}^{+\infty} \frac{1}{r!} D^r = \sum_{r=0}^{+\infty} \frac{1}{r!} \left(\sum_{k=1}^p \lambda_k^r \pi_k \right) = \sum_{k=1}^p \left(\sum_{r=0}^{+\infty} \frac{\lambda_k^r}{r!} \right) \pi_k = \sum_{k=1}^p e^{\lambda_k} \pi_k$$

et :

$$\begin{aligned} e^V &= \sum_{r=0}^{+\infty} \frac{1}{r!} V^r = \sum_{r=0}^{+\infty} \frac{1}{r!} \left(\sum_{k=1}^p (A - \lambda_k I_n)^r \pi_k \right) \\ &= \sum_{k=1}^p \left(\sum_{r=0}^{\alpha_k - 1} \frac{1}{r!} (A - \lambda_k I_n)^r \pi_k \right) = \sum_{k=1}^p \left(\sum_{r=0}^{\alpha_k - 1} \frac{1}{r!} (A - \lambda_k I_n)^r \right) \pi_k \end{aligned}$$

(la restriction de $A - \lambda_k I_n$ à $\ker(A - \lambda_k I_n)^{\alpha_k}$ est nilpotente d'ordre α_k).

Enfin avec les propriétés des projecteurs, on déduit que :

$$e^A = \sum_{k=1}^p e^{\lambda_k} \left(\sum_{r=0}^{\alpha_k-1} \frac{1}{r!} (A - \lambda_k I_n)^r \right) \pi_k$$

7.3 Un algorithme de calcul de l'exponentielle d'une matrice

Soit $A \in \mathcal{M}_n(\mathbb{C})$ de valeurs propres $\lambda_1, \dots, \lambda_n$ dans $\mathbb{C}$. On lui associe la suite de matrices $(B_k)_{1 \leq k \leq n}$ définie par :

$$\begin{cases} B_1 = I_n \\ B_{k+1} = (A - \lambda_k I_n) B_k \quad (1 \leq k \leq n-1) \end{cases}$$

On a alors $(A - \lambda_n I_n) B_n = 0$. En effet, on a $B_k = \prod_{j=1}^{k-1} (A - \lambda_j I_n)$ pour tout $k \in \{2, \dots, n\}$ et pour $k = n$, $(A - \lambda_n I_n) B_n = \prod_{j=1}^n (A - \lambda_j I_n) = \chi_A(A) = 0$ d'après le théorème de Cayley-Hamilton.

Théorème 7.5.

Avec les notations qui précèdent, si $y = (y_k)_{1 \leq k \leq n}$ est la solution du problème de Cauchy :

$$\begin{cases} y'(t) = Jy(t) \quad (t \in \mathbb{R}) \\ y(0) = e_1 \end{cases} \quad (7.7)$$

$$\text{où } J = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 1 & \lambda_2 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & \lambda_{n-1} & 0 \\ 0 & \cdots & 0 & 1 & \lambda_n \end{pmatrix} \text{ et } e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ on a alors :}$$

$$\forall t \in \mathbb{R}, e^{tA} = \sum_{k=1}^n y_k(t) B_k$$

Preuve. Il suffit de montrer que la fonction $y = \sum_{k=1}^n y_k B_k$ est solution du même problème de Cauchy que $t \mapsto e^{tA}$.

Pour $t = 0$ on a $y(0) = \sum_{k=1}^n y_k(0) B_k = B_1 = I_n$ et pour tout $t \in \mathbb{R}$:

$$\begin{aligned} y'(t) &= \sum_{k=1}^n y'_k(t) B_k = \lambda_1 y_1(t) + \sum_{k=2}^n (y_{k-1}(t) + \lambda_k y_k(t)) B_k \\ &= \sum_{k=1}^n \lambda_k y_k(t) B_k + \sum_{k=1}^{n-1} y_k(t) B_{k+1} \end{aligned}$$

avec :

$$\begin{aligned} \sum_{k=1}^{n-1} y_k(t) B_{k+1} &= \sum_{k=1}^{n-1} y_k(t) (A - \lambda_k I_n) B_k \\ &= A \sum_{k=1}^{n-1} y_k(t) B_k - \sum_{k=1}^{n-1} \lambda_k y_k(t) B_k \end{aligned}$$

ce qui donne en définitive :

$$\begin{aligned} y'(t) &= \lambda_n y_n(t) B_n + Ay(t) - y_n(t) AB_n \\ &= Ay(t) - y_n(t) (A - \lambda_n I_n) B_n = Ay(t) \end{aligned}$$

On peut donc conclure que $e^{tA} = \sum_{k=1}^n y_k(t) B_k$ pour tout réel t . □

7.4 Equations différentielles linéaires d'ordre n à coefficients constants

On considère l'équation différentielle d'ordre n linéaire à coefficients complexes constants :

$$y^{(n)}(t) = a_1 y(t) + a_2 y'(t) + \cdots + a_n y^{(n-1)}(t) \quad (t \in \mathbb{R}) \quad (7.8)$$

Une telle équation différentielle peut se ramener à un système différentiel linéaire $Y' = AY$ de matrice :

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{pmatrix}$$

(matrice compagnon du polynôme $X^n - \sum_{k=0}^{n-1} a_k X^k$).

En effet, si $y \in \mathcal{C}^n(\mathbb{R}, \mathbb{C})$ est une solution de l'équation différentielle (7.8), la fonction vectorielle $Y \in \mathcal{C}^1(\mathbb{R}, \mathbb{C}^n)$ définie par :

$$\forall t \in \mathbb{R}, Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n-1)}(t) \end{pmatrix}$$

est alors solution de :

$$\forall t \in \mathbb{R}, Y'(t) = \begin{pmatrix} y_2(t) \\ \vdots \\ y_n(t) \\ \sum_{k=1}^n a_k y_k(t) \end{pmatrix} = AY(t)$$

$$\text{où } A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \ddots & \ddots & \cdot \\ 0 & 0 & \cdots & 0 & 1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{pmatrix}. \text{ Réciproquement soit } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \text{ dans } \mathcal{C}^1(\mathbb{R}, \mathbb{C}^n) \text{ solution du système différentiel } Y' = AY. \text{ En posant } y = y_1, \text{ on a :}$$

$$y' = y_2, \dots, y^{(n-1)} = y_n \text{ et } y^{(n)} = a_1 y_1 + a_2 y_2 + \cdots + a_n y_n$$

c'est-à-dire que y est solution de (7.8).

Théorème 7.6.

Les solutions de l'équation différentielle (7.8) sont les fonctions de la forme $y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t)$, où $\lambda_1, \dots, \lambda_p$ sont les racines deux à deux distinctes du polynôme $\pi_A(X) = X^n - \sum_{k=1}^n a_k X^{k-1}$ de multiplicités respectives $\alpha_1, \dots, \alpha_p$ et chaque P_k est une fonction polynomiale de degré inférieur ou égal à $\alpha_k - 1$.

Preuve. On sait que toute solution du système différentiel $Y' = AY$ s'écrit

$$Y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t), \text{ où } \lambda_1, \dots, \lambda_p \text{ sont les valeurs propres deux à deux distinctes}$$

de A et pour tout $k \in \{1, \dots, p\}$, P_k est un polynôme à coefficients dans $\mathbb{C}^n$ de degré strictement inférieur à la multiplicité α_k de λ_k comme racine du polynôme minimal de A (théorème 7.4).

Avec les notations de ce paragraphe, le polynôme minimal de la matrice A est défini par $\pi_A(X) = X^n - \sum_{k=1}^n a_k X^{k-1}$ (théorème 1.7), donc toute solution de

l'équation différentielle (7.8) est de la forme $y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t)$, où $\lambda_1, \dots, \lambda_p$ sont les racines deux à deux distinctes du polynôme π_A et pour tout $k \in \{1, \dots, p\}$, P_k est un polynôme de degré strictement inférieur à α_k . $\square$

Le polynôme π_A est appelé le polynôme caractéristique de l'équation différentielle d'ordre n (7.8).

7.5 Systèmes différentiels linéaires à coefficients non constants

Soient $[\alpha, \beta]$ un intervalle réel fermé borné, t_0 un point donné dans $[\alpha, \beta]$, $K : [\alpha, \beta]^2 \rightarrow \mathbb{C}^n$ et $h : [\alpha, \beta] \rightarrow \mathbb{C}^n$ deux applications continues. On désigne $E = \mathcal{C}^0([\alpha, \beta], \mathbb{C}^n)$ l'espace des fonctions continues de $[\alpha, \beta]$ dans $\mathbb{C}^n$ muni de la norme de la convergence uniforme et on définit l'application $\varphi : E \rightarrow E$ par :

$$\forall y \in E, \forall t \in [\alpha, \beta], \varphi(y)(t) = \int_{t_0}^t K(t, u) y(u) du + h(t)$$

Lemme 7.11 *Avec les notations qui précèdent, il existe un entier p strictement positif tel que l'application $\varphi^p = \underbrace{\varphi \circ \dots \circ \varphi}_{p \text{ fois}}$ soit contractante. En conséquence,*

l'application φ admet un unique point fixe dans E .

Preuve. Il est clair que pour toute fonction y continue sur $[\alpha, \beta]$ la fonction $\varphi(y)$ est également continue sur $[\alpha, \beta]$. Pour y, z , dans E et t dans $[\alpha, \beta]$ on a :

$$\begin{aligned} \|\varphi(y)(t) - \varphi(z)(t)\| &= \left\| \int_{t_0}^t K(t, u) (y(u) - z(u)) du \right\| \\ &\leq \lambda \left| \int_{t_0}^t \|y(u) - z(u)\| du \right| \leq \lambda \|y - z\|_{\infty} |t - t_0| \end{aligned}$$

où $\lambda = \sup_{(t, u) \in [\alpha, \beta]^2} \|K(t, u)\|$. Par récurrence sur $k \geq 1$, on déduit que :

$$\forall t \in [\alpha, \beta], \|\varphi^k(y)(t) - \varphi^k(z)(t)\| \leq \frac{(\lambda |t - t_0|)^k}{k!} \|y - z\|_{\infty}$$

En effet, on vient de voir que le résultat est vrai pour $k = 1$ et le supposant acquis pour $k \geq 1$, on a pour tout $t \in [\alpha, \beta]$:

$$\begin{aligned} \|\varphi^{k+1}(y)(t) - \varphi^{k+1}(z)(t)\| &\leq \lambda \left| \int_{t_0}^t \|\varphi^k(y)(u) - \varphi^k(z)(u)\| du \right| \\ &\leq \lambda \frac{\lambda^k}{k!} \|y - z\|_{\infty} \left| \int_{t_0}^t |u - t_0|^k du \right| \end{aligned}$$

avec $\left| \int_{t_0}^t |u - t_0|^k du \right| = \frac{|t - t_0|^{k+1}}{k+1}$, ce qui donne :

$$\|\varphi^{k+1}(y)(t) - \varphi^{k+1}(z)(t)\| \leq \frac{(\lambda |t - t_0|)^{k+1}}{(k+1)!} \|y - z\|_{\infty}$$

Il en résulte que :

$$\forall k \geq 1, \|\varphi^k(y) - \varphi^k(z)\|_\infty \leq \frac{(\lambda(\beta - \alpha))^k}{k!} \|y - z\|_\infty$$

puis, avec $\lim_{k \rightarrow +\infty} \frac{(\lambda(\beta - \alpha))^k}{k!} = 0$, on en déduit que pour $k \geq 1$ assez grand l'application φ^k est contractante. Le théorème du point fixe itéré (l'espace $(E, \|\cdot\|_\infty)$ est complet) nous dit alors que φ admet un unique point fixe dans E . $\square$

Lemme 7.12 Soient $[\alpha, \beta]$ un intervalle réel fermé borné et

$$A : [\alpha, \beta] \rightarrow \mathcal{M}_n(\mathbb{C}), \quad b : [\alpha, \beta] \rightarrow \mathbb{C}^n$$

deux applications continues. Pour tout $t_0 \in [\alpha, \beta]$ et tout $y_0 \in \mathbb{C}^n$ le problème de Cauchy :

$$\begin{cases} y'(t) = A(t)y(t) + b(t) & (t \in [\alpha, \beta]) \\ y(t_0) = y_0 \end{cases} \quad (7.9)$$

admet une unique solution.

Preuve. Dire que $y \in \mathcal{C}^1([\alpha, \beta], \mathbb{C}^n)$ est solution de (7.9) équivaut à dire que pour tout $t \in [\alpha, \beta]$ on a :

$$y(t) = \int_{t_0}^t y'(u) du + y(t_0) = \int_{t_0}^t A(u)y(u) du + \int_{t_0}^t b(u) du + y_0$$

c'est-à-dire que $y = \varphi(y)$ où φ est définie par :

$$K : (t, u) \mapsto A(u), \quad h : t \mapsto \int_{t_0}^t b(u) du + y_0$$

On déduit alors de ce qui précède que (7.9) a une unique solution. $\square$

Dans ce qui suit I est un intervalle réel et $A : I \rightarrow \mathcal{M}_n(\mathbb{C}), b : I \rightarrow \mathbb{C}^n$ sont deux applications continues.

Théorème 7.7.

Pour tout $t_0 \in I$ et tout $y_0 \in \mathbb{C}^n$ le problème de Cauchy :

$$\begin{cases} y'(t) = A(t)y(t) + b(t) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (7.10)$$

admet une unique solution.

Preuve. Pour tout intervalle fermé borné $J \subset I$ contenant t_0 , on note y_J la solution de (7.9) sur J . Si J_1 et J_2 sont deux tels intervalles, du fait de l'unicité de la solution du problème de Cauchy sur un intervalle compact on déduit que $y_{J_1}(t) = y_{J_2}(t)$ pour tout $t \in J_1 \cap J_2$. On peut donc définir sur I la fonction y par $y(t) = y_J(t)$ où J est n'importe quel intervalle compact contenu dans I et tel que $t_0 \in J$. Il est alors clair que y est solution de (7.10) et que ce problème admet une unique solution dans $\mathcal{C}^1(I, \mathbb{C}^n)$. $\square$

Corollaire 7.2 : L'ensemble $\mathcal{S}_I$ des solutions sur l'intervalle I du système différentiel $y'(t) = A(t)y(t)$ est un espace vectoriel de dimension n sur $\mathbb{C}$. Une famille $(y_k)_{1 \leq k \leq n}$ de solutions est une base de $\mathcal{S}_I$ si, et seulement si, il existe $t_0 \in I$ tel que la famille de vecteurs $(y_k(t_0))_{1 \leq k \leq n}$ soit une base de $\mathbb{C}^n$.

Preuve. Du théorème 7.7 avec $b = 0$, on déduit que $\mathcal{S}_I$ est non vide et que, pour tout $t_0 \in I$, l'application linéaire $\psi : y \mapsto y(t_0)$ réalise un isomorphisme de $\mathcal{S}_I$ sur $\mathbb{C}^n$. Il en résulte que $\mathcal{S}_I$ est un espace vectoriel de dimension n sur $\mathbb{C}$. Si la famille de vecteurs $(y_k(t_0))_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$, son image réciproque $(y_k)_{1 \leq k \leq n}$ par l'isomorphisme ψ est alors une base de $\mathcal{S}_I$. $\square$

Ce résultat nous dit que si l'application $w : t \mapsto \det(y_1(t), \dots, y_n(t))$ s'annule en un point, elle est alors identiquement nulle.

Corollaire 7.3 : Soit u une solution sur l'intervalle I du système différentiel avec second membre :

$$y'(t) = A(t)y(t) + b(t) \quad (7.11)$$

Toute autre solution de ce système s'écrit $z = u + y$, où y est solution du système différentiel homogène $y'(t) = A(t)y(t)$ sur l'intervalle I .

Preuve. Il est clair que pour toute solution z de (7.11) la fonction $y = z - u$ est solution de $y' = Ay$. On déduit que si $(y_k)_{1 \leq k \leq n}$ est une base de solutions de $y' = Ay$ sur I , toute solution de $y' = Ay + b$ s'écrit alors $z = \sum_{k=1}^n \lambda_k y_k + u$, où les scalaires λ_k sont uniquement déterminés par $z(t_0)$, t_0 étant donné dans I . $\square$

Définition 7.1. Soit $(y_k)_{1 \leq k \leq n}$ une base de l'espace vectoriel des solutions sur l'intervalle I du système différentiel $y'(t) = A(t)y(t)$. On appelle wronskien de cet ensemble de solutions, la fonction :

$$\begin{aligned} w : I &\rightarrow \mathbb{C} \\ t &\mapsto w(t) = \det(y_1(t), \dots, y_n(t)) \end{aligned}$$

Théorème 7.8.

Avec les notations qui précèdent, le wronskien w est solution de l'équation différentielle :

$$w'(t) = \text{Tr}(A(t))w(t) \quad (t \in I)$$

Pour tout $t_0 \in I$, on a $w(t) = w(t_0)e^{g(t)}$ pour tout $t \in I$, en notant $g(t) = \int_{t_0}^t \text{Tr}(A(x))dx$. Le wronskien d'une base de solutions sur I du système $y'(t) = A(t)y(t)$ ne s'annule jamais.

Preuve. Les y_j étant de classe $\mathcal{C}^1$ et l'application déterminant de classe $\mathcal{C}^\infty$ on déduit que w est de classe $\mathcal{C}^1$ avec :

$$w' = \sum_{j=1}^n \det(y_1, \dots, y_{j-1}, y_j', \dots, y_n) = \sum_{j=1}^n \det(y_1, \dots, y_{j-1}, Ay_j, \dots, y_n)$$

(exercice 2.5). Pour tout $t \in I$ l'application :

$$\varphi_t : (y_1, \dots, y_n) \mapsto \sum_{j=1}^n \det(y_1, \dots, y_{j-1}, A(t)y_j, \dots, y_n)$$

étant n -linéaire alternée sur $(\mathbb{C}^n)^n$ est proportionnelle au déterminant, c'est-à-dire qu'il existe une constante $\lambda(t) \in \mathbb{C}$ telle que :

$$\forall Y = (y_1, \dots, y_n) \in (\mathbb{C}^n)^n, \varphi_t(Y) = \lambda(t) \det(Y)$$

Cette constante est égale à $\varphi_t(I_n)$, soit en notant $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$:

$$\lambda(t) = \sum_{j=1}^n \det(e_1, \dots, e_{j-1}, A(t)e_j, \dots, e_n) = \sum_{j=1}^n a_{jj}(t) = \text{Tr}(A(t))$$

On a donc en définitive, $w'(t) = \text{Tr}(A(t))w(t)$ pour tout $t \in I$.

Soit, pour t_0 fixé dans I , $g : t \mapsto \int_{t_0}^t \text{Tr}(A(x)) dx$ la primitive de $\text{Tr}(A)$ nulle en t_0 . On a $w'(t) = g'(t)w(t)$ et $(w(t)e^{-g(t)})' = 0$ pour tout $t \in I$. On déduit donc qu'il existe une constante $C \in \mathbb{C}$ telle que $w(t) = Ce^{g(t)}$ et $t = t_0$ donne $C = w(t_0)$. On a donc $w(t) = w(t_0)e^{g(t)}$, pour tout $t \in I$.

On sait que $(y_k)_{1 \leq k \leq n}$ est une base de solutions si, et seulement si, la famille de vecteurs $(y_k(t_0))_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$, ce qui équivaut à dire que $w(t_0) \neq 0$, ce qui est encore équivalent à $w(t) \neq 0$ pour tout $t \in I$. On a donc ainsi montré que le wronskien d'une base de solutions ne s'annule jamais sur I . $\square$

7.6 Méthode de variation des constantes

Théorème 7.9.

Si la famille $(y_k)_{1 \leq k \leq n}$ est une base de l'espace vectoriel des solutions du système $y'(t) = \bar{A}(t)y(t)$ sur l'intervalle I , la fonction y définie sur

I par $y(t) = \sum_{i=1}^n \lambda_i(t) y_i(t)$ est alors solution particulière sur l'intervalle

I du système différentiel avec second membre $y'(t) = A(t)y(t) + b(t)$ si, et seulement si, pour tout $t \in I$, le vecteur $(\lambda_k'(t))_{1 \leq k \leq n}$ est solution du

système linéaire $\sum_{i=1}^n \lambda_i'(t) y_i(t) = b(t)$.

Preuve. La fonction $y = \sum_{i=1}^n \lambda_i y_i$ est solution sur l'intervalle I du système différentiel $y' = Ay + b$ si, et seulement si, on a $\sum_{i=1}^n \lambda'_i y_i + \sum_{i=1}^n \lambda_i y'_i = Ay + b$, avec $\sum_{i=1}^n \lambda_i y'_i = \sum_{i=1}^n \lambda_i A y_i = Ay$, ce qui équivaut à dire que pour tout $t \in I$ le vecteur $(\lambda'_k(t))_{1 \leq k \leq n}$ est solution du système linéaire $\sum_{i=1}^n \lambda'_i(t) y_i(t) = b(t)$. Le déterminant de ce système est le wronskien de la base de solutions $(y_k)_{1 \leq k \leq n}$, il n'est donc jamais nul et on a une unique solution. $\square$

Corollaire 7.4 : Soient $p, q, r : I \rightarrow \mathbb{C}$ des fonctions continues. L'ensemble des solutions sur l'intervalle I de l'équation différentielle d'ordre 2 :

$$y'' = py' + qy \quad (7.12)$$

est un espace vectoriel de dimension 2 sur $\mathbb{C}$. Si (u, v) est une base de solutions sur l'intervalle I , de l'équation différentielle $y'' = py' + qy$, la fonction y définie par $y(t) = \lambda(t)u(t) + \mu(t)v(t)$ est alors solution sur l'intervalle I , de l'équation différentielle $y'' = py' + qy + r$ si, pour tout $t \in I$, $(\lambda'(t), \mu'(t))$ est solution du système linéaire de deux équations à deux inconnues :

$$\begin{cases} \lambda'(t)u(t) + \mu'(t)v(t) = 0 \\ \lambda'(t)u'(t) + \mu'(t)v'(t) = r(t) \end{cases}$$

Preuve. La fonction y est solution de (7.12) si, et seulement si, la fonction vectorielle $z = \begin{pmatrix} y \\ y' \end{pmatrix}$ est solution du système différentiel $z' = Az$ où $A = \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$ est continue de I dans $M_2(\mathbb{C})$. On en déduit alors que l'ensemble des solutions de (7.12) est un espace vectoriel sur $\mathbb{C}$ de dimension 2.

Si (u, v) est une base de solutions sur I , de l'équation $y'' = py' + qy$, le couple $\left(z = \begin{pmatrix} u \\ u' \end{pmatrix}, w = \begin{pmatrix} v \\ v' \end{pmatrix}\right)$ est alors une base de solutions sur I du système différentiel $z' = Az$ (il est facile de vérifier que (z, w) est libre).

La fonction $Z = \lambda z + \mu w$ est solution de $Z' = AZ + b$ si, et seulement si, pour tout $t \in I$ le vecteur $\begin{pmatrix} \lambda'(t) \\ \mu'(t) \end{pmatrix}$ est solution de $\lambda'(t)z(t) + \mu'(t)w(t) = b(t)$, soit :

$$\begin{cases} \lambda'(t)u(t) + \mu'(t)v(t) = 0 \\ \lambda'(t)u'(t) + \mu'(t)v'(t) = r(t) \end{cases}$$

Dans ce cas on a alors, en posant $y = \lambda u + \mu v$:

$$Z = \lambda z + \mu w = \begin{pmatrix} \lambda u + \mu v \\ \lambda u' + \mu v' \end{pmatrix} = \begin{pmatrix} y \\ y' \end{pmatrix}$$

et $Z' = \begin{pmatrix} y' \\ y'' \end{pmatrix} = AZ + b = \begin{pmatrix} y' \\ py' + qy + r \end{pmatrix}$, c'est-à-dire que la fonction $y = \lambda u + \mu v$ est solution sur I de $y'' = py' + qy + r$. $\square$

Le wronskien du système différentiel $z' = Az$ est $w(t) = w(t_0)e^{g(t)}$, en notant $g(t) = \int_{t_0}^t \text{Tr}(A(x)) dx = \int_{t_0}^t p(x) dx$. Ce wronskien étant $w(t) = \det(z(t), w(t))$, on en déduit que $u(t)v'(t) - u'(t)v(t) = w(t_0)e^{g(t)}$. Ce résultat peut être utilisé pour calculer une solution de base v connaissant une solution u .

7.7 Surjectivité et injectivité de l'exponentielle matricielle

On a vu que pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ la matrice e^A est inversible (lemme 7.7) et avec $\det(e^A) = e^{\text{Tr}(A)}$, on déduit que si A est à coefficients réels, e^A est alors dans l'ensemble $GL_n^+(\mathbb{R})$ des matrices réelles d'ordre n de déterminant strictement positif.

On note $\mathcal{N}_n(\mathbb{K})$ l'ensemble des matrices à coefficients réels ou complexes d'ordre n qui sont nilpotentes.

Définition 7.2. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est unipotente si la matrice $A - I_n$ est nilpotente.

On note $\mathcal{L}_n(\mathbb{K})$ l'ensemble des matrices à coefficients réels ou complexes d'ordre n qui sont unipotentes.

Lemme 7.13 Si $f(z) = \sum_{k=0}^{+\infty} a_k z^k$ est une série entière de rayon de convergence $R > 0$, alors pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\rho(A) < R$ la série de terme général $a_k A^k$ est convergente.

Preuve. Voir l'exercice 3.16. $\square$

Avec les notations du lemme précédent la matrice $f(A) = \sum_{k=0}^{+\infty} a_k A^k$, pour $\rho(A) < R$, est en fait un polynôme en A , dont les coefficients dépendent de A . En effet, $f(A)$ est dans l'adhérence de $\mathbb{C}[A]$ qui est fermé dans $\mathcal{M}_n(\mathbb{C})$ en tant que sous-espace vectoriel d'un espace vectoriel normé de dimension finie.

En partant du développement en série entière de la fonction logarithme complexe :

$$\ln(1+z) = \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{k} z^k \quad (|z| < 1)$$

on peut définir la fonction logarithme matricielle, sur l'ensemble des matrices A dans $\mathcal{M}_n(\mathbb{C})$ telles que $\rho(A) < 1$ par :

$$\ln(I_n + A) = \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{k} A^k$$

En particulier on a $\ln(I_n) = 0$ et pour toute matrice nilpotente A , en notant p un entier naturel tel que $A^{p+1} = 0$, on a $\ln(I_n + A) = \sum_{k=1}^p \frac{(-1)^{k-1}}{k} A^k$.

Lemme 7.14 *Pour $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\rho(A) < 1$, on a $e^{\ln(I_n + A)} = I_n + A$.*

Preuve. Pour tout $x \in \mathbb{R}$, on a $e^x = \sum_{k=0}^{+\infty} \alpha_k x^k$ avec $\alpha_k = \frac{1}{k!}$ pour tout entier

naturel k et pour tout réel x tel que $|x| < 1$, on a $\ln(1+x) = \sum_{j=1}^{+\infty} \beta_j x^j$ avec

$\beta_j = \frac{(-1)^{j-1}}{j}$ pour tout entier naturel j non nul. On peut alors écrire pour $k \geq 1$ et $|x| < 1$, que $(\ln(1+x))^k = \sum_{j=k}^{+\infty} \beta_{k,j} x^j$ et :

$$e^{\ln(1+x)} = 1 + \sum_{k=1}^{+\infty} \alpha_k \left(\sum_{j=k}^{+\infty} \beta_{k,j} x^j \right) = 1 + x + \sum_{j=2}^{+\infty} \left(\sum_{k=1}^j \alpha_k \beta_{k,j} \right) x^j$$

Avec $e^{\ln(1+x)} = 1 + x$, on déduit alors que :

$$\forall k \geq 2, \sum_{k=1}^j \alpha_k \beta_{k,j} = 0$$

En écrivant que $e^{\ln(I_n + A)} = I_n + A + \sum_{j=2}^{+\infty} \left(\sum_{k=1}^j \alpha_k \beta_{k,j} \right) A^j$, pour $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\rho(A) < 1$, on en déduit que $e^{\ln(I_n + A)} = I_n + A$. $\square$

Lemme 7.15 *Pour toute matrice $A \in \mathcal{N}_n(\mathbb{C})$ on a $e^A \in \mathcal{L}_n(\mathbb{C})$ et :*

$$\forall t \in \mathbb{R}, \ln(e^{tA}) = tA$$

Preuve. Si $A \in \mathcal{N}_n(\mathbb{C})$, il existe $p \in \mathbb{N}$ tel que $A^{p+1} = 0$ et $e^A = I_n + V$, avec $V = A \sum_{k=1}^p \frac{1}{k!} A^{k-1}$ qui est nilpotente. On a donc $e^A \in \mathcal{L}_n(\mathbb{C})$. Pour tout réel

t , on a également $e^{tA} = I_n + V(t) \in \mathcal{L}_n(\mathbb{C})$ avec $V(t) = \sum_{k=1}^p \frac{1}{k!} t^k A^k$ telle que $V(t)^{p+1} = 0$. La fonction V est dérivable sur $\mathbb{R}$ ainsi que la fonction φ définie par :

$$\varphi(t) = \ln(e^{tA}) - tA = \sum_{k=1}^{p+1} \frac{(-1)^{k-1}}{k} V(t)^k - tA$$

avec $\varphi'(t) = V'(t) \left(\sum_{k=1}^{p+1} (-1)^{k-1} V(t)^{k-1} \right) - A$.

Il est facile de vérifier que $(I_n + V(t)) \left(\sum_{k=1}^{p+1} (-1)^{k-1} V(t)^{k-1} \right) = I_n$, c'est-à-dire que :

$$\sum_{k=1}^{p+1} (-1)^{k-1} V(t)^{k-1} = (I_n + V(t))^{-1} = (e^{tA})^{-1} = e^{-tA}$$

et avec $V'(t) = (e^{tA} - I_n)' = Ae^{tA}$, on déduit que $\varphi'(t) = 0$ pour tout réel t . On a donc :

$$\forall t \in \mathbb{R}, \varphi(t) = \varphi(0) = \ln(I_n) = 0$$

ce qui équivaut à $\ln(e^{tA}) = tA$ pour tout réel t . $\square$

Théorème 7.10.

L'exponentielle matricielle réalise une bijection de $\mathcal{N}_n(\mathbb{C})$ sur $\mathcal{L}_n(\mathbb{C})$ d'inverse le logarithme matricielle.

Preuve. On sait déjà que l'exponentielle matricielle envoie $\mathcal{N}_n(\mathbb{C})$ dans $\mathcal{L}_n(\mathbb{C})$ et que pour tout $A \in \mathcal{N}_n(\mathbb{C})$, on a $e^{\ln(I_n + A)} = I_n + A$ avec $B = \ln(I_n + A) \in \mathcal{N}_n(\mathbb{C})$, ce qui prouve que l'exponentielle matricielle réalise une surjection de $\mathcal{N}_n(\mathbb{C})$ sur $\mathcal{L}_n(\mathbb{C})$. Si A_1, A_2 dans $\mathcal{N}_n(\mathbb{C})$ sont telles que $e^{A_1} = e^{A_2}$, alors $\ln(e^{A_1}) = \ln(e^{A_2})$, c'est-à-dire, d'après le lemme précédent avec $t = 1$, que $A_1 = A_2$. L'exponentielle matricielle restreinte à $\mathcal{N}_n(\mathbb{C})$ est donc injective. $\square$

Corollaire 7.5 : *Pour tout nombre complexe λ non nul et pour toute matrice $A \in \mathcal{N}_n(\mathbb{C})$ il existe $X \in \mathcal{M}_n(\mathbb{C})$ telle que $e^X = \lambda I_n + A$.*

Preuve. Soient $\lambda \in \mathbb{C}$ et $A \in \mathcal{N}_n(\mathbb{C})$. On sait que la fonction exponentielle complexe est surjective de $\mathbb{C}$ sur $\mathbb{C}^*$, il existe donc un nombre complexe μ tel que $\lambda = e^\mu$ et en posant $X = \mu I_n + \ln\left(I_n + \frac{1}{\lambda}A\right)$, on a :

$$e^X = e^{\mu I_n} e^{\ln(I_n + \frac{1}{\lambda}A)} = \lambda I_n \left(I_n + \frac{1}{\lambda}A \right) = \lambda I_n + A.$$

$\square$

Théorème 7.11.

L'exponentielle matricielle réalise une surjection de $\mathcal{M}_n(\mathbb{C})$ sur $GL_n(\mathbb{C})$.

Preuve. Le théorème de réduction de Jordan nous dit que toute matrice A dans $GL_n(\mathbb{C})$ est semblable à une matrice bloc de la forme :

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_p \end{pmatrix}$$

avec $J_k = \lambda_k I_n + V_k$, pour tout entier k compris entre 1 et p , la matrice V_k étant nilpotente et λ_k étant valeur propre de A . Comme la matrice A est inversible, tous les λ_k sont non nuls et on peut trouver des matrices à coefficients complexes X_k telles que $e^{X_k} = J_k$.

Si $A = PJP^{-1}$ avec P inversible, en définissant la matrice X par :

$$X = P \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & X_p \end{pmatrix} P^{-1},$$

on a :

$$e^X = P \begin{pmatrix} e^{X_1} & 0 & \cdots & 0 \\ 0 & e^{X_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{X_p} \end{pmatrix} P^{-1} = PJP^{-1} = A$$

□

Il est facile de vérifier que l'exponentielle matricielle définie sur $\mathcal{M}_n(\mathbb{C})$ n'est pas injective. Par exemple, pour tout entier relatif k on a $e^{2ik\pi I_n} = I_n$, c'est-à-dire que l'équation $e^X = I_n$ a une infinité de solutions dans $\mathcal{M}_n(\mathbb{C})$. En fait les solutions sont les matrices semblables aux matrices $2i\pi K$ où K est une matrice diagonale d'ordre n à coefficients entiers relatifs (exercice 7.14).

Le théorème précédent nous permet de retrouver la connexité de $GL_n(\mathbb{C})$.

Corollaire 7.6 : $GL_n(\mathbb{C})$ est connexe par arcs.

Preuve. Soient A_1, A_2 deux matrices dans $GL_n(\mathbb{C})$. Il existe deux matrices X_1, X_2 dans $\mathcal{M}_n(\mathbb{C})$ telles que $e^{X_1} = A_1$ et $e^{X_2} = A_2$. L'application φ définie sur $[0, 1]$ par $\varphi(t) = e^{(1-t)X_1 + tX_2}$ est alors un chemin continu dans $GL_n(\mathbb{C})$ qui relie A_1 et A_2 . □

On déduit aussi l'existence d'une racine p -ième pour toute matrice inversible.

Corollaire 7.7 : Soit p un entier naturel non nul. Pour toute matrice $A \in GL_n(\mathbb{C})$ il existe une matrice $X \in GL_n(\mathbb{C})$ telle que $X^p = A$.

Preuve. Il existe une matrice $Y \in \mathcal{M}_n(\mathbb{C})$ telle que $e^Y = A$. En posant $X = e^{\frac{1}{p}Y}$, on a $X \in GL_n(\mathbb{C})$ et $X^p = e^Y = A$. □

Si on se restreint aux matrices complexes hermitiennes [resp. réelles symétriques] l'exponentielle est injective et son image est l'ensemble des matrices complexes hermitiennes définies positives [resp. réelles symétriques définies positives].

Théorème 7.12.

L'exponentielle matricielle réalise une bijection de l'ensemble des matrices complexes hermitiennes sur l'ensemble des matrices complexes hermitiennes définies positives.

Preuve. Il est facile de vérifier que l'exponentielle d'une matrice hermitienne est également hermitienne. De plus si X est hermitienne, ses valeurs propres $\lambda_1, \dots, \lambda_n$ sont alors réelles et celles de e^X sont les réels strictement positifs $e^{\lambda_1}, \dots, e^{\lambda_n}$, en conséquence la matrice e^X est définie positive.

Soit $A = UDU^*$ une matrice hermitienne définie positive où U est unitaire et D diagonale de valeurs propres réelles strictement positives $\mu_1, \dots, \mu_n$. On note P le polynôme d'interpolation de Lagrange défini par $P(\mu_i) = \ln(\mu_i)$ pour tout $i \in \{1, \dots, n\}$ (le degré de P est $p-1$ où p est le nombre de valeurs propres distinctes de A). La matrice $X = P(A) = UP(D)U^*$ est hermitienne et telle que $e^X = A$. Si Y est une autre matrice hermitienne telle que $e^Y = A$, Y commute alors avec A , donc avec X qui est polynomiale en A . En définitive les matrices X et Y commutent et sont diagonalisables, on sait alors qu'elles sont simultanément diagonalisables (exercice 2.3), c'est-à-dire qu'il existe une matrice P dans $GL_n(\mathbb{C})$ telle que $X = P\Delta P^{-1}$ et $Y = P\Lambda P^{-1}$ avec Δ et Λ diagonales à coefficients réels. De $e^X = e^Y$ on déduit alors que $e^\Delta = e^\Lambda$ et $\Delta = \Lambda$ du fait que ces matrices sont diagonales à coefficients réels et que l'exponentielle réelle est bijective de $\mathbb{R}$ sur $\mathbb{R}_+^*$. On a donc $X = Y$, c'est à dire que l'équation $e^X = A$ a une unique solution hermitienne. $\square$

7.8 Exercices

Exercice 7.1. Comparer e^{A+B} , $e^A e^B$ et $e^B e^A$ pour $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$.

Solution. Le polynôme caractéristique de ces deux matrices est $X(X-1)$. Avec le théorème de Cayley-Hamilton on déduit alors qu'elles vérifient les équations $M^2 = M$ et donc $M^k = M$ pour tout entier k supérieur ou égal à 2. Il en résulte que :

$$e^A = I_2 + (e-1)A = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix}, \quad e^B = I_2 + (e-1)B = \begin{pmatrix} e & 1-e \\ 0 & 1 \end{pmatrix}$$

Ce qui donne :

$$e^A e^B = \begin{pmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{pmatrix}, \quad e^B e^A = \begin{pmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{pmatrix} \neq e^A e^B$$

D'autre part la matrice $A+B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ a pour polynôme caractéristique $X(X-2)$, ce qui entraîne, en utilisant le théorème de Cayley-Hamilton, que $(A+B)^2 = 2(A+B)$ et :

$$e^{A+B} = I_2 + \frac{1}{2}(e^2-1)(A+B) = \begin{pmatrix} e^2 & 0 \\ 0 & 1 \end{pmatrix}$$

est différent de $e^A e^B$ et de $e^B e^A$.

Exercice 7.2. Soient A, B dans $\mathcal{M}_n(\mathbb{C})$. Montrer que $e^{t(A+B)} = e^{tA}e^{tB}$ pour tout réel t si, et seulement si, A et B commutent.

Solution. Pour tout réel t on a :

$$e^{t(A+B)} = I_n + t(A+B) + \frac{t^2}{2}(A+B)^2 + o(t^2)$$

$$e^{tA}e^{tB} = I_n + t(A+B) + \frac{t^2}{2}(A+2AB+B) + o(t^2)$$

et de l'unicité du développement limité d'ordre 2, on déduit que l'égalité $e^{t(A+B)} = e^{tA}e^{tB}$ est réalisée si, et seulement si, $(A+B)^2 = A+2AB+B$ ce qui équivaut à $AB = BA$.

Exercice 7.3. Pour A, B dans $\mathcal{M}_n(\mathbb{C})$, on note $[A, B] = AB - BA$.

1. Soient A, B dans $\mathcal{M}_n(\mathbb{C})$ telles que $[A, [A, B]] = 0$. Montrer que :

$$\forall t \in \mathbb{R}, e^{tA}Be^{-tA} = B + t[A, B]$$

2. Soient A, B dans $\mathcal{M}_n(\mathbb{C})$ telles que $[A, [A, B]] = [B, [B, A]] = 0$. Montrer que :

$$\forall t \in \mathbb{R}, e^{-t(A+B)}e^{tA}e^{tB} = e^{\frac{t^2}{2}[A, B]}$$

3. Soient A, B dans $\mathcal{M}_n(\mathbb{C})$ telles que $[A, [A, B]] = [B, [B, A]] = 0$. Montrer que $e^{A+B} = e^Ae^Be^{-\frac{1}{2}[A, B]}$.

Solution.

1. En notant $Y(t) = e^{tA}Be^{-tA}$ pour tout réel t , on a $Y'(t) = e^{tA}[A, B]e^{-tA}$. Si de plus on a $[A, [A, B]] = 0$, les matrices A et $[A, B]$ commutent alors, ce qui entraîne que les matrices e^{tA} et $[A, B]$ commutent de sorte que :

$$\forall t \in \mathbb{R}, Y'(t) = [A, B]e^{tA}e^{-tA} = [A, B]$$

et, avec $Y(0) = B$, on déduit que $Y(t) = B + t[A, B]$.

2. En notant $Y(t) = e^{-t(A+B)}e^{tA}e^{tB}$, $Z(t) = e^{\frac{t^2}{2}[A, B]}$ pour tout réel t , on a $Z'(t) = t[A, B]Z(t)$ et :

$$\begin{aligned} Y'(t) &= e^{-t(A+B)}(-Be^{tA}e^{tB} + e^{tA}Be^{tB}) \\ &= e^{-t(A+B)}(-B + e^{tA}Be^{-tA})e^{tA}e^{tB} \end{aligned}$$

Avec l'hypothèse $[A, [A, B]] = 0$, on déduit que $-B + e^{tA}Be^{-tA} = t[A, B]$ et :

$$\forall t \in \mathbb{R}, Y'(t) = te^{-t(A+B)}[A, B]e^{tA}e^{tB}$$

Enfin avec l'hypothèse $[A, [A, B]] = [B, [B, A]] = 0$, on déduit que $[A, B]$ et $A+B$ commutent, ce qui entraîne que $[A, B]$ et $e^{-t(A+B)}$ commutent de sorte que $Y'(t) = t[A, B]Y(t)$ pour tout $t \in \mathbb{R}$. Avec $Y(0) = Z(0) = I_n$ on déduit que Y et Z sont solutions du même problème de Cauchy ce qui donne $Y = Z$.

3. En faisant $t = 1$ dans l'égalité précédente on a $e^{-(A+B)}e^Ae^B = e^{\frac{1}{2}[A,B]}$, ce qui équivaut à $e^{A+B} = e^Ae^Be^{-\frac{1}{2}[A,B]}$.

Exercice 7.4. Montrer que pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ on a

$$\lim_{p \rightarrow +\infty} \left(I_n + \frac{1}{p}A \right)^p = e^A.$$

Solution. On pose, pour tout entier naturel p , $A_p = \left(I_n + \frac{1}{p}A \right)^p = \sum_{k=0}^p \frac{\binom{p}{k}}{p^k} A^k$.

On note $\|\cdot\|$ une norme sous-multiplicative sur $\mathcal{M}_n(\mathbb{C})$. Pour tout entier naturel p , on a $\|e^A - A_p\| \leq \left\| \sum_{k=0}^p \left(\frac{1}{k!} - \frac{\binom{p}{k}}{p^k} \right) A^k \right\| + \sum_{k=p+1}^{+\infty} \frac{\|A\|^k}{k!}$ avec $\lim_{p \rightarrow +\infty} \sum_{k=p+1}^{+\infty} \frac{\|A\|^k}{k!} = 0$.

En remarquant que $\frac{1}{k!} - \frac{\binom{p}{k}}{p^k} = \frac{1}{k!} \left(1 - \frac{p(p-1) \cdots (p-k+1)}{p^k} \right) \geq 0$, on déduit que :

$$\|e^A - A_p\| \leq \sum_{k=0}^p \left(\frac{1}{k!} - \frac{\binom{p}{k}}{p^k} \right) \|A\|^k + \sum_{k=p+1}^{+\infty} \frac{\|A\|^k}{k!} = e^{\|A\|} - \left(1 + \frac{\|A\|}{p} \right)^p \xrightarrow{p \rightarrow +\infty} 0$$

Exercice 7.5. Soit $A \in \mathcal{M}_n(\mathbb{C})$ diagonalisable de valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_p$ ($1 \leq p \leq n$). Montrer que :

$$\forall t \in \mathbb{R}, e^{tA} = \sum_{k=1}^p \alpha_k e^{\lambda_k t} \prod_{\substack{j=1 \\ j \neq k}}^p (A - \lambda_j I_n)$$

où les $\alpha_k \in \mathbb{C}$ s'expriment en fonction des valeurs propres de A .

Solution. En reprenant les notations du théorème 7.3, le vecteur C_k est la projection de $y(0)$ sur $\ker(A - \lambda_k I_n)$. En désignant par p_k cette projection, on déduit alors du théorème 7.3 que $e^{tA} = \sum_{k=1}^p e^{\lambda_k t} p_k$. D'autre part, on sait que ces projec-

teurs s'écrivent $p_k = \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (A - \lambda_j I_n)$ où $\alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)}$ (exercice 2.2). Ce

qui donne en définitive $e^{tA} = \sum_{k=1}^p \alpha_k e^{\lambda_k t} \prod_{\substack{j=1 \\ j \neq k}}^p (A - \lambda_j I_n)$.

Exercice 7.6. Soit $A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{C})$.

1. Calculer e^{tA} pour tout $t \in \mathbb{R}$.
2. Résoudre le système différentiel $y' = Ay$

Solution.

1. Le polynôme caractéristique de A est $\chi_A(X) = X(X-1)^3$. L'espace propre associé à la valeur propre 0 est $E_0 = \mathbb{C} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$ et l'espace propre associé à la valeur propre 1 est $E_1 = \mathbb{C} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \oplus \mathbb{C} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. La matrice A n'est donc pas diagonalisable et comme $\mathbb{C}^4 = \ker(A) \oplus \ker(A - I_4)^3$, on déduit que :

$$2 < \dim(\ker(A - I_4)^2) \leq \dim(\ker(A - I_4)^3) = 3$$

Donc $\ker(A - I_4)^2$ est de dimension 3 et $\mathbb{C}^4 = \ker A \oplus \ker(A - I_4)^2$ (le polynôme minimal de A est $\pi_A(X) = X(X-1)^2$).

Pour tout $x \in \ker A$ et tout $y \in \ker(A - I_4)^2$, on a $e^{tA}x = x$ et $e^{tA}y = e^{te^{(A-I_4)}y} = e^t(y + t(A - I_4)y)$. Tout vecteur $z \in \mathbb{C}^4$ s'écrit $z = x + y$ où $x = (A - I_4)^2 z \in \ker A$, $y = A(2I_4 - A)z \in \ker(A - I_4)^2$ (l'identité de Bézout $(-X + 2)X + (X - 1)^2 = 1$ permet d'obtenir les projections sur les sous espaces caractéristiques) et :

$$\begin{aligned} e^{tA}z &= x + e^t(y + t(A - I_4)y) \\ &= (A - I_4)^2 z + e^t(A(2I_4 - A)z + t(A - I_4)A(2I_4 - A)z) \end{aligned}$$

Écrivant que $A(2I_4 - A) = A - A(I_4 - A)$ et :

$$(A - I_4)A(2I_4 - A) = -A(A - I_4)^2 + A(A - I_4) = A(A - I_4)$$

on déduit que $e^{tA}z = (A - I_4)^2 z + e^t(Az + (t-1)A(A - I_4)z)$. Enfin de

$$(A - I_4)^2 = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, (A - I_4)A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ on déduit}$$

$$\text{que } e^{tA} = \begin{pmatrix} e^t & 0 & 1 - e^t & e^t - 1 \\ 0 & e^t & te^t & 0 \\ 0 & 0 & e^t & 0 \\ 0 & 0 & e^t - 1 & 1 \end{pmatrix}.$$

2. Les solutions de $y' = Ay$ sont alors données par :

$$y(t) = e^{tA} y_0 = \begin{pmatrix} \alpha e^t + (\gamma - \delta)(1 - e^t) \\ (\beta + \gamma t)e^t \\ \gamma e^t \\ \gamma(e^t - 1) + \delta \end{pmatrix}$$

Exercice 7.7. Soit $A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{C})$. Calculer e^{tA} , en utilisant l'algorithme du paragraphe 7.3.

Solution. On a $\chi_A(X) = X(X-1)^3$, donc $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = \lambda_4 = 1$, $B_1 = I_4$, $B_2 = A$, $B_3 = (A - I_4)A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ et $B_4 = (A - I_4)^2 A = 0$ (le polynôme minimal de A est $X(X-1)^2$). La solution de (7.7) est donnée par :

$$\begin{cases} y_1(t) = 1 \\ y_2(t) = e^t - 1 \\ y_3(t) = (t-1)e^t + 1 \\ y_4(t) = \left(\frac{(t-1)^2 + 1}{2} \right) e^t - 1 \end{cases}$$

(le calcul de y_4 n'est pas nécessaire puisque $B_4 = 0$). On a alors :

$$e^{tA} = I_4 + (e^t - 1)A + ((t-1)e^t + 1)(A - I_4)A = \begin{pmatrix} e^t & 0 & 1 - e^t & e^t - 1 \\ 0 & e^t & te^t & 0 \\ 0 & 0 & e^t & 0 \\ 0 & 0 & e^t - 1 & 1 \end{pmatrix}$$

Exercice 7.8. En utilisant les résultats de l'exercice 1.7, calculer l'exponentielle de la matrice $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ d'ordre n supérieur ou égal à 3 définie par :

$$\forall i \in \{1, \dots, n\}, \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, \dots, n\} \setminus \{i\} \end{cases}$$

où α, β sont des nombres complexes donnés.

Solution. Pour $\alpha = 0$, on a $A(\alpha, \beta) = \beta I_n$ et $e^{A(\alpha, \beta)} = e^\beta I_n$. On suppose que α est non nul et dans ce cas on a (exercice 1.7) :

$$\forall k \in \mathbb{N}, A(\alpha, \beta)^k = \frac{1}{n} (\lambda_2^k A(1, 1) - \lambda_1^k A(1, 1 - n))$$

avec $\lambda_1 = \beta + (n-1)\alpha$ et $\lambda_2 = \beta - \alpha$. Ce qui donne :

$$e^{A(\alpha, \beta)} = \frac{1}{n} (e^{\lambda_2} A(1, 1) - e^{\lambda_1} A(1, 1 - n))$$

Exercice 7.9. Montrer que le système différentiel $y' = Ay$ a toutes ses solutions bornées sur $\mathbb{R}$ si, et seulement si, A est diagonalisable avec toutes ses valeurs propres imaginaires pures.

Solution. Si A est diagonalisable à valeurs propres imaginaires pures, toute solution de $y' = Ay$ est alors de la forme $y(t) = \sum_{k=1}^p e^{i\mu_k t} C_k$ avec les μ_k réels

et pour tout $t \in \mathbb{R}$ on a $\|y(t)\| \leq \sum_{k=1}^p |e^{i\mu_k t}| \|C_k\| = \sum_{k=1}^p \|C_k\| = M$, c'est-à-

dire que toute solution de $y' = Ay$ est bornée. Si A admet une valeur propre λ de partie réelle non nulle, pour tout vecteur propre associé $x \in \mathbb{C}^n \setminus \{0\}$, la fonction y définie par $y(t) = e^{\lambda t} x$ est alors une solution de $y' = Ay$ telle que $\|y(t)\| = e^{\operatorname{Re}(\lambda)t} \|x\|$ pour tout $t \in \mathbb{R}$ et pour $\operatorname{Re}(\lambda) > 0$ [resp. $\operatorname{Re}(\lambda) < 0$] on a $\lim_{t \rightarrow +\infty} \|y(t)\| = +\infty$ [resp. $\lim_{t \rightarrow -\infty} \|y(t)\| = +\infty$], c'est donc une solution non bornée. Si A n'est pas diagonalisable, il existe alors une valeur propre λ de A telle que $\ker(A - I_n) \subsetneq \ker(A - I_n)^2$ et on peut trouver un vecteur non nul $x \in \mathbb{C}^n$ tel que $(A - I_n)x \neq 0$ et $(A - I_n)^2 x = 0$. La solution du problème de Cauchy $y' = Ay$, $y(0) = x$ est alors $y(t) = e^{\lambda t} e^{t(A - I_n)} x = e^{\lambda t} (x + t(A - I_n)x)$ et on a $\|y(t)\| \geq e^{\operatorname{Re}(\lambda)t} (|t| \|(A - I_n)x\| - \|x\|)$, c'est donc une solution non bornée. On a donc ainsi montré que le système différentiel $y' = Ay$ a toutes ses solutions bornées sur $\mathbb{R}$ si, et seulement si, A est diagonalisable avec toutes ses valeurs propres imaginaires pures.

Exercice 7.10. Montrer que si $A \in \mathcal{M}_n(\mathbb{C})$ est telle que $A^* = -A$, toutes les solutions de $y' = Ay$ sont alors bornées sur $\mathbb{R}$.

Solution. Si $A^* = -A$, la matrice iA est alors hermitienne, donc diagonalisable à valeurs propres réelles. Les valeurs propres de A sont donc imaginaires pures et avec le résultat de l'exercice 7.9 on déduit que toutes les solutions de $y' = Ay$ sont bornées sur $\mathbb{R}$.

Exercice 7.11. Donner une condition nécessaire et suffisante sur la matrice A pour que le système $y' = Ay$ ait toutes ses solutions bornées sur $\mathbb{R}^+$.

Solution. On a vu avec l'exercice 7.9 que s'il existe une valeur propre λ de A de partie réelle strictement positive, on peut alors construire une solution de $y' = Ay$ non bornée sur $\mathbb{R}^+$. Supposons que toutes les valeurs propres de A soient de parties réelles négatives ou nulles. On note $\lambda_1, \dots, \lambda_r$ les valeurs propres de parties réelles nulles (quand il en existe) et $\lambda_{r+1}, \dots, \lambda_p$ celles de parties réelles strictement

négatives (quand il en existe), les λ_k étant deux à deux distincts. Le polynôme minimal de A s'écrit alors $\pi_A(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$ et toute solution de $y' = Ay$

est de la forme $y(t) = \sum_{k=1}^p e^{\lambda_k t} P_k(t)$ avec $P_k(t) = \sum_{j=0}^{\alpha_k-1} \frac{t^j}{j!} (A - \lambda_k I_n)^j x_k$. Une telle

fonction est bornée sur $\mathbb{R}^+$ si, et seulement si, $\alpha_1 = \dots = \alpha_r = 1$. En définitive, toutes les solutions de $y' = Ay$ sont bornées sur $\mathbb{R}^+$ si, et seulement si, toutes les valeurs propres de A sont de parties réelles négatives ou nulles, celles de partie réelle nulle étant racines simples du polynôme minimal de A . Ce qui est encore équivalent à dire que la matrice A est semblable à une matrice de l'une des trois formes

suivantes $\begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & e^{i\theta_n} \end{pmatrix}, \begin{pmatrix} e^{i\theta_1} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \ddots & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & e^{i\theta_r} & 0 & \dots & 0 \\ 0 & \dots & 0 & J_{r+1} & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \ddots & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & J_p \end{pmatrix},$ ou

$\begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & J_p \end{pmatrix}$ avec $J_k = \begin{pmatrix} \lambda_k & \varepsilon_1 & \dots & 0 \\ 0 & \lambda_k & \varepsilon_2 & \dots & 0 \\ \cdot & \cdot & \ddots & \ddots & \cdot \\ 0 & \dots & 0 & \lambda_k & \varepsilon_{n-1} \\ 0 & \dots & 0 & \lambda_k & \end{pmatrix}, \operatorname{Re}(\lambda_k) < 0,$

$\varepsilon_i \in \{0, 1\}$.

Exercice 7.12. Pour A et B dans $\mathcal{M}_n(\mathbb{C})$ on désigne par $\Phi_{A,B}$ l'application de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$ définie par :

$$\Phi_{A,B}(X) = AX + XB$$

L'espace vectoriel $\mathcal{M}_n(\mathbb{C})$ est muni d'une norme matricielle induite par une norme vectorielle sur $\mathbb{C}^n$.

1. Donner les solutions du problème de Cauchy :

$$\begin{cases} Y'(t) = \Phi_{A,B}(Y)(t) & (t \in \mathbb{R}) \\ Y(0) = C \end{cases} \quad (7.13)$$

où C est donnée dans $\mathcal{M}_n(\mathbb{C})$ (on peut utiliser $Z(t) = Y(t)e^{-tB}$).

2. On suppose que toutes les valeurs propres de A et B sont de parties réelles strictement négatives, on se donne une matrice $C \in \mathcal{M}_n(\mathbb{C})$ et Y est la solution de (7.13).

(a) Montrer qu'il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que $\|Y(t)\| \leq \beta e^{-\alpha t}$ pour tout $t \in \mathbb{R}$.

(b) Montrer qu'il existe $X \in \mathcal{M}_n(\mathbb{C})$ telle que $C = \Phi_{A,B}(X)$.

(c) En déduire que $\Phi_{A,B}$ est un isomorphisme.

Solution.

1. Le problème de Cauchy (7.13) admet une unique solution $Y \in \mathcal{C}^\infty(\mathbb{R}, \mathcal{M}_n(\mathbb{C}))$.
On définit $Z \in \mathcal{C}^\infty(\mathbb{R}, \mathcal{M}_n(\mathbb{C}))$ par $Z(t) = Y(t)e^{-tB}$ et pour tout $t \in \mathbb{R}$ on a :

$$Z'(t) = (AY(t) + Y(t)B)e^{-tB} - Y(t)Be^{-tB} = AZ(t)$$

avec $Z(0) = C$. On a donc $Z(t) = e^{tA}C$ et la solution du problème (7.13) est $Y(t) = e^{tA}Ce^{tB}$.

2.

- (a) En reprenant les notations de la démonstration du théorème 7.4, on a :

$$\|e^{tA}\| = \left\| \sum_{k=1}^p e^{\lambda_k t} \sum_{j=0}^{\alpha_k-1} \frac{t^j}{j!} (A - \lambda_k I_n)^j p_k \right\| \leq e^{-\mu t} \nu \sum_{j=0}^{n-1} \frac{t^j}{j!}$$

où $-\mu = \max_{1 \leq k \leq p} \operatorname{Re}(\lambda_k) < 0$ et $\nu = \max_{\substack{1 \leq k \leq p \\ 0 \leq j \leq \alpha_k-1}} \|A - \lambda_k I_n\|^j \|p_k\|$. Avec

$\lim_{t \rightarrow +\infty} e^{\frac{\mu}{2}t} \|e^{tA}\| = 0$, on déduit qu'il existe $\beta_1 > 0$ tel que $\|e^{tA}\| \leq e^{-\alpha_1 t} \beta_1$,

où $\alpha_1 = \frac{\mu}{2}$. On a une majoration analogue pour la matrice B avec des constantes α_2 et β_2 . On déduit alors la majoration de la solution de (7.13) :

$$\forall t \in \mathbb{R}, \|Y(t)\| \leq \|e^{tA}\| \|e^{tB}\| \|C\| \leq e^{-\alpha t} \beta$$

avec $\alpha = \alpha_1 + \alpha_2$ et $\beta = \beta_1 \beta_2 \|C\|$.

- (b) Soit Y solution de (7.13). En intégrant on a :

$$Y(t) = A \int_0^t Y(u) du + \int_0^t Y(u) du B + C$$

De $\|Y(t)\| \leq e^{-\alpha t} \beta$ on déduit que l'intégrale $\int_0^{+\infty} Y(u) du$ est convergente et que $\lim_{t \rightarrow +\infty} Y(t) = 0$. En faisant tendre t vers l'infini dans l'identité ci-dessus, on déduit que :

$$-C = A \int_0^{+\infty} Y(u) du + \int_0^{+\infty} Y(u) du B$$

c'est-à-dire $C = \Phi_{A,B}(X)$ avec $X = -\int_0^{+\infty} Y(u) du \in \mathcal{M}_n(\mathbb{C})$.

- (c) Le résultat de la question précédente se traduit en disant que $\Phi_{A,B}$ est surjective. Cette application étant un endomorphisme de $\mathcal{M}_n(\mathbb{C})$ on déduit que c'est un isomorphisme. Une autre démonstration consiste à utiliser le fait que $\operatorname{sp}(\Phi_{A,B}) = \operatorname{sp}(A) + \operatorname{sp}(B)$.

Exercice 7.13. Montrer que les fonctions $u(t) = \int_0^{+\infty} e^{-x^2} \cos(tx^2) dx$ et $v(t) = \int_0^{+\infty} e^{-x^2} \sin(tx^2) dx$ sont bien définies, de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et solutions d'un système différentiel que l'on résoudra.

Solution. De la convergence de l'intégrale $\int_0^{+\infty} e^{-x^2} dx$ on déduit que la fonction $w : t \mapsto \int_0^{+\infty} e^{(it-1)x^2} dx$ est bien définie sur tout $\mathbb{R}$. Il en résulte que les fonctions $u = \operatorname{Re}(w)$ et $v = \operatorname{Im}(w)$ sont bien définies sur $\mathbb{R}$. En dérivant formellement on $w'(t) = \int_0^{+\infty} ix^2 e^{(it-1)x^2} dx$. De la convergence de l'intégrale $\int_0^{+\infty} x^2 e^{-x^2} dx$ on déduit que la fonction $y : t \mapsto \int_0^{+\infty} ix^2 e^{(it-1)x^2} dx$ est bien définie sur tout $\mathbb{R}$. Pour $t \in \mathbb{R}$ et $h \in \mathbb{R}^*$, on a :

$$\frac{w(t+h) - w(t)}{h} - y(t) = \int_0^{+\infty} e^{(it-1)x^2} \left(\frac{e^{ihx^2} - 1}{h} - ix^2 \right) dx$$

avec :

$$\left| \frac{e^{ihx^2} - 1}{h} - ix^2 \right| = \left| \frac{1}{h} \sum_{n \geq 2} \frac{(ihx^2)^n}{n!} \right| \leq |h| x^4 \sum_{n \geq 2} \frac{(|h|x^2)^{n-2}}{n!} \leq |h| x^4 e^{|h|x^2}$$

et pour $|h| \leq \frac{1}{2}$, $\left| \frac{e^{ihx^2} - 1}{h} - ix^2 \right| \leq |h| x^4 e^{\frac{x^2}{2}}$, ce qui donne en définitive :

$$\left| \frac{w(t+h) - w(t)}{h} - y(t) \right| \leq |h| \int_0^{+\infty} x^4 e^{-\frac{x^2}{2}} dx \xrightarrow{h \rightarrow 0} 0$$

La fonction w est donc dérivable sur $\mathbb{R}$ de dérivée y . On en déduit alors que les fonctions u et v sont dérivables sur $\mathbb{R}$ de dérivées :

$$u'(t) = - \int_0^{+\infty} x^2 e^{-x^2} \sin(tx^2) dx, \quad v'(t) = \int_0^{+\infty} x^2 e^{-x^2} \cos(tx^2) dx$$

Une intégration par parties (où on dérive $e^{(it-1)x^2}$ par rapport à la variable x) donne :

$$w(t) = -2(t+i)w'(t) \quad (7.14)$$

ce qui équivaut au système :

$$\begin{cases} u(t) = -2(tu'(t) - v'(t)) \\ v(t) = -2(u'(t) + tv'(t)) \end{cases}$$

encore équivalent au système différentiel :

$$\begin{cases} 2(1+t^2)u'(t) = -tu(t) - v(t) \\ 2(1+t^2)v'(t) = u(t) - tv(t) \end{cases}$$

On a de plus les conditions initiales $u(0) = \int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ et $v(0) = 0$. L'équation (7.14) entraîne $(w^2(t+i))' = 0$, soit $w^2(t+i) = C$, où C est une constante complexe. Avec $w(0) = \frac{\sqrt{\pi}}{2}$, on déduit que $C = i\frac{\pi}{4}$ et :

$$u^2 - v^2 = \frac{\pi}{4} \frac{1}{1+t^2}, \quad 2uv = \frac{\pi}{4} \frac{t}{(1+t^2)} \quad (7.15)$$

C'est à dire que $(u^2, -v^2)$ est solution de l'équation polynomiale de degré 2 :

$$r^2 - \frac{\pi}{4} \frac{1}{1+t^2} r - \left(\frac{\pi}{8}\right)^2 \frac{t^2}{(1+t^2)^2} = 0$$

ce qui donne $u^2(t) = \frac{\pi}{8} \frac{1 + \sqrt{1+t^2}}{(1+t^2)}$ et $v^2(t) = \frac{\pi}{8} \frac{\sqrt{1+t^2} - 1}{(1+t^2)}$. La fonction continue u ne s'annule jamais, elle garde donc un signe constant et ce signe est celui de $u(0)$, c'est-à-dire le signe positif. De (7.15) on déduit que la fonction uv est impaire du signe de t . Ce qui donne en définitive :

$$u(t) = \frac{1}{2} \sqrt{\frac{\pi}{2}} \sqrt{\frac{1 + \sqrt{1+t^2}}{(1+t^2)}}$$

$$v(t) = -\frac{\text{signe}(t)}{2} \sqrt{\frac{\pi}{2}} \sqrt{\frac{\sqrt{1+t^2} - 1}{(1+t^2)}}$$

Exercice 7.14. Déterminer toutes les solutions dans $\mathcal{M}_n(\mathbb{C})$ de l'équation $e^X = I_n$.

Solution. Soit $X \in \mathcal{M}_n(\mathbb{C})$ telle que $e^X = I_n$. On a la décomposition de Dunford $X = D + V$ avec D diagonalisable, V nilpotente et $DV = VD$. Cette décomposition permet d'obtenir celle de e^X , soit $e^X = e^D + e^D(e^V - I_n)$ avec e^D diagonalisable et $e^D(e^V - I_n)$ nilpotente (lemme 7.6). Avec l'unicité de cette décomposition, on déduit que $e^D = I_n$ et $e^V = I_n$. La première égalité avec D diagonalisable entraîne que D est semblable à une matrice $2i\pi K$ où K est une matrice diagonale à coefficients entiers et la deuxième égalité entraîne $V = 0$. En effet si p est le plus petit entier naturel non nul tel que $V^p = 0$, en supposant $p \geq 2$, l'égalité $e^V = I_n + \sum_{k=1}^{p-1} \frac{1}{k!} V^k = I_n$ entraîne $\sum_{k=1}^{p-1} \frac{1}{k!} V^k = 0$ et en multipliant par V^{p-2} , on aboutit à $V^{p-1} = 0$ ce qui est en contradiction avec la définition de p . On a donc $p = 1$ et $V = 0$. On a donc ainsi montré que les solutions dans $\mathcal{M}_n(\mathbb{C})$ de l'équation $e^X = I_n$ sont les matrices semblables à une matrice $2i\pi K$ où K est une matrice diagonale à coefficients entiers.

Bibliographie

- [1] R. BURLISCH, J. STOER — *Introduction to numerical analysis*. Springer-Verlag (1980)
- [2] A. CHAMBERT-LOIR, S. FERMIGIER, V. MAILLOT — *Exercices de Mathématiques pour l'Agrégation. Analyse 1*. Masson (1995).
- [3] P. G. CIARLET — *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson (1982).
- [4] J. DIEUDONNE — *Éléments d'analyse. Fondements de l'analyse moderne*. Gauthier-Villars (1972).
- [5] F. R. GANTMACHER. *Théorie des matrices I*. Dunod (1966).
- [6] F. R. GANTMACHER. *Théorie des matrices II*. Dunod (1966).
- [7] S. GONNORD, N. TOSEL — *Topologie et analyse fonctionnelle*. Ellipses (1996).
- [8] A. GRAMAIN — *Géométrie élémentaire*. Hermann (1997).
- [9] R. A. HORN, C. A. JOHNSON — *Matrix analysis*. Cambridge University Press (1985).
- [10] A. S. HOUSEHOLDER — *The theory of matrices in numerical analysis*. Dover (1975).
- [11] P. LASCAUX, R. THEODOR — *Analyse numérique matricielle appliquée à l'art de l'ingénieur*. Masson (1987).
- [12] R. MNEIMNE, F. TESTARD — *Introduction à la théorie des groupes de Lie classiques*. Hermann (1986).
- [13] J. MOUSSA — *Une preuve élémentaire des formules de Newton*. Revue de Mathématiques Spéciales (Novembre, Décembre 1995).
- [14] M. PARODI — *La localisation des valeurs caractéristiques des matrices et ses applications*. Gauthier-Villars (1959).
- [15] V. PRASOLOV. *Problèmes et théorèmes d'algèbre linéaire*. Cassini (2007).
- [16] J. E. ROMBALDI — *Algorithmique numérique et Ada*. Masson (1994).
- [17] J. E. ROMBALDI — *Problèmes corrigés d'analyse numérique*. Masson (1996).
- [18] J. E. ROMBALDI. *Éléments d'analyse réelle, deuxième édition*. EDP Sciences (2019).
- [19] F. RELICH — *Perturbation theory of eigenvalues problems*. Gordon and Breach (1950).

- [20] F. RIDEAU — *Exercices de calcul différentiel*. Hermann (1979).
- [21] P. TAUVEL — *Exercices de mathématiques pour l'agrégation*. Algèbre 2. Masson (1994).
- [22] P. TAUVEL — *Algèbre pour l'agrégation interne*. Masson (1996).
- [23] C. TISSERON — *Géométries affine, projective et euclidienne*. Hermann (1983).
- [24] B. FLANNERY, W. PRESS, S. TEUKOLSKY, W. VETTERLING — *Numerical recipes*. Cambridge University Press (1988).

Index

équations différentielles, 240

adjointe (matrice), 48

Bézout, 2

Bernoulli, 212

Birkhoff, 153

Carathéodory, 145

Cauchy (problème de), 229

Cauchy-Schwarz (inégalité de), 35, 43, 47

Cayley-Hamilton, 13

Cholesky, 175

conditionnement, 95

convexe, 143

Courant-Fischer, 100

Cramer, 161

Crout (méthode de), 174

décomposition des noyaux, 17

décomposition LR, 173

décomposition polaire, 51, 79

décomposition singulière, 64

déflation, 212

déterminants principaux, 171

descente (méthode de), 190

diagonale strictement dominante, 8

diagonalisable (endomorphisme), 33

diagonalisable (matrice), 33

différentiable (fonction), 188

dilatation (matrice de), 164

doublement stochastique (matrice), 142

Dunford, 20

enveloppe convexe, 144

espaces caractéristiques, 18

euclidien (espace), 35

exponentielle (d'une matrice), 233

fonctionnelle quadratique, 189

Frobenius (matrice de), 172

Gauss (pivots de), 168

Gauss-Jordan, 176

Gauss-Seidel, 179

Gelfand, 90

Gerschgorin (disques de), 8

Gerschgorin-Hadamard, 7

Givens, 218

gradient (vecteur), 188

gradient conjugué (méthode du), 193

Gram-Schmidt, 37, 48

groupe topologique, 77

Hadamard (inégalité de), 63

hausdorffien, 96

hermitien (espace), 47

hermitienne (matrice), 49

hermitienne définie positive (matrice), 49

hermitienne positive (matrice), 49

Hessenberg (matrice de), 223

Hilbert (matrice de), 113, 163

Householder, 89

Householder (méthode de), 44

Householder (matrice de), 44

hyper-quadratique, 189

hyperplan affine, 143

hyperplan d'appui, 148

idéal annulateur, 2

indice (d'une valeur propre), 5

irréductible (matrice), 135

isotrope (cône), 43

Jacobi, 178, 213

Jordan, 52

Jordan (forme réduite de), 55

- Krein-Milman, 151
- Krylov, 15
- Leverrier, 17
- logarithme matriciel, 247
- méthode itérative, 177
- matrice élémentaire, 165
- matrice compagnon, 10
- matrice de Frobenius, 10
- mesure (d'un angle), 36
- Minkowski (inégalité de), 36, 47
- Newton (formules de), 15
- nilpotent, 6
- normale (matrice), 49
- norme matricielle, 73
- orientation, 168
- orthogonal (endomorphisme), 40
- orthogonale (famille), 37
- orthogonale (matrice), 38
- orthogonaux (vecteurs), 36, 48
- orthonormée (famille), 37
- Ostrowski, 10
- Ostrowski-Reich, 182
- parallélogramme, 36
- permutation (matrice de), 134, 166
- Perron-Frobenius, 129, 133, 138
- point extrême (d'un convexe), 149
- polyèdre, 143
- polynôme caractéristique, 3
- polynôme minimal, 2
- positive (matrice), 123
- primitive (matrice), 139
- produit scalaire euclidien, 35
- produit scalaire hermitien, 47
- projecteurs spectraux, 20
- projection orthogonale, 146
- puissance inverse, 212
- puissance itérée, 209
- QR (décomposition), 39
- réductible (matrice), 135
- résidu (vecteur), 189
- résultant, 83
- Rayleigh-Ritz (quotient de), 96
- Rayleigh-Ritz (théorème de), 97
- rayon spectral, 86
- relaxation (méthode de), 181
- relaxation par blocs, 187
- Richardson, 206
- rotation, 38
- Schur (norme de), 65
- signature, 174
- Souriau, 17
- sous espace cyclique, 13
- sous espace propre, 2, 3
- sous-matrices principales, 61, 171
- sous-multiplicative (norme), 76
- spectre, 2, 3
- stochastique (matrice), 141
- Sylvester (matrice de), 83
- symétrique (matrice), 42
- symétrique définie positive (matrice), 42
- symétrique positive (matrice), 42
- systèmes différentiels, 229, 242
- taux asymptotique de convergence, 178
- taux moyen de convergence, 178
- Tchebychev, 186
- transvection (matrice de), 164
- trigonalisable (endomorphisme), 31
- trigonalisable (matrice), 31
- unipotente (matrice), 247
- unitaire (matrice), 48
- unitaire (polynôme), 2
- valeur propre, 2, 3
- valeurs singulières, 87
- variation des constantes, 245
- vecteur propre, 2, 3
- Weyl, 101
- Wielandt, 141
- wronskien, 244
- Young-Varga, 185