Algèbre approfondie pratique mais théorique

Adrien Blanchet

25 octobre 2024

Introduction

Objectif pratique

L'objectif pratique de ce cours est de munir la lectrice/le lecteur d'un bagage précis et étayé pour pouvoir utiliser trois monstres sacrés des mathématiques appliquées :

- la diagonalisation ou plus généralement la réduction d'endomorphismes,
- les projections orthogonales,
- l'étude du signe d'une forme quadratique.

Ces trois objets sont indispensables à la résolution dans de nombreux domaines des applications des mathématiques. Nous expliquerons en quoi dans les trois sections qui vont suivre.

Réduction d'endomorphismes

La réduction d'endomorphisme consiste à simplifier la description d'endomorphismes. Cette théorie remonte au XVIII^{ème} siècle où Leonhard Euler et Joseph-Louis Lagrange découvrent l'importance des "axes principaux" qui deviendront dans notre terminologie les espaces propres. La diagonalisation émerge comme une clé de voûte, une technique qui transcende les complexités pour révéler la simplicité inhérente à certaines transformations linéaires. En effet, le procédé se ramène à la décomposition de l'espace vectoriel en une somme directe de droites vectorielles sur lesquelles l'endomorphisme se réduit à une homothétie. La décomposition en valeurs propres ouvre la porte à des perspectives inédites, transformant des systèmes en apparence intriqués en des structures élémentaires, et offrant ainsi un éclairage nouveau sur les phénomènes mathématiques.

La diagonalisation apparaît lorsque l'on cherche à calculer la puissance d'une matrice, ou les itérées d'une application linéaire. En statistique et en économétrie la réduction d'endomorphismes est un premier pas vers l'"analyse en composantes principales" qui est massivement utilisée dans les applications. Cependant les instruments que nous développerons pour la théorie (les valeurs propres, les vecteurs propres, les espaces propres, etc.) ont en eux même certaines caractéristiques fondamentales de la matrice. Par exemple, en théorie élémentaire des graphes cela permet de définir des groupes d'individus dans une matrice qui recense les interactions entre individus. La diagonalisation permet ainsi une compréhension plus claire des interactions économiques sous-jacentes. Les valeurs propres peuvent représenter des indicateurs cruciaux, tels que les taux de croissance ou les facteurs influençant la stabilité d'un système économique. Ainsi, notre exploration de la diagonalisation devient une quête pour éclairer les dynamiques économiques, fournissant des outils précieux pour l'analyse et la prévision.

(Mal)heureusement toutes les matrices ne sont pas diagonalisables, mais cette théorie peut être étendue à l'ensemble des matrices (au moins dans $\mathbb{C}$) en s'intéressant à la trigonalisation. Nous ne développerons pas dans ce cours la théorie complète de la trigonalisation mais nous verrons comment l'utiliser rigoureusement en dimension 3.

Les théorèmes qui seront utilisés dans les autres cours :

- Critère de dimension : Théorème 1.20
- Critère de multiplicité : Théorème 1.23

Projection orthogonale

En parallèle, notre parcours mathématique nous conduira à la découverte de la projection orthogonale, une opération élémentaire mais puissante qui trouve des applications dans des domaines

aussi variés que la résolution de systèmes linéaires et l'analyse des données. Cette technique, véritable joyau de l'algèbre, nous permettra de distinguer les composantes essentielles d'un espace vectoriel, jetant ainsi une lumière nouvelle sur la nature des transformations linéaires et de leurs conséquences.

La théorie que nous développerons ici formalise et étend la notion de projection graphique qui a été introduite au collège. Fondamentalement, la projection orthogonale permet de déterminer la meilleure approximation d'un vecteur dans un sous-espace vectoriel. On pourra ainsi définir la distance d'un vecteur à un sous-espace vectoriel. C'est une question absolument naturelle en statistique et encore plus en économétrie : dans un ensemble de données si une entrée est manquante par exemple, il est intéressant de compléter cette donnée par sa meilleure approximation dans l'espace vectoriel définie par les autres données.

La projection orthogonale est aussi cruciale en optimisation et dans de nombreux algorithmes d'étude de données en grandes dimensions puisque les projections orthogonales sont l'essence de nombreux algorithmes d'optimisation et de classification en statistique, en apprentissage automatique et en exploration des données : méthode des "moindre carrés", de régression linéaire, de régression sur la quartile, la régression non-paramétrique, les algorithmes de valeurs propres, la décomposition QR, la décomposition en valeurs singulières (traitement du signal météorologie), etc. Plus proche des mes travaux de recherche, la projection orthogonale sur l'intersection de convexes est le cœur de l'algorithme de Dykstra qui est utilisée pour résoudre des problèmes de transport optimal.

Dans le domaine de l'économie spatiale, la projection orthogonale peut être utilisée pour optimiser l'allocation des ressources, minimisant les coûts tout en maximisant l'efficacité. Cette technique offre ainsi un cadre mathématique pour aborder des questions cruciales telles que la planification urbaine et la gestion des ressources.

Nous revisiterons pour la partie théorique la géométrie vue au Collège. La notion de produit scalaire s'étend ainsi en grande généralité aussi bien aux vecteurs de $\mathbb{R}^N$, qu'aux matrices, aux polynômes et même aux fonctions. Les structures de perpendicularité, d'angle s'étendent alors à tous ces ensemble. Nous revisiterons dans une grande généralité certains théorèmes fameux comme le théorème de Pythagore, l'inégalité de Cauchy-Schwarz ou l'inégalité triangulaire.

Les concepts principaux qui seront utilisés dans les autres cours :

- Définition d'une base orthonormée : Définition 4.10
- Procédé de Gram-Schmidt : Démonstration du Théorème 4.13
- Expression de la projection orthogonale dans le cas d'une base orthogonale : Théorème 4.30
- Théorème de projection orthogonale : Théorème 4.36

Signe d'une forme quadratique

Une forme quadratique est un polynôme homogène de degré 2. Si l'étude d'une forme quadratique est très proche de la géométrie, c'est essentiellement pour l'étude locale des points critiques qui apparaîtront dans les problèmes d'optimisation que l'on étudiera principalement le signe d'une forme quadratique. Cet outil permet en effet d'étendre la notion de positivité de la dérivée seconde à des fonctions d'un espace de dimension supérieure ou égale à 2. On définira dans le même temps la notion de positivité d'une matrice carrée qui n'est pas un concept absolument évident à première vue.

Cette notion apparaît aussi pour étudier la matrice de covariance d'une distribution de probabilité multivariée ou pour certains problèmes d'optimisation de portefeuille. De nombreux problèmes de résolution de systèmes linéaires, ou d'inversion de matrice ont des algorithmes particulièrement efficaces lorsque la matrice est définie positive : optimisation semi-définie positive, optimisation conique, algorithme des points intérieurs, lagrangien augmenté, méthode des faisceaux, décomposition de Cholesky, simulation de Monte-Carlo, etc. Ainsi toute une branche de la recherche en optimisation consiste à sa ramener à des problèmes comportant des matrices définies positives.

En plongeant dans le signe d'une forme quadratique, nous ouvrons la porte à une compréhension approfondie des choix économiques individuels et collectifs. Les formes quadratiques ap-

paraissent souvent dans la modélisation des fonctions d'utilité ou des coûts, et comprendre leur signe offre des perspectives cruciales. L'analyse du signe d'une forme quadratique peut aider à évaluer la nature des décisions individuelles ou des politiques publiques. Par exemple, déterminer si une fonction d'utilité est concave ou convexe peut avoir des implications significatives sur les comportements de consommation et les politiques fiscales.

Les concepts principaux qui seront utilisés dans les autres cours :

- Forme positive et négative : Définition 3.24
- Méthode de réduction en carrés de Gauss : Section 3.4.2
- Signature : critère des valeurs propres : Proposition 5.8
- Signature : Critère des valeurs propres en dimension 2 : Proposition 5.9
- Signature : Méthode des mineurs principaux : Proposition 5.10

Objectif théorique

Le cours est construit dans l'objectif de donner à l'étudiant.e appliqué.e les outils qui lui seront utiles pour les applications et suivre confortablement les cours des années à venir. Nous essaierons d'y parvenir en allant aussi directement que possible vers la maîtrise de ces outils tout en restant parfaitement rigoureux. Ce cours reste donc intrinsèquement un cours de mathématiques et l'ensemble des résultats nécessaires sera démontré dans toute leur rigueur mathématique à un niveau parfaitement accessible à l'étudiant.e motivé.e de niveau L2.

Nous illustrerons les concepts et résultats par des exemples concrets et chaque chapitre comporte une section d'exercice corrigés. Nous nous efforcerons également de présenter des contre-exemples, car ils mettent en évidence la limite d'un résultat. Nous essaierons également d'exposer les extensions possibles d'un concept et les difficultés à l'étendre. Nous ferons notre possible pour donner aux étudiant.e.s une vision géométrique et l'intuition derrière les concepts, car nous considérons que cela donne un meilleur aperçu qui aide à comprendre les aspects techniques. Nous ferons dans le même temps de notre mieux pour situer les concepts mathématiques dans l'histoire et dans le paysage appliqué : pourquoi ils étaient importants lorsque la théorie a été développée, où ils sont actuellement utilisés et quelles sont les orientations prometteuses en termes de recherche ou de débouchés professionnels.

Précisons donc, qu'en plus de permettre de maîtriser les techniques utiles pour les applications, ce cours a pour but génial de faire des mathématiques en soi. Sur le chemin nous rencontrerons des résultats théoriques passionnants et surtout nous nous ébattons joyeusement dans ce monde merveilleux des mathématiques abstraites.

Objectif philosophique

Les mathématiques sont aussi souvent utilisées comme un processus de sélection. Passionné de philosophie au lycée, je me désespérais de mes notes moyennes en dissertation alors que mon travail en mathématiques se traduisait inmanquablement par des notes à l'image du travail fourni. Je garde de ces traumatismes une certaine frustration mais aussi une passion débordante pour la philosophie, une tendance à m'y vautrer dès qu'un auditoire m'est acquis. Reste que les mathématiques me semblent effectivement un outil très efficace pour évaluer la rigueur des étudiants, pour sélectionner ceux qui sont les plus aptes à maîtriser rapidement de nouveaux concepts et pour identifier les esprits les plus brillants, compétences qui sont très valorisées sur le marché de l'emploi ainsi que dans la recherche académique.

Dans les années 30, certains textes mathématiques fondamentaux furent écrits par Nicolas Bourbaki. Afin de fournir les éléments biographiques requis pour publier ses travaux dans les Comptes rendus de l'Académie des sciences André Weil le présente comme "Monsieur Nicolas Bourbaki, membre canonique de l'académie royale de Poldévie, grand maître de l'ordre des compacts, conservateur des uniformes, lord protecteur des filtres". Si ce personnage s'est révélé n'être qu'un personnage imaginaire fruit de l'imagination fertile de mathématiciens de l'École Normale

Supérieure, il a durablement marqué la culture mathématiques et a largement été à la source de l'excellence mathématique française. Ce groupe considérait problématique :

- l'émiettement des mathématiques en spécialités étanches ;
- la pré-éminence d'une analyse foisonnante mais manquant de rigueur ;
- l'ignorance de branches actives à l'étranger.

Nous sommes redevables à Bourbaki d'un travail de clarification des concepts, de précision dans la formulation, d'une recherche — parfois aride — de structure, de classification systématique et exhaustive des mathématiques. Cette approche a une influence considérable dans de nombreuses autres disciplines : notamment, l'ouliipo (ouvroir de littérature potentielle), le structuralisme lacanien, ou celui de Lévi-Strauss en ethnologie mais aussi en politique, en psychanalyse, en linguistique etc. En économie, Gérard Debreu, a été bercé dans cette abstraction féconde lors de son passage à la rue d'Ulm, son approche axiomatique. Ainsi il publie en 1954 une contribution majeure intitulée *Existence of an Equilibrium for a Competitive Economy* en collaboration avec Kenneth Arrow qui caractérise les conditions d'existence d'un équilibre général en économie de marché en se fondant sur une méthode topologique. En 1959, il publie *La Théorie de la valeur, une analyse axiomatique de l'équilibre économique*, qui constitue une reformulation en termes rigoureux de la théorie de l'équilibre général inaugurée par Léon Walras en 1874. Cette orientation axiomatique des sciences économiques, avec ses contradictions et ses contradicteurs trop tus, est fondatrice de la façon de faire des sciences économiques à TSE et dans le monde.

Au lycée pour un exposé, j'avais contacté Adrien Douady¹, membre du groupe Bourbaki, qui m'avait donné rendez-vous dans un café du Luxembourg. Alors que je l'attendais, je vis arriver un homme timide, la barbe fournie, habillé de rouge de la chemise de bûcheron aux santiags pointues, le sourire joueur, c'était Adrien Douady. Lorsque je lui ai demandé pourquoi il avait choisi de faire des mathématiques il m'a expliqué qu'il avait grandi pendant la guerre et que le monde mathématique était pour lui un refuge rassurant où tout est à sa place. Un de ses élèves témoigne "Si je me souviens si bien de son cours après toutes ces années, ce n'est pas parce qu'il venait pieds nus dans l'amphi, mais à cause de la façon si singulière qu'il avait de faire des mathématiques. [...] Avec lui j'ai compris ce que veut dire comprendre. Pour lui quelque chose était compris quand cela se mettait à aller de soi, à couler de source : il comprenait avec son corps". L'approche bourbakiste consiste à décortiquer dans un concept la structure minimale nécessaire pour développer la théorie. Si le formalisme développé ici n'a pas le degré d'abstraction de ces éminents maîtres, il me semblait important de leur rendre hommage pour tenter de nous placer dans cette lignée.

Ce cours a dans cet esprit finalement pour ambition de donner à l'étudiant.e intéressé.e un formidable espace de jeu et d'exploration. Certains passent des heures à s'acharner sur de subtils sudoku, d'autres payent pour quelques heures d'escape games commerciaux. Parcourez ce polycopié, comprenez en l'essence, cherchez les démonstrations, faites les exercices, émerveillez vous des ramrques hors-programme, émerveillez-vous sans cesse et réfléchissez. Ce monde mathématiques est effectivement parfait, et dans un monde de plus en plus incertain où les injustices font rage et les révoltes grondent sans exploser, il est très apaisant de se plonger entre deux luttes, deux réflexions et deux prises de conscience, dans ce monde mathématiques où tout est juste, à sa place, prêt à jouer. Nous vous invitons à explorer les mystères et les merveilles de l'algèbre, où chaque page offre une invitation à la découverte et à l'illumination. Puisqu'à la fin les mathématiques sont un jeu, certain.e.s parmi les plus sages et les plus privilégié.e.s diront que la Vie elle-même est un jeu. Alors jouons.

Notation

- Lorsque A est $\mathbb{N}$ ou $\mathbb{R}$, A^* désigne l'espace A privé de l'élément $\{0\}$.
- $\mathcal{M}_N(\mathbb{K})$ est l'ensemble des matrices carrés de taille N dont les coefficients sont dans $\mathbb{K}$.
- $\{u, v\}$ est la famille ordonnée dont le premier élément est u et le second v .
- Vect désigne l'espace vectoriel engendré.

1. Adrien Douady https://fr.wikipedia.org/wiki/Adrien_Douady

- $\ker$ désigne le noyau d'une application linéaire comme d'une matrice.
- Im désigne l'image d'une application linéaire comme d'une matrice.
- id désigne l'application identité.

Programme

Afin de faciliter la lecture de ce polycopié chaque résultat et chaque démonstration est précédé d'un drapeau de couleur :

- ▶ signifie que ce point est à connaître absolument (en pratique l'étudiant.e qui ne connaît pas ce résultat ne mérite pas 10/20 aux contrôles),
- signifie qu'il est bon de maîtriser ce point (en pratique l'étudiant.e qui vise la note tout juste honorable de 15/20 aux contrôles doit étudier ces parties).
- signifie que cette partie est hors-programme. L'étudiant.e joueu.r.se (qui ira loin) peut s'aventurer dans ces parties dans un esprit noble de curiosité saine et sans danger.

L'ensemble des méthodes des exercices corrigés en fin de chapitre sont des exercices "types" qu'il est indispensable de connaître et de comprendre. L'étudiant.e qui sait reproduire sans erreur ces exercices s'assurera d'une note d'au moins 14/20.

Pour distinguer les étudiant.e.s les meilleur.e.s les contrôles contiennent, en général, un exercice dont la résolution nécessite d'avoir pris du recul sur le cours et de savoir utiliser les connaissances acquises pour résoudre un exercice qui ne fait pas partie des exercices types.

Table des matières

I	Réduction d'endomorphismes	11
1	Diagonalisation	13
1.1	Rappel sur les applications linéaires	17
1.1.1	Représentation matricielle	17
1.1.2	Changement de base	19
1.2	Diagonalisation et éléments propres	20
1.3	Théorème de diagonalisation : critère de dimension	24
1.4	Théorème de diagonalisation : critère de multiplicité	27
1.5	Applications	28
1.5.1	Application aux itérés d'une application linéaire	28
1.5.2	Application à la résolution de systèmes linéaires	28
1.5.3	Application à la résolution de suites linéaires d'ordre supérieur	29
1.5.4	Application à la géométrie d'une application linéaire	29
1.5.5	Application aux racines d'une matrices	30
1.6	Exercices	30
1.7	Corrigés des exercices	32
2	Trigonalisation	37
2.1	Polynôme d'endomorphisme	38
2.2	Théorème de trigonalisation	38
2.3	Théorème de Cayley-Hamilton	40
2.3.1	Application au calcul de l'inverse d'une matrice	41
2.3.2	Application à la détermination du polynôme minimal	42
2.4	Théorème de diagonalisation : critère du polynôme minimal	42
2.5	Technique de trigonalisation : méthode de Jordan dans $\mathcal{M}_3$	44
2.6	Exercices	48
2.7	Corrigés des exercices	49
II	Algèbre bilinéaire	53
3	Formes bilinéaires et formes quadratiques	55
3.1	Dual	56
3.2	Forme bilinéaire	60
3.3	Formes quadratiques	63
3.4	Signe d'une forme quadratique	66
3.4.1	Définitions	67
3.4.2	Méthode de réduction en carrés de Gauss	68
3.5	Diagonalisation de formes quadratiques	71
3.6	Exercices	73
3.7	Corrigés des exercices	75

4	Projection orthogonale dans un espaces Euclidien	83
4.1	Espaces Euclidiens	83
4.1.1	Produit scalaire	83
4.1.2	Inégalités fonctionnelles	85
4.1.3	Base orthonormée	87
4.1.4	Espace orthogonal	93
4.2	Projection orthogonale	96
4.2.1	Expression de la projection orthogonale	99
4.2.2	Expression matricielle de la projection orthogonale	100
4.2.3	Applications à la distance d'un vecteur à un sous-espace vectoriel	102
4.3	Exercices	103
4.4	Corrigés des exercices	105
III	Pour aller plus haut	113
5	Connections entre applications linéaires et formes bilinéaires et applications	115
5.1	Théorème spectral	115
5.2	Théorème de représentation de Riez-Fréchet	116
5.3	Loi d'inertie de Sylvester	117
5.4	Applications	118
5.4.1	Application du théorème spectral à la diagonalisation d'une matrice symétrique	118
5.4.2	Application du théorème de représentation de Riesz-Fréchet à la signature d'une forme quadratique	119
5.5	Exercices	122
5.6	Corrigés des exercices	123

Première partie

Réduction d'endomorphismes

Chapitre 1

Diagonalisation

Objectif du chapitre :

- déterminer si une matrice est diagonalisable,
- déterminer une matrice inversible P et une matrice diagonale D telles que $A = PDP^{-1}$,
- calculer la puissance d'une matrice.

Rappelons qu'une matrice peut représenter :

- une application linéaire,
- une matrice de covariance en probabilité,
- une forme bilinéaire (c'est l'objet de la deuxième partie du cours),
- mais aussi, par ex en sciences sociales, les interactions entre les individus d'une population,
- etc.

Nous nous concentrerons dans cette partie sur l'identification d'une matrice à son application linéaire, mais les espaces que nous allons étudier peuvent aussi donner des informations très importantes sur la structure sous-jacente de la matrice. L'espace propre, par ex, que nous allons définir dans un instant, peut donner des informations sur les groupes que constituent une population dans l'interprétation d'une matrice comme représentant les interactions entre des individus.

Nous allons nous intéresser dans ce chapitre (et le suivant) à des matrices carrées. L'application linéaire associée à une matrice carrée peut être vue comme un endomorphisme d'un espace vectoriel E . Considérons $f : E \rightarrow E$ où E est un $\mathbb{R}$ -espace vectoriel. Une application possible de ce chapitre consiste à calculer les itérées de f . En terme de matrice cela correspond à calculer des puissances de la matrice correspondante.

Exemple 1. *Deux individus sont en interaction. Nous allons modéliser cette relation de la façon suivante : Pour le premier individu, cette interaction consiste à fournir à son partenaire une certaine quantité de A (attention), B (bienveillance) et C (compassion) et d'en tirer une nouvelle quantité de A , B et C . Notons $f(A, B, C)$ la quantité de A , B et C que reçoit l'individu lorsqu'il en a fourni A , B et C . Nous considérerons, dans le cas d'un individu en situation de dépendance affective par exemple, que les quantités A , B et C ne dépendent que de cette relation. Soit a , b et c la quantité initiale de chacune de ces quantités que possède l'individu considéré au début de la relation. Nous aimerions savoir comment va évoluer cette relation, au sens où l'on aimerait savoir la quantité d'attention, de bienveillance et de compassion que va avoir cet individu après un grand nombre interactions. Si f est une application linéaire, mathématiquement cela consiste à calculer*

$$f(f \cdots f(a, b, c) = f^{(n)}(a, b, c) .$$

En terme matriciel, si on note M la représentation matricielle de f , il s'agit de calculer M^n .

Rappelons quand même que les relations entre individus sont bien plus complexes que ne le modélise bêtement cet exercice mathématiques et que l'individu est irrémédiablement amené à souffrir s'il ne s'occupe pas sérieusement de sa dépendance affective.

Le calcul de la puissance d'une matrice, ou des itérées d'une application linéaire est, avec vos connaissances actuelles, possible uniquement dans des cas très particuliers. Par exemple, lorsque le calcul des premières puissances de M laisse apparaître une récurrence évidente qu'il s'agit alors de prouver proprement.

Exemple 2. Pour tout $n \in \mathbb{N}^*$, calculer la puissance n -ème de

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

▷ On calcule

$$M^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad M^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \quad M^4 = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$$

Il semble raisonnable de conjecturer que pour tout $n \in \mathbb{N}^*$,

$$M^n := \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

Pour le prouver on peut le faire par récurrence. Soit pour $n \in \mathbb{N}^*$ la proposition :

$$\mathcal{P}(n) := "M^n := \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}"$$

- $\mathcal{P}(1)$ est vraie.
- Supposons que $\mathcal{P}(n)$ soit vraie à un rang n donné. On calcule alors

$$M^{n+1} := MM^n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n+1 \\ 0 & 1 \end{pmatrix}$$

Donc la proposition $\mathcal{P}$ est héréditaire. Ainsi $\mathcal{P}$ est vraie pour tout $n \in \mathbb{N}^*$.

Exemple 3. Pour tout $n \in \mathbb{N}^*$, calculer la puissance n -ème de

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

▷ On calcule

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

(Il s'agit en fait d'une rotation d'angle $\pi/2$). Donc pour tout $k \in \mathbb{N}^*$

$$M^{2k} = (M^2)^k = (-1)^k I \quad \text{et} \quad M^{2k+1} = MM^{2k} = (-1)^k M.$$

Finalement

$$M^n = \begin{cases} (-1)^{n/2} I & \text{pour tout } n \text{ pair} \\ M & \text{pour tout } n \text{ impair} \end{cases}$$

Exemple 4. Pour tout $n \in \mathbb{N}^*$, calculer la puissance n -ème de

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

▷ On calcule (par bloc par ex)

$$M^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad M^3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = M$$

(En fait M est une matrice de permutation) donc pour tout

$$M^n = \begin{cases} M & \text{pour tout } n \text{ pair} \\ M^2 & \text{pour tout } n \text{ impair} \end{cases}$$

Le calcul des puissances de certaines matrices triangulaires est aussi aisé en utilisant la formule du binôme de Newton vue en première année :

Proposition 1.1 (► Formule du binôme de Newton). *Soient A et B deux matrices de taille n . Si A et B commutent c-à-d $AB = BA$ alors*

$$(A + B)^p = \sum_{k=0}^p \binom{p}{k} A^k B^{p-k} = \sum_{k=0}^p \binom{p}{k} A^{p-k} B^k .$$

Exemple 5. *Pour tout $n \in \mathbb{N}^*$, calculer la puissance n -ème de*

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

► *On voit que M peut être écrite sous la forme $M = I + N$ avec*

$$N := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

N est nilpotente d'ordre 2 car $N^2 = 0$. On a $NI = IN$ donc on peut appliquer la formule du binôme de Newton, on obtient

$$M^n = (I + N)^n = I + nN + \underbrace{\frac{n(n-1)}{2}N^2 + \frac{n(n-1)(n-2)}{3!}N^3 + \dots}_{=0} = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

Exemple 6.

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

► *On voit que M peut être écrite sous la forme $M = 2I + N$ où*

$$N := \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

On calcule (par bloc par ex)

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad N^3 = 0 .$$

Donc N est une matrice nilpotente d'ordre 3. Comme $2IN = N2I$ on peut appliquer la formule du binôme de Newton et obtenir

$$M^n = (2I + N)^n = 2^n I + 2^{n-1} n N + 2^{n-2} \frac{n(n-1)}{2} N^2 = \begin{pmatrix} 2^n & n2^{n-1} & -n2^{n-1} + 2^{n-3}n(n-1) \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix}$$

On peut vérifier que la formule est correcte pour $n = 1$:

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

Le calcul de la puissance d'une matrice triangulaire n'est pas toujours facile car l'utilisation du binôme de Newton nécessite que les matrices diagonales et nilpotentes commutent ce qui n'est pas toujours le cas. Voir chapitre 2. Une autre classe de matrice pour lesquelles il est facile de calculer la puissance est celle des matrices diagonales. En effet, il a été vu en première année de licence que

Proposition 1.2 (► Puissance de matrice diagonale). Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonale $\text{diag}(\lambda_1, \dots, \lambda_p)$. On a

$$M^k = \text{diag}(\lambda_1^k, \dots, \lambda_p^k).$$

On peut combiner ce résultat au binôme de Newton pour calculer la puissance de matrices qui sont la somme d'une matrice diagonale et d'une nilpotente :

Exemple 7. Calculer la puissance n -ème de

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

► On voit M peut être écrite sous la forme $D + N$ où

$$N := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad D := \text{diag}(-1, 2, 2).$$

De plus N est une matrice nilpotente d'ordre 2.

Comme $DN = ND$ (on verra dans le chapitre suivant une technique, dite de Jordan, qui permet de transformer par changement de variable n'importe quelle matrice en une matrice triangulaire supérieure où on est assuré que la matrice diagonale et la matrice nilpotente commutent) on peut appliquer la formule du binôme de Newton et obtenir

$$\begin{aligned} M^n &= (D + N)^n = D^n + nD^{n-1}N \\ &= \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 2^n \end{pmatrix} + n \begin{pmatrix} (-1)^{n-1} & 0 & 0 \\ 0 & 2^{n-1} & 0 \\ 0 & 0 & 2^{n-1} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 2^n \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & n2^{n-1} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix} \end{aligned}$$

On vérifie pour $n = 1$ que l'on a bien $M^1 = M$.

À noter que par extension toutes les matrices qui peuvent s'écrire sous la forme PDP^{-1} où D est une matrice diagonale et P une matrice inversible (on dit que la matrice est diagonalisable comme nous le verrons bientôt) sont aussi faciles à évaluer car on peut prouver par récurrence, voir Proposition 1.27 que l'on a pour tout $k \in \mathbb{N}^*$,

$$M^k = PD^kP^{-1}.$$

Remarque (►). On pourra pousser le raisonnement pour calculer une exponentielle de matrice. Pour les étudiant.e.s intéressé.e.s notons qu'il est possible de prouver que la série

$$\sum_{k=0}^N \frac{M^k}{k!}$$

converge normalement (prouver que $\|AB\| \leq \|A\|\|B\|$). On peut appeler la limite de cette série "exponentielle de M ", notée $\exp(M)$ (qui est l'analogue de la série exponentielle pour les nombres réels). Cette exponentielle peut être utilisée, par exemple, pour résoudre les équations différentielles ordinaires linéaires d'ordre 1 en dimensions supérieures de la forme $Y' = MY$ où $M \in \mathcal{M}_n$ et $Y \in \mathcal{M}_{n,1}$. Cette équation différentielle a pour base de solution $\{\exp(M)\}$. Ces équations différentielles ordinaires linéaires d'ordre 1 en dimensions quelconque sont d'autant plus intéressantes que toute équation différentielle ordinaire linéaire peut se ramener à un système de ce type. Par exemple

l'équation différentielle linéaire $y'' + ay' + by = c$ où a , b et c sont des réels peut s'écrire sous la forme $Y' = MY$ avec

$$M := \begin{pmatrix} -a & -b \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad Y := \begin{pmatrix} y' \\ y \end{pmatrix}$$

Notez d'ailleurs que c'est une façon élégante de voir que l'espace des solutions est un espace vectoriel de dimension 2.

Ce chapitre va consister à montrer qu'une grande classe de matrice peut s'écrire sous la forme PDP^{-1} où D est une matrice diagonale et P une matrice inversible où D et P seront déterminées explicitement. Dans le même temps nous caractériserons des quantités qui donnent beaucoup d'information sur les propriétés de la matrices : le spectre, les valeurs propres, les vecteurs propres et les espaces propres.

Dans le cas où la matrice n'est pas diagonalisable, il est toujours possible de calculer la puissance de la matrice, au moins dans le corps des complexes en utilisant des techniques dites de *trigonalisation de Jordan* et le binôme de Newton, voir Chapitre 2.

1.1 Rappel sur les applications linéaires

Dans une section préliminaire nous rappellerons, et fixerons les notations, sur les applications linéaires étudiées en L1.

Définition 1.3 (► Applications linéaires). *Soient E et F deux espaces vectoriels. Une application $f : E \rightarrow F$ est une application linéaire de E vers F si*

- $\forall (x, y) \in E^2 \quad f(x + y) = f(x) + f(y)$,
- $\forall x \in E, \forall \lambda \in \mathbb{R} \quad f(\lambda x) = \lambda f(x)$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F . On notera $\mathcal{L}(E)$, au lieu de $\mathcal{L}(E, E)$, l'ensemble des endomorphismes de E .

1.1.1 Représentation matricielle

En dimension finie il est remarquable que pour caractériser une application linéaire partout il suffit de connaître l'évaluation de cette application en un nombre fini de vecteurs : les vecteurs de base. En effet, soit x un vecteur de E de coordonnées $(x_i)_{i \in \{1, \dots, N\}}$ dans la base $\{e_i\}_{i \in \{1, \dots, N\}}$. On a

$$f(x) = f\left(\sum_{i=1}^N x_i e_i\right) = \sum_{i=1}^N x_i f(e_i).$$

On peut donc caractériser cet endomorphisme par une matrice c'est l'objet de la définition suivante :

Définition 1.4 (► Représentation matricielle d'une application linéaire). *Soit E et F deux espaces vectoriels de dimension finie et $f \in \mathcal{L}(E, F)$. On considère $\mathcal{E} := \{e_1, \dots, e_N\}$ une base de E et $\mathcal{F}$ une base de F . La représentation matricielle de f dans les bases $\mathcal{E}, \mathcal{F}$ est la matrice dont la colonne i est constituée des coordonnées de $f(e_i)$ écrite dans la base $\mathcal{F}$. On note cette matrice $\text{mat}(f, \mathcal{E}, \mathcal{F})$ ou $\mathcal{M}_{\mathcal{F}}^{\mathcal{E}}(f)$. Dans le cas d'un endomorphisme on note simplement lorsqu'il n'y a pas d'ambiguïté $\text{mat}(f, \mathcal{E})$ ou $\mathcal{M}_{\mathcal{E}}(f)$.*

Réciproquement, soit $A \in \mathcal{M}_n(\mathbb{R})$, l'application

$$\begin{aligned} \mathcal{M}_{n,1} &\rightarrow \mathcal{M}_{n,1} \\ X &\mapsto A \cdot X \end{aligned}$$

est une application linéaire.

Remarque (►). *On fera souvent l'identification entre un vecteur de $\mathbb{R}^N$ et la matrice colonne donc entre $\mathcal{M}_{n,1}$ et $\mathbb{R}^n$.*

Remarque (►). On remarquera que dans la réciproque les bases sous-jacentes sont les bases canoniques de E et de F donc l'identification entre matrice et application linéaire est faite les bases de E et de F étant prescrites. On parle d'isomorphisme non-canonique.

Exemple 8. Déterminer la forme de l'application linéaire associée à la matrice

$$M := \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{pmatrix}$$

▷ On calcule

$$\begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x - z \end{pmatrix}$$

Donc l'application linéaire associée à M est

$$f : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^2 \\ (x, y, z) & \mapsto & (x + 2y, 3x - z) \end{array}$$

Exemple 9. Considérons

$$f : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^2 \\ (x, y, z) & \mapsto & (x + y - z, x - z) \end{array}$$

1. Quelle est la représentation matricielle de f dans les bases canoniques de $\mathcal{C}_3$ de $\mathbb{R}^3$ et $\mathcal{C}_2$ de $\mathbb{R}^2$?
2. Quelle est la représentation matricielle de f dans les bases $\mathcal{E} := \{(1, 1, 0), (0, 1, 0), (1, 1, 1)\}$, $\mathcal{F} := \{(1, 1), (1, 0)\}$.

▷

1. On calcule

$$f(1, 0, 0) = (1, 1) \quad f(0, 1, 0) = (1, 0) \quad f(0, 0, 1) = (-1, -1).$$

Notez que lorsque l'on ne précise pas la base dans laquelle est écrit le vecteur il est par défaut écrit dans les bases canoniques. On obtient donc

$$\text{mat}(f, \mathcal{C}_3, \mathcal{C}_2) = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

2. On calcule

$$\begin{aligned} f(1, 1, 0) &= (2, 1) = 1 * (1, 1) + 1 * (1, 0) = (1, 1)_{\mathcal{F}} \\ f(0, 1, 0) &= (1, 0) = 0 * (1, 1) + 1 * (1, 0) = (0, 1)_{\mathcal{F}} \\ f(1, 1, 1) &= (1, 0) = (0, 1)_{\mathcal{F}} \end{aligned}$$

Donc

$$\text{mat}(f, \mathcal{E}, \mathcal{F}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Dans la deuxième question, la détermination des coordonnées de l'image dans $\mathcal{F}$ est équivalente en terme de calcul à la résolution d'un système linéaire ou de l'inversion d'une matrice. On peut le faire à la main et c'est en général le plus efficace si on ne cherche à déterminer la représentation matricielle de f que dans une base différente de la base canonique. Si on cherche à trouver la représentation matricielle de plusieurs applications linéaires dans ces nouvelles bases, ou à plus forte raison si on cherche à faire des démonstrations abstraites il est intéressant de formaliser ces changements de base. C'est ce que nous allons faire dans la section qui suit.

1.1.2 Changement de base

Définition 1.5 (► Matrice de passage). Soient E un espace vectoriel de dimension finie et $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E . On appelle matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$ la matrice de l'application identité id dans les bases $\mathcal{E}'$ et $\mathcal{E}$. On la note $P_{\mathcal{E}}^{\mathcal{E}'}$:

$$P_{\mathcal{E}}^{\mathcal{E}'} := \text{mat}(\text{id}, \mathcal{E}', \mathcal{E}).$$

Remarque (►). Notez que la base de l'espace de départ est $\mathcal{E}'$ et que celle d'arrivée est $\mathcal{E}$ donc la dénomination matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$ est un peu trompeuse.

Exemple 10. Écrire la matrice de passage de la base canonique de $\mathcal{C}_3$ de $\mathbb{R}^3$ à la base $\mathcal{E} := \{(1, 1, 0), (0, 1, 0), (1, 1, 1)\}$.

► On remarquera pour commencer que $\mathcal{E}$ forme bien une base de $\mathbb{R}^3$ car le déterminant de sa matrice est égal à 1 (exercice). On a donc

$$P_{\mathcal{C}_3}^{\mathcal{E}} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Exemple 11. Écrire la matrice de passage de la base canonique de $\mathbb{R}_2[X]$ à la base $\mathcal{E} := \{1, X - 1, 2X^2 - X + 1\}$.

► On a

$$P_{\mathcal{C}_3}^{\mathcal{E}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

Exemple 12. Écrire la matrice de passage de la base $\mathcal{E} := \{(1, 1, 0), (0, 1, 0), (1, 1, 1)\}$ à la base canonique $\mathcal{C}_3$ de $\mathbb{R}^3$.

► C'est la question réciproque de l'exercice 10. Il faut déterminer les coordonnées des vecteurs de la base canonique dans la base $\mathcal{E}$:

$$\begin{aligned} (1, 0, 0) &= 1 * (1, 1, 0) + (-1) * (0, 1, 0) + 0 * (1, 1, 1) = (1, -1, 0)_{\mathcal{E}} \\ (0, 1, 0) &= (0, 1, 0)_{\mathcal{E}} \\ (0, 0, 1) &= (-1, 0, 1)_{\mathcal{E}} \end{aligned}$$

Donc

$$P_{\mathcal{E}}^{\mathcal{C}_3} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Dans ce dernier exemple on a résolu le système linéaire et donc en fait on a calculé l'inverse de la matrice. C'est en fait un résultat générique :

Proposition 1.6 (► Inverse d'une matrice de passage). Soit E un espace vectoriel de dimension finie et $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E . On a $P_{\mathcal{E}'}^{\mathcal{E}} = \left[P_{\mathcal{E}}^{\mathcal{E}'} \right]^{-1}$.

► *Démonstration.* On a

$$P_{\mathcal{E}}^{\mathcal{E}'} \cdot P_{\mathcal{E}'}^{\mathcal{E}} = \text{mat}(\text{id}, \mathcal{E}', \mathcal{E}) \cdot \text{mat}(\text{id}, \mathcal{E}, \mathcal{E}') = \text{mat}(\text{id}, \mathcal{E}', \mathcal{E}') = I_n.$$

□

Changement de base : vecteur

Proposition 1.7 (► Changement de base : vecteur). Soient E un espace vectoriel de dimension finie et $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E . On considère x un vecteur de E . Soit $X_{\mathcal{E}}$ les coordonnées du vecteur x dans la base $\mathcal{E}$ et $X_{\mathcal{E}'}$ les coordonnées du vecteur x dans la base $\mathcal{E}'$. On a

$$X_{\mathcal{E}} = P_{\mathcal{E}}^{\mathcal{E}'} \cdot X_{\mathcal{E}'}.$$

► *Démonstration.* On a

$$P_{\mathcal{E}}^{\mathcal{E}'} \cdot X_{\mathcal{E}'} = \text{mat}(\text{id}, \mathcal{E}', \mathcal{E}) \cdot X_{\mathcal{E}'} = X_{\mathcal{E}}.$$

□

Changement de base : application linéaire

Proposition 1.8 (► Changement de base : application linéaire). *Soient E et F deux espaces vectoriels de dimension finie. Soient $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E et $\mathcal{F}$ et $\mathcal{F}'$ deux bases de F . On considère $f \in \mathcal{L}(E, F)$. On a*

$$\text{mat}(f, \mathcal{E}, \mathcal{F}) = P_{\mathcal{F}'}^{\mathcal{F}} \text{mat}(f, \mathcal{E}', \mathcal{F}') P_{\mathcal{E}'}^{\mathcal{E}}.$$

Remarque (►). *Pour se souvenir de la formule on peut appliquer $X_{\mathcal{E}}$ des deux côtés de l'égalité. À droite il faut faire agir $P_{\mathcal{E}'}^{\mathcal{E}}$ sur $X_{\mathcal{E}}$ pour obtenir un vecteur "écrit" dans les coordonnées de $\mathcal{E}'$ qui sont celles que "comprend" $\text{mat}(f, \mathcal{E}', \mathcal{F}')$. On obtient alors à droite un vecteur $Y_{\mathcal{F}'}$ alors qu'à gauche on avait ce même vecteur écrit dans la base $\mathcal{F}$. Il faut donc pré-multiplier le membre de droite par $P_{\mathcal{F}'}^{\mathcal{F}}$.*

$$\text{mat}(f, \mathcal{E}, \mathcal{F}) X_{\mathcal{E}} = P_{\mathcal{F}'}^{\mathcal{F}} \text{mat}(f, \mathcal{E}', \mathcal{F}') P_{\mathcal{E}'}^{\mathcal{E}} X_{\mathcal{E}}.$$

► Dans le cas d'un endomorphisme de E cette formule peut s'écrire

$$\text{mat}(f, \mathcal{E}) = P \text{mat}(f, \mathcal{E}') P^{-1} \quad \text{où} \quad P := P_{\mathcal{E}'}^{\mathcal{E}}. \quad (1.1)$$

Nous réutiliserons cette formule si on interprète la diagonalisation comme la réécriture d'une application linéaire dans une base appropriée.

1.2 Diagonalisation et éléments propres

Définition 1.9 (► Matrices semblables). *Soit $(A, B) \in \mathcal{M}_N(\mathbb{K})^2$. On dit que A et B sont semblables dans $\mathbb{K}$ s'il existe une matrice inversible $P \in \mathcal{M}_N(\mathbb{K})$ telle que*

$$A = P B P^{-1}.$$

Remarque (►). *L'ensemble des matrices inversible sur $\mathbb{K}$ n'est pas un sous-espace vectoriel de $\mathcal{M}_N(\mathbb{K})$ car la somme de deux matrices inversibles n'est pas nécessairement inversible. Par contre, c'est un groupe pour la multiplication matricielle. En effet, le produit de deux matrices inversibles est inversible, ce produit est également associatif, la matrice identité, qui est bien inversible, est l'élément neutre et toute matrice inversible admet un symétrique pour cette opération (c'est son inverse).*

Définition 1.10 (► Matrice diagonalisable). *Soit $A \in \mathcal{M}_N$. A est diagonalisable dans $\mathbb{K}$ si elle est semblable dans $\mathbb{K}$ à une matrice diagonale.*

Même si nous avons choisi d'avoir une présentation de la diagonalisation orientée vers les applications linéaires, cette définition est purement matricielle et s'applique en toute généralité. Par extension nous dirons qu'une application linéaire est diagonalisable si sa représentation matricielle est diagonalisable.

► Soit E un espace vectoriel de dimension N . On considère un endomorphisme f de E et $A \in \mathcal{M}_N$ sa représentation matricielle dans la base canonique de E . On va chercher à construire une base $\mathcal{E} := \{\varepsilon_i\}_{i \in \{1, \dots, N\}}$ dans laquelle la représentation matricielle de f est diagonale : $\text{mat}(f, \mathcal{E}) = \text{diag}(\lambda_i)_{i \in \{1, \dots, N\}}$. D'après la définition de la représentation matricielle, Définition 1.4, cela veut dire que

$$\text{mat}(f, \mathcal{E}) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}$$

donc que

$$\forall i \in \{1, \dots, N\}, \quad f(\varepsilon_i) = \lambda_i \varepsilon_i.$$

Définition 1.11 (► Valeurs propres, vecteurs propres). Soit $A \in \mathcal{M}_N(\mathbb{K})$. On appelle λ valeur propre de A dans le corps $\mathbb{K}$ s'il existe $X_\lambda \neq 0$ dans $\mathbb{K}^N$ tel que $AX_\lambda = \lambda X_\lambda$. Un tel X_λ est appelé vecteur propre associé à la valeur propre λ (pour la matrice A dans le corps $\mathbb{K}$). L'ensemble des valeurs propres de A est appelé spectre de A et noté $\text{Sp}(A)$.

Exemple 13. Soit

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

1. Montrer que $(1, 1)$ est un vecteur propre de A .

2. Montrer que $(-1, 1)$ est un vecteur propre de A .

▷ On calcule

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

donc $(1, 1)$ est un vecteur propre associé à la valeur propre 3. De même

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

donc $(-1, 1)$ est un vecteur propre associé à la valeur propre -1 .

► Afin de déterminer une telle base il faut pouvoir trouver des couples $(\lambda_i, \varepsilon_i) \in \mathbb{R} \times E \setminus \{0_E\}$ tels que

$$f(\varepsilon_i) = \lambda_i \varepsilon_i .$$

Autrement dit

$$\text{On cherche un } \varepsilon_i \neq 0_E \text{ tel que } f(\varepsilon_i) - \lambda_i \varepsilon_i = 0_E$$

ou de façon équivalente

$$\text{On cherche un } \varepsilon_i \neq 0_E \text{ tel que } (f - \lambda_i \text{id})\varepsilon_i = 0_E$$

ou encore

$$\text{On cherche un } \varepsilon_i \neq 0_E \text{ dans } \ker(f - \lambda_i \text{id})$$

D'après le cours d'algèbre linéaire de première année, on sait que cela est possible si $\dim(\ker(f - \lambda_i \text{id})) > 0$. Ce qui est équivalent à $\det(f - \lambda_i \text{id}) = 0$. On est donc réduits à trouver des zéros de $\det(f - \lambda_i \text{id})$. En terme matricielle on cherche à trouver des zéros de $\det(A - \lambda_i I)$. Notez que l'on a réduit le problème à deux inconnues $(\lambda_i, \varepsilon_i) \in \mathbb{R} \times E \setminus \{0_E\}$ à un problème à une inconnue λ_i .

Définition 1.12 (► Polynôme caractéristique). Soit $A \in \mathcal{M}_N$ on appelle polynôme caractéristique de A le polynôme

$$\chi_A : \begin{array}{l} E \rightarrow \mathbb{R} \\ X \mapsto \det(A - X I) . \end{array}$$

D'après le raisonnement que l'on a développé ci-dessus il est évident que

Proposition 1.13 (► Caractérisation des valeurs propres). Soit $A \in \mathcal{M}_N$. Un scalaire $\lambda \in \mathbb{K}$ est valeur propre si et seulement s'il est racine dans $\mathbb{K}$ du polynôme caractéristique χ_A .

Exemple 14. Déterminer le spectre de

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

Quand on a trouvé une valeur propre λ , tout élément non-nul de $\ker(A - \lambda I_N)$ est un vecteur propre :

Définition 1.14 (► Espace propre). Soient $A \in \mathcal{M}_N$ et $\lambda \in \mathbb{K}$. On appelle espace propre le sous-espace vectoriel $E_\lambda := \ker(A - \lambda I)$.

Proposition 1.15 (► Vecteur propre de l'espace propre). Soient $A \in \mathcal{M}_N$ et $\lambda \in \text{Sp}(A)$. Tout vecteur de E_λ est un vecteur propre associé à la valeur propre λ .

Remarque (►). Attention si λ est valeur propre on sait qu'il existe un vecteur propre non-nul dans l'espace propre. Il n'est donc pas possible d'avoir un espace propre associé à une valeur propre réduit au vecteur nul!

Exemple 15. Déterminer des vecteurs propres associés aux valeurs propres -1 et 3 de

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

Proposition 1.16 (► Matrices semblables et valeurs propres). Si deux matrices sont semblables alors elles ont le même polynôme caractéristique. En particulier elles ont les mêmes valeurs propres.

► *Démonstration.* Soient A et B deux matrices semblables dans $\mathbb{K}$. Alors il existe une matrice inversible P telle que $A = PBP^{-1}$. Or

$$\begin{aligned} \chi_A &= \det(A - XI) = \det(PBP^{-1} - XPIP^{-1}) = \det(P(B - XI)P^{-1}) \\ &= \det(P) \det(B - XI) \det(P^{-1}) = \det(B - XI) = \chi_B. \end{aligned}$$

□

Définition 1.17 (► Multiplicité algébrique). Soit $\lambda \in \mathbb{K}$ une valeur propre de la matrice $A \in \mathcal{M}_N(\mathbb{K})$. On appelle multiplicité algébrique de la valeur propre λ de A l'ordre de multiplicité de λ comme racine du polynôme caractéristique χ_A .

Remarque (►). Proposition 1.16 a une première conséquence pratique qui permet de déceler parfois des erreurs dans la détermination du spectre d'une matrice. Supposons que la matrice soit diagonalisable, alors un rapide calcul de combinatoire permet de voir que le polynôme caractéristique est de la forme

$$(-1)^n X^n + (-1)^{n-1} \text{Tr}(A) X^{n-1} + \dots + \det(A).$$

Comme le polynôme caractéristique est le même pour toute matrice semblable on doit nécessairement avoir :

- La trace de A doit être égal à la somme des valeurs propres multipliée par leur multiplicité algébrique,
- Le déterminant de A doit être égal au produit des valeurs propres multipliée par leur multiplicité algébrique.

Si le deuxième critère est un peu plus pénible à vérifier le premier est très pratique et permet d'éliminer aisément certaines erreurs. Cette remarque est valable aussi lorsque la matrice n'est pas diagonalisable, comme nous le verrons dans le chapitre sur la trigonalisation.

Exemple 16. Soit

$$A = \begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix}$$

1. Déterminer χ_A .
2. Déterminer le spectre de A dans $\mathbb{R}$.
3. Déterminer une base de chaque espace propre associé aux valeurs propres réelles de A .
4. Déterminer une base de vecteurs propres de A .

►

1. On calcule

$$\begin{aligned}\chi_A(\lambda) &= \det \left(\begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \det \left(\begin{pmatrix} 1-\lambda & -4 \\ -4 & 7-\lambda \end{pmatrix} \right) \\ &= \lambda^2 - 8\lambda - 9 = (\lambda - 9)(\lambda + 1).\end{aligned}$$

Attention à ne pas confondre le polynôme caractéristique $\det(A - XI)$ qui est un polynôme et l'évaluation de ce polynôme en λ qui est donné par $\det(A - \lambda I)$.

2. Les deux racines de χ_A sont 9 et -1 donc $\text{Sp}(A) = \{-1, 9\}$.

3. On calcule

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-1} = \ker(A + I) \Leftrightarrow \begin{pmatrix} 2 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow x - 2y = 0$$

donc $E_{-1} = \text{Vect}\{(2, 1)\}$. On peut vérifier que

$$\begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix} = (-1) * \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

donc $(2, 1)$ est bien un vecteur propre associé à la valeur propre -1 . On calcule

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_9 \Leftrightarrow \begin{pmatrix} -8 & -4 \\ -4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow 2x + y = 0$$

donc $E_9 = \text{Vect}\{(1, -2)\}$. On peut encore vérifier en effectuant le calcul comme ci-dessus. Il est recommandé de faire cette vérification lors des premiers calculs de vecteurs propres. Avec l'expérience et la rigueur on aura moins tendance à le faire par la suite.

4. La famille $\{(2, 1), (1, -2)\}$ forme une base de $\mathbb{R}^2$ constituée de vecteurs propres.

Exemple 17. Soit

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

1. Déterminer le spectre de A .

2. Déterminer une base de vecteurs propres de A .

▷

1. On calcule $\chi_A(\lambda) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 2)(\lambda - 1)^2$. Donc $\text{Sp}(A) = \{1, 2\}$. On vérifie que $2 + 1 + 1 = 4 = \text{Tr}(A)$.

2. On calcule $E_2 = \text{Vect}\{(1, 2, 1)\}$ et $E_1 = \text{Vect}\{(-1, 1, 0), (2, 0, 1)\}$. Or on peut vérifier (par le calcul du déterminant par exemple, mais nous verrons dans la section suivante que les espaces propres sont en somme directe) que $\{(1, 2, 1), (-1, 1, 0), (2, 0, 1)\}$ forme une base de $\mathbb{R}^3$.

Exemple 18. Soit

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

1. Déterminer le spectre de A .

2. Déterminer les espaces propres de A .

▷

1. On calcule $\chi_A(\lambda) = -\lambda^3 + 6\lambda^2 - 32 = -(\lambda + 2)(\lambda - 4)^2$. Donc $\text{Sp}(A) = \{-2, 4\}$ (on ne répète pas deux fois le 4 même s'il est racine double du polynôme caractéristique). On vérifie que $-2 + 2 \times 4 = 6 = \text{Tr}(A)$.

2. On calcule $E_4 = \text{Vect}\{(0, 1, 0)\}$ et $E_{-2} = \{(-3, -2, 3)\}$. Notez que l'on a que deux vecteurs propres de A donc ils ne forment pas une base de $\mathbb{R}^3$.

Nous n'avons pas beaucoup pris la peine ci-dessus de préciser le corps dans lequel nous travaillons. Il est pourtant important de bien prendre conscience que le fait de travailler dans $\mathbb{R}$ ou dans $\mathbb{C}$ n'est pas équivalent comme le montre l'exemple ci-dessous :

Exemple 19. Soit

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

1. Déterminer χ_A .
2. Déterminer le spectre de A dans $\mathbb{R}$.
3. Déterminer le spectre de A dans $\mathbb{C}$.

▷

1. On calcule $\chi_A = X^2 + 1$.
2. Le polynôme caractéristique n'a pas de racine dans $\mathbb{R}$ donc le spectre de A est vide dans $\mathbb{R}$.
3. Le polynôme caractéristique dans le corps des complexes a deux racines i et $-i$. donc $\text{Sp}(A) = \{i, -i\}$ dans $\mathbb{C}$.

On peut d'ailleurs remarquer que les deux valeurs propres de A dans la cas complexe son conjuguées. C'est toujours le cas comme le montre la proposition suivante

Proposition 1.18 (► Valeurs propres imaginaires pures). Soit $A \in \mathcal{M}_N(\mathbb{R})$. Si $\lambda \in \mathbb{C}$ est valeur propre associée au vecteur propre X_λ alors le conjugué de λ est aussi valeur propre et $\bar{X}_\lambda$ est vecteur propre associé à $\bar{\lambda}$.

► *Démonstration.* Comme $\lambda \in \mathbb{C}$ est valeur propre associée au vecteur propre X_λ on a $AX_\lambda = \lambda X_\lambda$. En prenant le conjugué on a

$$\overline{AX_\lambda} = \overline{\lambda X_\lambda} \Leftrightarrow A\bar{X}_\lambda = \bar{\lambda}\bar{X}_\lambda$$

ce qui signifie que $\bar{X}_\lambda$ est vecteur propre associé à $\bar{\lambda}$. □

Remarque (►). Il me semble important ici de relativiser l'efficacité pratique de la diagonalisation. Si la diagonalisation est une technique théoriquement fructueuse elle n'est pas, en pratique, si facile à mettre en œuvre qu'il pourrait sembler. Nous allons, en effet, utiliser la technique de diagonalisation pour diagonaliser de nombreuses matrices mais nous tenons à insister sur le fait que toutes les matrices que nous allons diagonaliser ont été sélectionnées avec soin pour que les calculs soient faciles. En particulier on voit par la proposition ci-dessus que la première étape consiste à déterminer les racines d'un polynôme. Or s'il est aisé de déterminer les racines d'un polynôme de degré 2, pour les polynômes de degré 3, il n'est pas toujours aisé de déterminer ses racines. En fait il existe une méthode générique pour déterminer les racines d'un polynôme de degré 3 et 4 : la méthode de Cardan (Niccolò Fontana, 1545) mais elle est hors-programme. En dimension supérieures le théorème d'Abel-Ruffini prouve que, pour tout entier n supérieur ou égal à 5, il n'existe pas de formule générale exprimant les solutions « par radicaux ». Donc en général, on ne saura pas déterminer le spectre d'une matrice de taille plus grande que 4 et on ne pourra pas développer notre technique de diagonalisation. Notons encore que dans les matrices de taille 3 que nous allons essayer de diagonaliser (et de trigonaliser) cette année il y aura toujours une racine évidente parmi $-2, -1, 0, 1, 2, 5, 9$.

1.3 Théorème de diagonalisation : critère de dimension

Proposition 1.19 (► Endomorphisme diagonalisable). Soit $f \in \mathcal{L}(E)$. L'endomorphisme f est diagonalisable dans $\mathbb{K}$ si et seulement s'il existe une base de vecteurs propres de f dans $\mathbb{K}$.

► *Démonstration.* Comme nous l'avons vu en début de section, s'il existe une base de vecteurs propres pour f , la représentation matricielle A de f dans cette base est une matrice diagonale D . La formule de changement de variable pour les applications linéaires (1.1) donne directement $A = PDP^{-1}$. $\square$

Remarque. Comme nous l'avons vu plus haut il est possible d'identifier une matrice et une application linéaire. Dans les preuves qui suivent, nous passerons allégrement de la matrice A à l'application linéaire sous-jacente f sans le rappeler à chaque fois.

Exemple 20. Diagonaliser

$$A = \begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix}$$

▷ En utilisant les résultats de l'exercice 16 on voit que dans la base $\mathcal{E} := \{(2, 1), (1, -2)\}$ la représentation matricielle de l'application linéaire sous-jacente associée à A est

$$D := \begin{pmatrix} -1 & 0 \\ 0 & 9 \end{pmatrix}$$

En utilisant la formule de changement de base dans une application linéaire (1.1) on obtient bien $A = PDP^{-1}$ avec $D = \text{diag}\{-1, 9\}$ et P est la matrice de passage de la base canonique à la base de vecteurs propres $\mathcal{E}$ données par les coordonnées des vecteurs de $\mathcal{E}$ donnés dans la base canonique c-à-d

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

On en calculera pas P^{-1} si ça n'est pas explicitement demandé.

Exemple 21. Diagonaliser

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

▷ En utilisant l'exercice 17 on voit que dans la base $\mathcal{E} := \{(1, 2, 1), (-1, 1, 0), (2, 0, 1)\}$ la représentation matricielle de l'application linéaire (sous-jacente associée à A) est

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

Remarquez que 1 apparaît deux fois dans la matrice D . En utilisant la formule de changement de base dans une application linéaire (1.1) on obtient bien $A = PDP^{-1}$ avec $D = \text{diag}\{2, 1, 1\}$ et P est la matrice de passage de la base canonique à la base de vecteurs propres $\mathcal{E}$ données par les coordonnées des vecteurs de $\mathcal{E}$ donnés dans la base canonique c-à-d

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

Exemple 22. La matrice suivante est-elle diagonalisable dans $\mathbb{R}$,

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

▷ Nous n'avons pas trouvé de base de vecteurs propres de $\mathbb{R}^3$ et la matrice n'est en fait pas diagonalisable comme nous allons le voir grâce au critère suivant.

Théorème 1.20 (► Critère de dimension). Soient $A \in \mathcal{M}_N$ et $\text{Sp}(A) = \{\lambda_i\}_{i \in \{1, \dots, p\}}$. A est diagonalisable si et seulement si

$$\sum_{i=1}^p \dim(E_i) = N .$$

La démonstration de ce théorème repose sur le lemme suivant :

Lemma 1.21 (► Somme directe de sous-espaces propres). Soient $A \in \mathcal{M}_N$ et $\{\lambda_i\}_{i \in \{1, \dots, p\}}$ des scalaires deux à deux distincts. Les sous-espaces propres $E_{\lambda_1}, \dots, E_{\lambda_p}$ sont en somme directe.

► *Démonstration.* Soit pour $k \in \mathbb{N}^*$ la proposition : $\mathcal{P}(k) := "E_{\lambda_1}, \dots, E_{\lambda_k}$ sont en somme directe"

- $\mathcal{P}(1)$ est vraie.
- Supposons que $\mathcal{P}(k)$ soit vraie à un rang k donné. C'est à dire que " $E_{\lambda_1}, \dots, E_{\lambda_k}$ sont en somme directe". On veut montrer que $E_{\lambda_1}, \dots, E_{\lambda_{k+1}}$ sont en somme directe ce qui revient à montrer que $E_{\lambda_1}, \dots, E_{\lambda_k} \cap E_{\lambda_{k+1}} = 0_E$. Prenons $x \in \left(\sum_{i=1}^k E_{\lambda_i}\right) \cap E_{\lambda_{k+1}}$. D'une part comme $x \in E_{\lambda_{k+1}}$ on a

$$Ax = \lambda_{k+1}x . \quad (1.2)$$

D'autre part comme $x \in \sum_{i=1}^k E_{\lambda_i}$, il existe $\{x_i\}_{i \in \{1, \dots, k\}} \in \times_{i=1}^k E_{\lambda_i}$ tel que $x = \sum_{i=1}^k x_i$. En conséquence

$$Ax = A \left(\sum_{i=1}^k x_i \right) = \sum_{i=1}^k f(x_i) = \sum_{i=1}^k \lambda_i x_i . \quad (1.3)$$

En soustrayant (1.2) et (1.3) on obtient

$$0 = \sum_{i=1}^k (\lambda_{k+1} - \lambda_i) x_i .$$

Comme $\{x_i\}_{i \in \{1, \dots, k\}} \in \times_{i=1}^k E_{\lambda_i}$ où $E_{\lambda_1}, \dots, E_{\lambda_k}$ sont en somme directe donc

$$\forall i \in \{1, \dots, k\}, \quad (\lambda_{k+1} - \lambda_i) x_i = 0 ; .$$

Or comme pour tout $i \in \{1, \dots, k\}$, $\lambda_{k+1} \neq \lambda_i$ ceci implique que pour tout $i \in \{1, \dots, k\}$ $x_i = 0$ et donc que $x = 0$. On a ainsi prouvé que $E_{\lambda_1}, \dots, E_{\lambda_k} \cap E_{\lambda_{k+1}} = 0_E$ et donc que $E_{\lambda_1}, \dots, E_{\lambda_{k+1}}$ sont en somme directe. Ce qui prouve que la proposition $\mathcal{P}$ est héréditaire. Ainsi $\mathcal{P}$ est vraie pour tout $k \in \mathbb{N}^*$. $\square$

- *Démonstration du Théorème 1.20.*
- ($\Rightarrow$) : Si A est diagonalisable il existe une base dans laquelle la représentation matricielle est $\text{diag}(\lambda_1, \dots, \lambda_N)$. Les λ_i ne sont pas nécessairement distincts. Quitte à changer l'ordre des vecteurs on peut supposer que les λ_i sont ordonnés par ordre croissant. Pour tout $i \in \{1, \dots, N\}$, notons α_i le nombre d'occurrence de λ_i dans la matrice diagonale. Pour tout $i \in \{1, \dots, N\}$, les α_i vecteurs de base correspondants aux colonnes qui contiennent un λ_i sont des vecteurs propres associés à la valeur propre λ_i . Ainsi l'espace propre E_{λ_i} est de dimension α_i . Donc $\sum_{i=1}^N \dim E_{\lambda_i} = \sum_{i=1}^N \alpha_i$. Or comme il y a N éléments λ_i sur la diagonale, cette somme est égale à la dimension N .
 - ($\Leftarrow$) : On a évidemment $\bigoplus_{i=1}^p E_{\lambda_i} \subset E$. D'après Lemme 1.21, $E_{\lambda_1}, \dots, E_{\lambda_{k+1}}$ sont en somme directe on a donc $\dim(\bigoplus_{i=1}^p E_{\lambda_i}) = \sum_{i=1}^p \dim E_{\lambda_i}$. D'après l'hypothèse cette somme est égale à N . On a donc $\bigoplus_{i=1}^p E_{\lambda_i} \subset E$ et $\dim(\bigoplus_{i=1}^p E_{\lambda_i}) = \dim E$ donc $E = \bigoplus_{i=1}^p E_{\lambda_i}$. Pour tout $i \in \{1, \dots, p\}$, définissons $\mathcal{B}_i$ une base de E_{λ_i} . La concaténation des bases $\{\mathcal{B}_i\}_{i \in \{1, \dots, p\}}$ est une base de vecteurs propres de E . Donc la matrice est diagonalisable. $\square$

On a même démontré :

Corollaire 1.22. Soient $A \in \mathcal{M}_N$ et $\text{Sp}(A) = \{\lambda_i\}_{i \in \{1, \dots, p\}}$. A est diagonalisable si et seulement si

$$E = \bigoplus_{i=1}^p E_{\lambda_i} .$$

1.4 Théorème de diagonalisation : critère de multiplicité

Rappelons que l'on appelle $m(\lambda)$ la multiplicité algébrique de λ .

Théorème 1.23 (► Critère de multiplicité). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. On a*

$$A \text{ est diagonalisable} \Leftrightarrow \begin{cases} \chi_A \text{ est scindé sur } \mathbb{K} \\ \text{Pour tout } \lambda \in \text{Sp}(A) \text{ on a } m(\lambda) = \dim E_\lambda \end{cases} .$$

La démonstration de ce théorème repose sur

Lemma 1.24 (► Dimension des espaces propres). *Soient $A \in \mathcal{M}_N(\mathbb{K})$ et χ_A son polynôme caractéristique. Si λ est une valeur propre de multiplicité $m(\lambda)$ dans le polynôme caractéristique alors*

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

► *Démonstration.* Tout d'abord par définition d'une valeur propre on a $\dim(E_\lambda) \geq 1$.

Notons $p := \dim E_\lambda$ et $\{\varepsilon_1, \dots, \varepsilon_p\}$ une base de E_λ . On complète cette base en une base $\{\varepsilon_1, \dots, \varepsilon_p, \varepsilon_{p+1}, \dots, \varepsilon_n\}$ de E . Dans cette base la représentation matricielle de f est de la forme

$$\begin{pmatrix} \lambda I_p & B \\ 0 & C \end{pmatrix}$$

En calculant le polynôme caractéristique de cette matrice triangulaire supérieure on obtient

$$\det((\lambda - X)I_p) \det(B - XI_{n-p}) = (\lambda - X)^p \det(B - XI_{n-p}) .$$

Ceci prouve que $(\lambda - X)^p$ divise χ_f . Par définition de la multiplicité d'une racine on a $m(\lambda) \geq p$. ◻

► *Démonstration du Théorème 1.23.* • ($\Rightarrow$) : Notons $\{\lambda_i\}_{i \in \{1, \dots, p\}}$ les valeurs propres distinctes de f et $\{m(\lambda_i)\}_{i \in \{1, \dots, p\}}$ leurs multiplicités algébriques respectives. Comme f est diagonalisable il existe une base $\mathcal{E}$ dans laquelle la représentation matricielle de f est diagonale :

$$D := \text{diag}\{\underbrace{\lambda_1, \dots, \lambda_1}_{\alpha_1}, \underbrace{\lambda_2, \dots, \lambda_2}_{\alpha_2}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{\alpha_p}\} .$$

On a évidemment

$$\chi_f = \chi_D = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i} .$$

On en déduit que χ_f est scindé sur $\mathbb{K}$ et que, pour tout $i \in \{1, \dots, p\}$, $m(\lambda_i) = \alpha_i$. De plus comme $\mathcal{E}$ est une base de vecteurs propres, les vecteurs propres correspondants aux valeurs propres λ_i forment une famille libre de α_i vecteurs de E_{λ_i} . Donc $\dim E_{\lambda_i} \geq \alpha_i$. Comme par ailleurs on a prouvé l'inégalité inverse dans le lemme 1.24, on a $\dim E_{\lambda_i} = \alpha_i$.

• ($\Leftarrow$) : Soient $\{\lambda_i\}_{i \in \{1, \dots, p\}}$ les valeurs propres distinctes de f et $\{m(\lambda_i)\}_{i \in \{1, \dots, p\}}$ leurs multiplicités algébrique. On a

$$\chi_f = \prod_{i=1}^p (\lambda_i - X)^{m(\lambda_i)} .$$

D'après Lemme 1.21, les espaces propres sont en somme directe donc

$$\dim(\oplus_{i=1}^p E_{\lambda_i}) = \sum_{i=1}^p \dim(E_{\lambda_i}) = \sum_{i=1}^p m(\lambda_i) = \deg(\chi_f) = \dim E .$$

Comme $\oplus_{i=1}^p E_{\lambda_i} \subset E$ et $\dim \oplus_{i=1}^p E_{\lambda_i} = \dim E$ on a $\oplus_{i=1}^p E_{\lambda_i} = E$. Soit $\mathcal{E}_i$ une base de E_{λ_i} . La concaténation des bases $\{\mathcal{E}_i\}_{i=1, \dots, p}$ forme une base de vecteurs propres de E . Ce qui implique que f est diagonalisable. ◻

Corollaire 1.25 (► Polynôme caractéristique scindé à racines simples). Soit $A \in \mathcal{M}_N(\mathbb{K})$. Si A a N valeurs propres distinctes dans $\mathbb{K}$ alors A est diagonalisable dans $\mathbb{K}$.

► *Démonstration.* D'après Lemme 1.24 on a $1 \leq \dim(E_{\lambda_i}) \leq 1$ donc $\dim(E_{\lambda_i}) = 1$. On peut appliquer Théorème 1.20 pour conclure. $\square$

1.5 Applications

1.5.1 Application aux itérés d'une application linéaire

Définition 1.26 (Itérés d'une application linéaire). Soit $f \in \mathcal{L}(E)$. On peut définir les itérés de f par récurrence pour tout $k \in \mathbb{N}^*$

$$\begin{cases} f^{(1)} = f \\ f^{(k+1)} = f \circ f^{(k)} \end{cases}.$$

C'est l'interprétation que l'on a suivit tout au long de ce chapitre. Or on a

Proposition 1.27 (► Puissance d'une matrice diagonalisable). Soit $M \in \mathcal{M}_n(\mathbb{R})$ telle qu'il existe une matrice diagonale D et une matrice inversible P telles que $M = PDP^{-1}$. On a pour tout $k \in \mathbb{N}^*$,

$$M^k = PD^k P^{-1}.$$

► *Démonstration.* Soit pour $k \in \mathbb{N}^*$ la proposition : $\mathcal{P}(k) := "M^k = PD^k P^{-1}"$.

- $\mathcal{P}(1)$ est vraie.
- Supposons que $\mathcal{P}(k)$ soit vraie à un rang k donné. On calcule alors

$$M^{k+1} := MM^k = PDP^{-1}PD^k P^{-1} = PD^{k+1}P^{-1}$$

Donc la proposition $\mathcal{P}$ est héréditaire. Ainsi $\mathcal{P}$ est vraie pour tout $k \in \mathbb{N}^*$. $\square$

Exemple 23. Calculer les itérés de

$$f : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (2x + y - 2z, 2x + 3y - 4z, x + y - z) \end{array}$$

▷ Cela revient à calculer les puissances de

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

Or d'après Proposition 1.27, $A^n = PD^n P^{-1}$.

1.5.2 Application à la résolution de systèmes linéaires

La résolution de systèmes linéaires en plus grande dimension se ramène à l'écriture d'une équation linéaire matricielle de la forme $X_{n+1} = AX_n$. C'est l'analogie des suites géométriques réelles en dimensions supérieures. on a

Proposition 1.28 (► Suite géométrique matricielle). Soit $A \in \mathcal{M}_N(\mathbb{K})$. Considérons le système $X_{n+1} = AX_n$. Alors pour tout $n \in \mathbb{N}^*$, $X_n = A^n X_0$.

- *Démonstration.* Soit la proposition $\mathcal{P}(n) : "X_n = A^n X_0"$.
- $\mathcal{P}(1)$ est vraie.
 - Supposons que $\mathcal{P}(n)$ soit vraie. Alors on a $X_{n+1} = AX_n = AA^n X_0 = A^{n+1} X_0$. Donc $\mathcal{P}$ est héréditaire et $\mathcal{P}$ est vraie pour tout $n \geq 1$

$\square$

On est donc ramenés au calcul de la puissance d'une matrice.

Exemple 24. *Considérons le système linéaire*

$$\begin{cases} x_{n+1} &= 2x_n + y_n - 2z_n \\ y_{n+1} &= 2x_n + 3y_n - 4z_n \\ z_{n+1} &= x_n = y_n - z_n \end{cases}$$

où x_0 , y_0 et z_0 sont donnés.

▷ Ce système est équivalent à $X_{n+1} = AX_n$ avec

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

D'après Proposition 1.28 on a $X_n = A^n X_0$ où d'après Proposition 1.27, $A^n = PD^nP^{-1}$.

1.5.3 Application à la résolution de suites linéaires d'ordre supérieur

Rappelons qu'une suite linéaire est une suite de la forme

$$x_{n+k} = f(x_{n+k-1}, \dots, x_1)$$

où f est une application linéaire. Ces suites récurrentes peuvent s'écrire sous la forme $X_{n+1} = AX_n$. On peut alors utiliser le résultat de la section précédente.

Exemple 25. *On considère la suite récurrente linéaire d'ordre 2 suivante*

$$x_{n+2} = x_{n+1} - x_n$$

où x_0 et x_1 sont donnés.▷ Ce système est équivalent à $X_{n+1} = AX_n$ avec

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

D'après Proposition 1.28 on a $X_n = A^n X_0$ où d'après Proposition 1.27, $A^n = PD^nP^{-1}$.

1.5.4 Application à la géométrie d'une application linéaire

Exemple 26. *Décrire géométriquement l'application*

$$f : \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (x, y) & \mapsto & (2x, -2y) \end{array}$$

▷ La représentation matricielle de f est donnée par

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

Le polynôme caractéristique est $\chi_A = X(X - 2)$ donc $\text{Sp}(A) = \{0, 2\}$. On a $E_2 = \text{Vect}(1, -1)$ et $E_0 = \text{Vect}\{(0, 1)\}$. Donc A est diagonalisable et dans la base de vecteurs propres $\{(1, -1), (0, 1)\}$ la représentation matricielle de f est donnée par

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

Dans cette base l'application linéaire est la composée (commutative) d'une projection sur le deuxième vecteur de base et d'une homothétie de rapport 2.

Dans la base canonique d'origine f représente donc la composée entre une homothétie de rapport 2 et une projections sur l'axe $\text{Vect}\{(1, -1)\}$.

1.5.5 Application aux racines d'une matrices

Si A est diagonalisable de la forme $A = PDP^{-1}$ avec $D = \text{diag}(\lambda_1, \dots, \lambda_N)$ alors on peut définir $M = \text{diag}(\lambda_1^{1/k}, \dots, \lambda_N^{1/k})$ (avec la restriction $\lambda_i > 0$ pour tout $i \in \{1, \dots, N\}$ si k est pair). On a alors $M^k = D$. Ainsi $B = PMP^{-1}$ vérifie, pour tout $k \geq 1$, $B^k = PM^kP^{-1} = PDP^{-1} = A$.

1.6 Exercices

Rappels

Exercice 1.1. 1. Montrer que $\mathcal{F} := \{(1, 1, 2), (1, 0, 2), (0, 1, 1)\}$ forme une base de $\mathbb{R}^3$.

2. Soit x le vecteur dont les coordonnées dans la base $\mathcal{F}$ est $(1, 1, 0)$. Déterminer ses coordonnées dans la base canonique de $\mathbb{R}^3$.

3. Soit x le vecteur dont les coordonnées dans la base canonique de $\mathbb{R}^3$ sont $(1, 0, 1)$. Déterminer ses coordonnées dans la base $\mathcal{F}$.

Exercice 1.2. Soit

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \\ (x, y) \mapsto (x + y, x, x - 2y)$$

1. Montrer que f est une application linéaire.

2. Donner la représentation matricielle de f dans les bases canoniques de $\mathbb{R}^2$ et $\mathbb{R}^3$.

3. Montrer que $\mathcal{E} := \{(1, 1), (1, -1)\}$ forme une base de $\mathbb{R}^2$ et $\mathcal{F} := \{(1, 1, 1), (1, -1, 0), (1, 0, -1)\}$ forme une base de $\mathbb{R}^3$.

4. Donner la représentation matricielle de f dans les bases $\mathcal{E}, \mathcal{F}$.

Exercice 1.3. Soit le système défini par récurrence par

$$\begin{cases} x_{n+1} = x_n + 2y_n + z_n \\ y_{n+1} = y_n + 2z_n \\ z_{n+1} = z_n \end{cases}$$

et où x_0, y_0 et z_0 sont donnés.

1. Ecrire le système sous la forme

$$X_{n+1} = AX_n \quad \text{où} \quad X_n := \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

2. Calculer A^n pour tout $n \in \mathbb{N}$, en utilisant la formule du binôme de Newton.

3. Exprimer x_n, y_n et z_n en fonction de n, x_0, y_0 et z_0 .

Diagonalisation

Exercice 1.4. On considère l'application

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \\ (x, y, z) \mapsto \begin{pmatrix} 2x + 2y \\ x + 2y + z \\ 2y + 2z \end{pmatrix}$$

1. Déterminer la représentation matricielle A de f dans la base canonique de $\mathbb{R}^3$.

2. Déterminer le spectre de A .

3. Déterminer des bases des espaces propres de A .

4. Déterminer la représentation matricielle de f dans une base de vecteurs propres.

5. A est-elle diagonalisable dans $\mathbb{R}$? Dans $\mathbb{C}$? (essayer d'utiliser divers résultats du cours).
6. Déterminer leur décomposition en produit $A = P D P^{-1}$ où D est une matrice diagonale et P est une matrice inversible.

Exercice 1.5. Pour chacune des matrices suivantes

1. La matrice est-elle diagonalisable dans $\mathbb{R}$? Dans $\mathbb{C}$? (essayer d'utiliser divers résultats du cours).
2. Déterminer leur décomposition en produit $A = P D P^{-1}$ où D est une matrice diagonale et P est une matrice inversible.

$$A_1 := \begin{bmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{bmatrix} \quad A_2 := \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix} \quad A_3 := \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$A_4 := \begin{bmatrix} -3 & -1 & 6 \\ 6 & 4 & -12 \\ 1 & 1 & -2 \end{bmatrix} \quad A_5 := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 1 \\ 2 & -1 & 1 & 1 \\ 3 & -1 & -1 & 3 \end{bmatrix} \quad A_6 := \begin{bmatrix} 3 & -1 & 7 & 14 \\ 4 & -1 & 7 & -15 \\ 0 & 0 & 3 & -4 \\ 0 & 0 & 2 & -3 \end{bmatrix}$$

$$A_7 := \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \quad A_8 := \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} \quad A_9 := \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & 1 & 2 \end{bmatrix}$$

Applications de la diagonalisation

Exercice 1.6. Le système défini par récurrence par

$$\begin{cases} x_{n+1} = 2x_n + 2y_n \\ y_{n+1} = x_n + 2y_n + z_n \\ z_{n+1} = 2y_n + 2z_n \end{cases}$$

où x_0, y_0, z_0 et t_0 sont donnés. Exprimer x_n, y_n, z_n et t_n en fonction de n, x_0, y_0, z_0 et t_0 .

Exercice 1.7. Soit la suite définie pour tout $n \in \mathbb{N}$ par

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

où u_0 et u_1 sont donnés.

1. Déterminer A tel que $X_{n+1} = AX_n$ où $X_n = (u_{n+1} \ u_n)^T$.
2. Calculer A^n pour tout $n \in \mathbb{N}$.
3. Exprimer, pour tout $n \geq 2$, u_n en fonction de u_0 et u_1 .

Parce qu'il n'y a pas que $\mathbb{R}^N$

Exercice 1.8. On considère

$$\begin{aligned} \varphi : \mathbb{R}_2[X] &\rightarrow \mathbb{R}_2[X] \\ P &\mapsto P - (X+1)P' \end{aligned}$$

1. Justifier que φ définit un endomorphisme de $\mathbb{R}_n[X]$.
2. Déterminer la représentation matricielle de φ dans la base canonique de E .
3. Déterminer la représentation matricielle de φ dans la base $\{1, X-1, (X-1)^2\}$.
4. Diagonaliser φ .
5. En déduire les vecteurs propres de φ .

Exercice 1.9. Soit $E = \mathbb{R}^{\mathbb{N}}$ l'ensemble des suites réelles. Soit $f : E \rightarrow E$ l'endomorphisme qui, à toute suite $(u_n)_n$ associe la suite $(v_n)_n$ définie par :

$$v_0 = u_0 \quad \text{et} \quad v_n = \frac{u_n + u_{n-1}}{2}.$$

Déterminer les éléments propres de f .

1.7 Corrigés des exercices

Rappels

Corrigé Exercice 1.1. 1. On calcule

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

Donc le déterminant est différent de zéro et la famille forme une base de $\mathbb{R}^3$.

2. On a

$$X_{\mathcal{C}_4} = P_{\mathcal{C}_4}^{\mathcal{F}} X_{\mathcal{F}}.$$

Or la matrice de la question précédente représente la matrice de passage de $\mathcal{C}_4$ à $\mathcal{F}$:

$$P_{\mathcal{C}_4}^{\mathcal{F}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

Donc

$$X_{\mathcal{C}_4} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

Donc les coordonnées du vecteur dans la base canonique est $(2, 1, 4)$.

3. On a

$$X_{\mathcal{F}} = P_{\mathcal{F}}^{\mathcal{C}_4} X_{\mathcal{C}_4}.$$

La matrice de passage de $\mathcal{F}$ à $\mathcal{C}_4$ est l'inverse de P la matrice de passage de $\mathcal{C}_4$ à $\mathcal{F}$ et vaut

$$P_{\mathcal{F}}^{\mathcal{C}_4} = \begin{pmatrix} 2 & 1 & -1 \\ -1 & -1 & 1 \\ -2 & 0 & 1 \end{pmatrix}$$

Finalement

$$X_{\mathcal{F}} = \begin{pmatrix} 2 & 1 & -1 \\ -1 & -1 & 1 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

On peut vérifier aisément que

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 1 * \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 0 * \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + (-1) * \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Corrigé Exercice 1.2. 1. Soit

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

L'application $X \mapsto AX$ est linéaire et elle est isomorphe à f donc f est une application linéaire.

2. C'est la matrice

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

3. On calcule

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

donc $\mathcal{E} = \{(1, 1), (1, -1)\}$ forme une base de $\mathbb{R}^2$. De même

$$\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = 3$$

donc $\mathcal{F} = \{(1, 1, 1), (1, -1, 0), (1, 0, -1)\}$ forme une base de $\mathbb{R}^3$.

4. On utilise la formule du cours :

$$\mathcal{M}_{\mathcal{F}}^{\mathcal{E}}(f) = P_F^{\mathcal{C}_4} \text{mat}(f, \mathcal{C}_3, \mathcal{C}_4) P_{\mathcal{C}_3}^{\mathcal{E}}.$$

où $\mathcal{C}_4$ est la base canonique de $\mathbb{R}^4$ et $\mathcal{C}_3$ est la base canonique de $\mathbb{R}^3$. Comme

$$P_F^{\mathcal{C}_4} = (P_{\mathcal{C}_4}^{\mathcal{F}})^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

On obtient

$$\text{mat}(f, \mathcal{E}, \mathcal{F}) = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & 4 \\ -1 & 1 \\ 5 & -5 \end{pmatrix}.$$

Corrigé Exercice 1.3. 1. Le système est équivalent à $X_{n+1} = AX_n$ avec

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

2. A peut-être écrite sous la forme $I + N$ où

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

On calcule

$$N^2 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et $N^k = 0$ pour tout $k \geq 3$.

Comme $IN = NI$ on peut appliquer la formule du binôme de Newton. On obtient

$$(I + N)^n = I + nN + \frac{n(n-1)}{2}N^2 = \begin{pmatrix} 1 & 2n & n + 2n(n-1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{pmatrix}$$

On remarque que A^1 est bien égale à A .

3. On a $X_n = A^n X_0$ ce qui donne

$$\begin{cases} x_n = x_0 + 2ny_0 + [n + 2n(n-1)]z_0 \\ y_n = y_0 + 2nz_0 \\ z_n = z_0 \end{cases}$$

Diagonalisation

Corrigé Exercice 1.4. 1. On a

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

2. On obtient $\chi_{A_1} = -X^3 + 6X^2 - 8X = -X(X-2)(X-4)$. Donc $\text{Sp}(A_1) = \{0, 2, 4\}$. Remarquons que $0 + 2 + 4 = 2 + 2 + 2 = \text{Tr}(A_1)$. Remarquons surtout que comme on est dans $\mathbb{R}^3$ et qu'il y a trois valeurs propres distinctes on sait déjà que la matrice est diagonalisable.
3. On calcule $E_0 = \ker(A_1 - 0I) = \text{Vect}\{(1, -1, 1)\}$. On peut aisément vérifier que $A_1(1, -1, 1) = (0, 0, 0) = 0 * (1, -1, 1)$. De même on calcule $E_2 = \ker(A_1 - 2I) = \text{Vect}\{(-1, 0, 1)\}$ et on peut vérifier que $A_1(-1, 0, 1) = 1 * (-1, 0, 1)$. De même $E_4 = \ker(A_1 - 4I) = \text{Vect}\{(1, 1, 1)\}$ et on peut vérifier $A_1(1, 1, 1) = (4, 4, 4) = 4 * (1, 1, 1)$.
4. La représentation matricielle de f dans la base de vecteurs propres est $D = \text{diag}\{0, 2, 4\}$.
5. On a déjà vu dans la question 1 lors de la détermination du spectre que la matrice était diagonalisable dans $\mathbb{R}$. On peut aussi utiliser le critère de multiplicité puisque le polynôme caractéristique est scindé et que la dimension des espace propre est nécessairement de dimension 1 puisque $1 \leq \dim E_\lambda \leq 1$ pour tout λ dans le spectre. On peut aussi utiliser le critère de dimension puisque la somme des dimensions des espaces propres est égal à 3 qui est la dimension de $\mathbb{R}^3$, l'espace dans lequel nous travaillons.
6. On a donc $A = PDP^{-1}$ avec $D = \text{diag}\{0, 2, 4\}$ et

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

Corrigé Exercice 1.5. • A_1 - On obtient $\chi_{A_1} = -X^3 + 4X^2 - 5X + 2 = -(X-2)(X-1)^2$. Donc $\text{Sp}(A_1) = \{1, 2\}$. Remarquons que $1 + 1 + 2 = \text{Tr}(A_1)$. À cette étape on ne sait pas si la matrice est diagonalisable. D'après le critère de multiplicité elle le sera si et seulement si $\dim E_1 = 2$ car le polynôme caractéristique est scindé.

On commence par calculer $E_1 = \text{Vect}\{(2, 0, 1), (-1, 1, 0)\}$. Comme E_1 est de dimension 2, la matrice est diagonalisable. On calcule alors $E_2 = \text{Vect}\{(1, 2, 1)\}$,

On a $D = \text{diag}\{1, 1, 2\}$ et

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

- A_2 - On obtient $\chi_{A_2} = -X^3 + 12X - 16 = -(X+4)(X-2)^2$. Donc $\text{Sp}(A_2) = \{-4, 2\}$. Remarquons que $-4 + 2 + 2 = \text{Tr}(A_2)$. À cette étape on ne sait pas si la matrice est diagonalisable. D'après le critère de multiplicité elle le sera si et seulement si $\dim E_2 = 2$ car le polynôme caractéristique est scindé.

On calcule $E_2 = \text{Vect}\{(1, 1, 0)\}$. Donc la matrice n'est pas diagonalisable car la multiplicité de la valeur propre 2 dans le polynôme caractéristique est égale à 2 et que la dimension de l'espace propre est 1.

- A_3 - On obtient $\chi_{A_3} = -X^3 + X^2 + X - 1 = -(X+1)(X-1)^2$. Donc $\text{Sp}(A_3) = \{-1, 1\}$. Remarquons que $-1 + 1 + 1 = \text{Tr}(A_3)$.

On $E_1 = \text{Vect}\{(0, 1, 0), (1, 0, 1)\}$. La dimension de E_1 est égal à la multiplicité de 1 dans le polynôme caractéristique et la dimension de E_{-1} est nécessairement 1 car $1 \leq \dim E_{-1} \leq 1$. Donc la matrice est diagonalisable. On calcule $E_{-1} = \text{Vect}\{(-1, 0, 1)\}$.

On a $D = \text{diag}\{-1, 1, 1\}$ et

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

- A_4 - On a $\chi_{A_4} = -X(X-1)(X+2)$ donc $\text{Sp}(A_4) = \{-2, 0, 1\}$. Donc la matrice est diagonalisable. On a finalement $D = \text{diag}\{-2, 0, 1\}$ et

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

- A_5 - On a $\chi_{A_5} = (X-1)^2(X-2)^2$ donc $\text{Sp}(A_5) = \{1, 2\}$. On calcule $\dim E_1 = 2$ et $\dim E_2 = 2$. Donc par le critère de multiplicité la matrice est diagonalisable. On a finalement $D = \text{diag}\{1, 1, 2, 2\}$ et

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

- A_6 - On a $\chi_{A_6} = (X+1)(X-1)^3$ donc $\text{Sp}(A_6) = \{-1, 1\}$. On calcule $\dim E_1 = 1$ donc par le critère de multiplicité la matrice n'est pas diagonalisable.
- A_7 - On a $\chi_{A_7} = X^2 + X + 1$ donc par le critère de multiplicité la matrice n'est pas diagonalisable dans $\mathbb{R}$. Dans $\mathbb{C}$, on a $\text{Sp}(A_7) = \{j, \bar{j}\}$ et la matrice est diagonalisable. On obtient $D = \text{diag}\{1, \bar{j}, j\}$ et

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

- A_8 - On a $\chi_{A_8} = X^2$ donc $\text{Sp}(A_8) = \{0\}$. La matrice n'est pas diagonalisable ni dans $\mathbb{R}$ ni dans $\mathbb{C}$ car les seules matrices diagonalisables ayant une seule valeur propre sont les matrices scalaires. En effet si A est diagonalisable avec $\text{Sp}(A) = \{\lambda\}$ alors on aurait $A = P(\lambda I)P^{-1} = \lambda I$.
- A_9 - On a $\chi_{A_9} = -(X-1)^2(X-2)$ donc $\text{Sp}(A_9) = \{1, 2\}$. On calcule l'espace propre associé à 1. Comme $\dim E_1 = 2$ la matrice n'est pas diagonalisable par le critère de multiplicité.

Applications de la diagonalisation

Corrigé Exercice 1.6. Pour tout $n \geq 1$, le système est équivalent à $X_{n+1} = AX_n$. D'après Proposition 1.28 on a donc pour tout $n \geq 1$, $X_n = A^n X_0$. Pour calculer A^n nous allons diagonaliser A . D'après Corrigé 1.4, on a

$$\begin{aligned} A = PDP^{-1} &= \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+1} + 4^n & 2 * 4^n & -2^{n+1} + 4^n \\ 4^n & 2 * 4^n & 4^n \\ -2^{n+1} + 4^n & 2 * 4^n & 2^{n+1} + 4^n \end{pmatrix} \end{aligned}$$

On vérifie que pour $n = 1$ on a bien $A^1 = A$. Donc

$$\begin{cases} x_n = [2^{n+1} + 4^n]x_0 + 2 * 4^n y_0 + [-2^{n+1} + 4^n]z_n \\ y_n = 4^n x_0 + 2 * 4^n y_0 + 4^n z_0 \\ z_n = [-2^{n+1} + 4^n]x_0 + 2 * 4^n y_0 + [2^{n+1} + 4^n]z_0 \end{cases}$$

Corrigé Exercice 1.7. 1. On pose

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

2. On diagonalise A . A est diagonalisable car son spectre est composé de deux éléments distincts. On a $A = PDP^{-1}$ avec $D = \text{diag}\{-1, 2\}$ et

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

Par la proposition du cours on a donc

$$A^n = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} (-1)^n + 2^{n+1} & 2(-1)^{n+1} + 2^{n+1} \\ (-1)^{n+1} + 2^n & 2(-1)^n + 2^n \end{pmatrix}$$

On remarque que $A^1 = A$.

3. Donc on a pour tout $n \geq 1$

$$u_n = [(-1)^n + 2^{n+1}]x_0 + [2(-1)^{n+1} + 2^{n+1}]y_0.$$

Parce qu'il n'y a pas que $\mathbb{R}^N$

Corrigé Exercice 1.8. 1. Pour tout P et Q dans $\mathbb{R}_2[X]$ et tout $\lambda \in \mathbb{R}$ on a

$$\begin{aligned} \varphi(P + \lambda Q) &= (P + \lambda Q) - (X + 1)(P + \lambda Q)' = P + \lambda Q - (X + 1)(P' + \lambda Q') \\ &= P + \lambda Q - (X + 1)P' - (X + 1)\lambda Q' = P - (X + 1)P' + \lambda Q - (X + 1)\lambda Q' \\ &= \varphi(P) + \lambda\varphi(Q). \end{aligned}$$

Donc φ est une application linéaire. De plus si $P \in \mathbb{R}_2[X]$ alors $P' \in \mathbb{R}_1[X]$ et $(X + 1)P' \in \mathbb{R}_2[X]$. Donc φ est un endomorphisme de $\mathbb{R}_2[X]$.

2. On a

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

3. On a

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

4. On calcule $\text{Sp}(A) = \{-1, 0, 1\}$. Donc A est diagonalisable. On a $A = PDP^{-1}$ avec $D = \text{diag}\{-1, 0, 1\}$ et

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

5. On a donc que $1 + 2X + X^2$ est vecteur propre associé à la valeur propre -1 , $1 + X$ est vecteur propre associé à la valeur propre 0 et 1 est vecteur propre associé à la valeur propre 1 .

Corrigé Exercice 1.9. Soit (u_n) un vecteur propre associé à la valeur propre λ . Alors on a $u_0 = \lambda u_0$ et, pour tout $n \geq 1$,

$$\frac{u_n + u_{n-1}}{2} = \lambda u_n \iff (1 - 2\lambda)u_n = -u_{n-1}.$$

On distingue alors trois cas :

- Si $\lambda = 1$, alors on a $u_0 = u_0$, puis pour tout $n \geq 1$, on a $u_n = u_{n-1}$. Réciproquement, toute suite constante et non-nulle est bien vecteur propre de ϕ pour la valeur propre 1 . On en déduit que 1 est une valeur propre de ϕ dont l'espace propre associé est constitué par les suites constantes.
- Si $\lambda = 1/2$, alors l'équation $u_0 = \lambda u_0$ donne $u_0 = 0$ et on a aussi, pour tout $n \geq 1$, $u_{n-1} = 0$ ce qui implique que (u_n) est la suite nulle et donc $1/2$ n'est pas valeur propre de ϕ .
- Dans tous les autres cas, alors $u_0 = \lambda u_0$ donne encore $u_0 = 0$ et pour tout $n \geq 1$,

$$u_n = \frac{1}{2\lambda - 1} u_{n-1}.$$

Ainsi, la suite (u_n) est là encore la suite nulle, et λ n'est pas valeur propre.

En conclusion, la seule valeur propre est 1 , et les seuls vecteurs propres sont les suites constantes.

Chapitre 2

Trigonalisation

Comme nous l'avons vu au chapitre 1, certaines matrices ne sont pas diagonalisables, ni dans $\mathbb{R}$ ni même dans sa clôture algébrique $\mathbb{C}$. Le résultat théorique principal de ce chapitre est qu'il est toujours possible, au moins dans le corps des complexes, d'obtenir une matrice semblable à une matrice triangulaire. Rappelons la définition d'une matrice triangulaire :

Définition 2.1 (► Matrice triangulaire). Soit $A \in \mathcal{M}_N(\mathbb{K})$. La matrice A de terme général $(a_{ij})_{(i,j) \in \{1, \dots, N\}^2}$ est triangulaire supérieure si pour tout $i > j$, $a_{ij} = 0$. Une matrice est triangulaire inférieure si A^T est triangulaire supérieure.

Exemple 2.1. La matrice

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

est triangulaire supérieure.

Il est alors tentant de penser que la formule de Newton permet alors de calculer aisément les puissances d'une matrice. Ceci n'est vrai, comme le montre l'exemple suivant :

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

En effet, M peut être écrite sous la forme $D + N$ où

$$N := \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad D := \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \text{diag}(-1, 2, 2).$$

Mais

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{alors que} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

On ne peut donc pas appliquer la formule du binôme de Newton comme précédemment pour cette matrice.

On peut donc calculer la puissance d'une matrice triangulaire en utilisant la formule du binôme de Newton uniquement lorsqu'elle peut être écrite somme commutative d'une matrice diagonale et d'une matrice nilpotente. En pratique la détermination de cette décomposition, peut être assez complexe en grande dimension. Nous allons, dans ce chapitre, introduire une technique simple pour trigonaliser des matrices en utilisant la méthode de Jordan. Cette méthode s'étend aux dimensions plus grandes. Si la technique s'avère parfois plus complexe, les techniques que nous verrons ici permettront de déterminer la réduite de Jordan de la plupart des matrices que l'on rencontre dans les exercices académiques.

En développant la théorie de ce chapitre nous revisiterons quelques questions que l'on sait résoudre par d'autres méthodes notamment le calcul de l'inverse d'une matrice en utilisant le théorème de Cayley-Hamilton, un critère de diagonalisation basé sur le polynôme minimal, etc.

2.1 Polynôme d'endomorphisme

Définition 2.2 (► Polynôme d'endomorphisme). Soient E un espace vectoriel, $u \in \mathcal{L}(E)$ et $P \in \mathbb{R}[X]$ de la forme $P = \sum_{k=0}^n a_k X^k$. On appelle polynôme de u et on note $P(u)$ l'endomorphisme défini par

$$\begin{aligned} E &\rightarrow E \\ u &\mapsto \sum_{k=0}^n a_k u^k \end{aligned}$$

où $u^0 = \text{id}$ et pour tout $k \in \mathbb{N}$, u^k désigne la k -ème itérée de u .

Cette définition et les suivantes s'étendent de façon évidente à l'ensemble des matrices.

Exemple 2.2. Soit $P = X^3 + 2X^2 - X + 1$. Le polynôme d'endomorphisme est $P(u) = u^3 + 2u^2 - u + \text{id}$.

Définition 2.3 (► Polynôme annulateur). Soient E un espace vectoriel, $u \in \mathcal{L}(E)$ et $P \in \mathbb{R}[X]$. Le polynôme P est dit polynôme annulateur de u si $P(u) \equiv 0$.

Exemple 2.3. Le polynôme $X^2 - 2X + 1$ est polynôme annulateur de

$$\begin{aligned} f : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (x + y, y) \end{aligned}$$

En effet la représentation matricielle de f dans la base canonique est :

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

Et on calcule

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

Donc $X^2 - 2X + 1$ est polynôme annulateur de f (et de A).

Remarque. L'ensemble des polynômes annulateurs de u est un idéal de $\mathbb{R}[X]$. Comme il n'est pas réduit à $\{0\}$, il admet un unique générateur unitaire, noté m_u , appelé polynôme minimal. Pour ne pas perdre les étudiant.e.s qui n'auront pas une maîtrise parfaite de leur cours d'algèbre sur les idéaux nous proposons ici une définition plus simple du polynôme minimal.

Définition 2.4 (► Polynôme minimal). Soient E un espace vectoriel, $u \in \mathcal{L}(E)$. Le polynôme minimal de u est le plus petit (au sens de la division euclidienne) polynôme annulateur de u dont le coefficient de plus haut degré est 1.

Comme nous avons déjà dit que le polynôme minimal permet de déterminer si une matrice est diagonalisable, nous imaginons l'enthousiasme de l'étudiant.e. passionné.e et son désir d'apprendre comment trouver le polynôme minimal d'une matrice quelconque. Pour l'étudiant.e impatient.e je dévoile l'astuce : il s'avère que le polynôme minimal est un diviseur du polynôme caractéristique. C'est une conséquence du théorème de Cayley-Hamilton que nous verrons Section 2.3. Il existe de nombreuses démonstrations de ce résultat important mais nous pensons que la plus simple repose sur un théorème de trigonalisation que nous nous proposons de démontrer dans la section suivante.

2.2 Théorème de trigonalisation

Définition 2.5 (► Trigonalisation). Soit $A \in \mathcal{M}_N(\mathbb{K})$. La matrice A est trigonalisable dans $\mathbb{K}$ si elle est semblable dans $\mathbb{K}$ à une matrice T triangulaire.

Par la formule de changement de base dans une application linéaire, Proposition 1.8, la trigonalisation consiste à déterminer une base dans laquelle la représentation matricielle (de l'application linéaire sous-jacente à A) est triangulaire.

Théorème 2.6 (► Critère de trigonalisation). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. La matrice A est trigonalisable dans $\mathbb{K}$ si et seulement si son polynôme caractéristique est scindé dans $\mathbb{K}$.*

► *Démonstration.* $\Rightarrow$) Si A est trigonalisable dans $\mathbb{K}$ alors A est semblable à une matrice T triangulaire. Appelons α_i les termes sur la diagonale de T . Alors

$$\chi_T = \prod_{i=1}^N (\alpha_i - X)$$

est scindé. Or d'après Proposition 1.16, le polynôme caractéristique est le même pour toute matrice semblable, donc χ_A aussi est scindé. On a montré dans le même temps que les valeurs propres d'une matrice triangulaire sont ses éléments diagonaux.

$\Leftarrow$) On va procéder par récurrence sur la dimension. Soit $\mathcal{P}$ la proposition définie pour $n \in \mathbb{N}^*$ par $\mathcal{P}(n)$: "tout endomorphisme d'un espace vectoriel de dimension n dont le polynôme caractéristique est scindé est trigonalisable".

— $\mathcal{P}(1)$ est vraie.

— Supposons que $\mathcal{P}(n)$ soit vraie à un rang n donné et considérons un endomorphisme d'un espace vectoriel E de dimension $n + 1$ dont le polynôme caractéristique χ_A est scindé. Comme χ_A est scindé, il y a au moins une racine a de χ_A , qui est aussi une valeur propre. Soit $u \neq 0$ le vecteur propre associé à a . On complète la famille $\{u\}$ en une base $\mathcal{E} := \{u, u_1, \dots, u_n\}$ de E . Soit $Q = \mathcal{P}_{\mathcal{E}}^{\mathcal{E}}$. Par la formule de changement de base, Proposition 1.8, on a

$$Q^{-1}AQ = \begin{pmatrix} a & v^T \\ 0 & B \end{pmatrix}$$

où $v \in \mathbb{K}^n$ et $B \in \mathcal{M}_n(\mathbb{K})$. On peut appliquer l'hypothèse de récurrence à B : il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ et une matrice triangulaire $T \in \mathcal{M}_n(\mathbb{K})$ telles que $B = PTP^{-1}$ ou $P^{-1}BT = T$. On pose

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

On a alors

$$(Q')^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix}$$

On calcule

$$(Q')^{-1}Q^{-1}AQQ' = \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} \begin{pmatrix} a & v^T \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} a & v^T P \\ 0 & T \end{pmatrix}$$

qui est triangulaire. Donc A est trigonalisable dans $\mathbb{K}$. □

Comme le corps $\mathbb{C}$ des complexes est algébriquement clôt, tout polynôme est scindé dans $\mathbb{C}$ et on a :

Corollaire 2.7 (► Trigonalisation dans $\mathbb{C}$). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. La matrice A est trigonalisable dans $\mathbb{C}$.*

Exemple 27. *Soit*

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

1. A est-elle diagonalisable dans $\mathbb{R}$?
2. A est-elle diagonalisable dans $\mathbb{C}$?

▷

1. On calcule $\chi_A = X^2 + 1$. Le polynôme caractéristique n'est pas scindé dans $\mathbb{R}$ donc A n'est pas trigonalisable dans $\mathbb{R}$
2. Toute matrice est trigonalisable dans $\mathbb{C}$.

Exemple 28. Trigonaliser

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

▷ On calcule $\chi_A = -(X - 3)(X - 2)^2$. On détermine $E_2 = \text{Vect}\{(4, 3, 4)\}$. La multiplicité algébrique de 2 est différente de la dimension de son espace propre donc par le critère de multiplicité, Théorème 1.23, A n'est pas diagonalisable. Par contre comme χ_A est scindée dans $\mathbb{R}$, A est trigonalisable dans $\mathbb{R}$. On détermine $E_3 = \text{Vect}\{(1, 1, 1)\}$. Quelque soit ε_3 tel que $\{(1, 1, 1), (4, 3, 4), \varepsilon_3\}$ forme une base A est semblable à

$$\begin{pmatrix} 3 & 0 & \cdot \\ 0 & 2 & \cdot \\ 0 & 0 & \cdot \end{pmatrix}$$

Donc quelque soit ε_3 , A est semblable à une matrice triangulaire supérieure. Si on prend, par exemple $\varepsilon_3 = (0, 0, 1)$ on a $A\varepsilon_3 = (-2, -3, 0) = -6(1, 1, 1) + (4, 3, 4) + 2(0, 0, 1)$ donc $A = PTP^{-1}$ avec

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

On peut écrire cette matrice sous la forme $D + N$ avec $D = \text{diag}\{3, 2, 2\}$ et

$$N = \begin{pmatrix} 0 & 0 & -6 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Par contre on a

$$DN = \begin{pmatrix} 0 & 0 & -18 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & -12 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = ND$$

Donc on ne peut pas appliquer la formule du binôme de Newton à cette matrice T . Nous verrons Section 2.5 la méthode de Jordan qui permet de trigonaliser de sorte de s'assurer de pouvoir appliquer la formule du binôme de Newton.

2.3 Théorème de Cayley-Hamilton

Théorème 2.8 (► Théorème de Cayley-Hamilton). Soient E un espace vectoriel et $f \in \mathcal{L}(E)$. On a $\chi_f(f) \equiv 0$.

En terme matricielle on a : pour tout $A \in \mathcal{M}_N(\mathbb{K})$, $\chi_A(A) \equiv 0$ La démonstration que nous proposons ici repose sur le résultat suivant :

Lemma 2.9 (►). Soit E un espace vectoriel et $f \in \mathcal{L}(E)$. Soit P un polynôme annulateur de f . Alors

$$\lambda \in \text{Sp}(f) \Rightarrow P(\lambda) = 0.$$

► *Démonstration.* Soit $P = \sum a_k X^k$ un polynôme annulateur de f c'est à dire que $0 \equiv \sum a_k f^k$. Comme $\lambda \in \text{Sp}(f)$, il existe $u \neq 0$ tel que $f(u) = \lambda u$. D'où

$$f^k(u) = f^{k-1}(f(u)) = f^{k-1}(\lambda u) = \lambda f^{k-1}(u) = \dots = \lambda^k u.$$

En injectant dans le polynôme annulateur on a

$$0 = \sum a_k f^k(u) = \sum a_k \lambda^k u = P(\lambda) u.$$

Or le vecteur propre u est différent de 0, ce qui implique que $P(\lambda) = 0$. □

► *Démonstration de Théorème 2.8.* Plaçons nous dans la clôture algébrique de $\mathbb{K}$, donc ici $\mathbb{C}$. D'après Corollaire 2.7, f est trigonalisable dans $\mathbb{C}$, c'est-à-dire qu'il existe $(\lambda_i)_{i \in \{1, \dots, N\}}$ pas forcément distincts et éventuellement complexes tels que

$$\chi_f = \prod_{k=1}^N (\lambda_k - X).$$

Soit $(e_i)_{i \in \{1, \dots, N\}}$ la base dans laquelle la représentation matricielle de f est triangulaire (supérieure). On a

$$(\lambda_1 \text{id} - f)(e_1) = 0.$$

Comme les $(\lambda_i \text{id} - f)$ commutent on prouve par récurrence que pour tout $j \in \{1, \dots, n\}$,

$$\chi_f(f)(e_j) = (\lambda_1 \text{id} - f) \circ (\lambda_2 \text{id} - f) \circ \dots \circ (\lambda_n \text{id} - f)(e_j) = 0.$$

Il en résulte que $\chi_f(f)$ est nul pour tout vecteur de la base et donc que $\chi_f(f) = 0$. $\square$

2.3.1 ► Application au calcul de l'inverse d'une matrice

Lorsque A est inversible, 0 n'est pas valeur propre donc le polynôme caractéristique est de la forme $a_0 + \sum_{k=1}^N a_k X^k$. D'après le théorème de Cayley-Hamilton on a

$$a_0 I + \sum_{k=1}^N a_k A^k = 0$$

D'où

$$\sum_{k=1}^N a_k A^k = -a_0 I$$

que l'on peut écrire

$$A \left[-\frac{1}{a_0} \sum_{k=1}^N a_k A^{k-1} \right] = I$$

D'où l'on déduit

$$A^{-1} = -\frac{1}{a_0} \sum_{k=1}^N a_k A^{k-1}.$$

Le calcul de l'inverse d'une matrice se réduit ainsi à la détermination de puissance d'une matrice, qu'il est souvent bien plus facile de calculer.

Remarquons aussi que, à l'inverse du processus de diagonalisation, il n'est pas nécessaire de savoir déterminer les racines du polynôme caractéristique pour appliquer cette méthode.

Exemple 29. *Calculer*

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1}$$

► On calcule $\chi_P = X^2 - 3X + 1$. D'après le théorème de Cayley-Hamilton on a

$$P^2 - 3P + I = 0$$

d'où

$$I = P(-P + 3I)$$

Donc

$$P^{-1} = -P + 3I = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

2.3.2 ► Application à la détermination du polynôme minimal

D'après Définition 2.4, le polynôme minimal est un diviseur des polynômes annulateurs. Or d'après le théorème de Cayley-Hamilton, Théorème 2.8, le polynôme caractéristique est un polynôme annulateur. Donc le polynôme minimal est à rechercher parmi les diviseurs du polynôme caractéristique. Par ailleurs Lemme 2.9 indique que les racines du polynôme minimal sont les mêmes que les racines du polynôme caractéristique. Dans $\mathbb{R}$ tout polynôme peut-être écrit sous la forme

$$C \prod_{i=1}^k (\lambda_i - X)^{\alpha_i} \prod_{i=1}^l (X^2 + b_i X + c_i)^{\beta_i}$$

où $(C, \lambda, b_i, c_i) \in \mathbb{R}^4$ et $(\alpha_i, \beta_i) \in \mathbb{N}^2$. Dans le cas où l'on travaille dans le corps des complexes on a évidemment $a_i = b_i = c_i = 0$.

D'après ce que nous avons dit ci-dessus, si le polynôme caractéristique est de la forme ci-dessus alors le polynôme minimal est nécessairement de la forme

$$m_A = \prod_{i=1}^k (\lambda_i - X)^{\tilde{\alpha}_i} \prod_{i=1}^l (X^2 + b_i X + c_i)^{\tilde{\beta}_i}$$

avec $1 \leq \tilde{\alpha}_i \leq \alpha_i$ et $1 \leq \tilde{\beta}_i \leq \beta_i$.

Exemple 30. Déterminer le polynôme minimal de

$$\begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$$

► Le polynôme caractéristique est $\chi_A = -(X - 2)^3$. Comme le polynôme minimal est un diviseur de χ_A et qu'il a les mêmes racines, le polynôme minimal est à chercher parmi

$$(X - 2)^3 \quad (X - 2)^2 \quad (X - 2)$$

Rappelons que l'on a normalisé le polynôme minimal de sorte que son coefficient dominant soit +1. On calcule

$$A - 2I = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{pmatrix} \neq 0$$

donc $X - 2$ n'est pas le polynôme minimal.

On calcule

$$(A - 2I)^2 = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{pmatrix}^2 = 0$$

Donc le polynôme minimal est $(X - 2)^2$.

2.4 Théorème de diagonalisation : critère du polynôme minimal

Nous présentons pour finir un dernier théorème de diagonalisation qui permet de déterminer si une matrice est diagonalisable sans même avoir à calculer les espaces propres. Ce théorème constitue aussi un premier pas dans la direction de la trigonalisation de Jordan.

Théorème 2.10 (► Théorème de diagonalisation : critère du polynôme minimal). Soit $A \in \mathcal{M}_N(\mathbb{K})$. A est diagonalisable dans $\mathbb{K}$ si et seulement si son polynôme minimal est scindé à racines simples dans $\mathbb{K}$.

► *Démonstration.* — ($\Rightarrow$) Soit $(\lambda_i)_{i \in \{1, \dots, p\}}$ le spectre de A et $(\varepsilon_i)_{i \in \{1, \dots, N\}}$ une base de vecteurs propres. Pour tout $i \in \{1, \dots, N\}$ et vecteur propre ε_i il existe une valeur propre λ_k associée au vecteur propre ε_i c-à-d tel que

$$(A - \lambda_k \text{id})\varepsilon_i = 0 .$$

Comme les $A - \lambda \text{id}$ commutent, pour tout $i \in \{1, \dots, N\}$,

$$\prod_{l=1}^p (A - \lambda_l \text{id})\varepsilon_i = 0 .$$

Donc

$$\prod_{l=1}^p (A - \lambda_l \text{id}) \equiv 0 .$$

Ainsi

$$Q = \prod_{l=1}^p (X - \lambda_l \text{id})$$

est un polynôme annulateur de A . Comme le polynôme minimal est un diviseur de tout polynôme annulateur, voir Lemme 2.9, le polynôme minimal m_A est un diviseur de Q . Ce qui implique que m_A est scindé à racines simples. On peut d'ailleurs noter que, comme d'après Lemme 2.9, toute valeur propre est racine du polynôme minimal, on a $m_A \equiv Q$.

— ($\Leftarrow$) D'après le corollaire du critère de dimension, Corollaire 1.22, la matrice A est diagonalisable si et seulement s'il existe $(\lambda_i)_{i \in \{1, \dots, p\}}$ telle que

$$E = \bigoplus_{i=1}^p E_{\lambda_i} .$$

De plus, le lemme 1.21 assure que les espaces propres sont en somme directe. Il faut donc juste démontrer que tout $x \in E$ peut s'écrire sous la forme $x = \sum_{i=1}^k x_i$ avec $x_i \in E_{\lambda_i}$. En utilisant la décomposition fractionnelle il existe $(\eta_i)_{i \in \{1, \dots, p\}}$ tel que l'on ait

$$\frac{1}{m_A} = \frac{1}{\prod_{i=1}^p (x - \lambda_i)} = \sum_{i=1}^p \frac{\eta_i}{x - \lambda_i}$$

On définit

$$Q_i := \eta_i \frac{m_A}{X - \lambda_i} .$$

De sorte que $(X - \lambda_i)Q_i = \eta_i m_A$ et que l'on a $\sum_{i=1}^p Q_i = 1$. On a alors la décomposition suivante

$$x = \sum_{i=1}^p Q_i(A) x$$

où pour tout $i \in \{1, \dots, p\}$, $Q_i(A) x \in E_{\lambda_i}$ puisque par définition de Q_i on a

$$(A - \lambda_i I)Q_i(A) x = \eta_i m_A(A) x = 0 .$$

Ce qui termine la démonstration. □

Exemple 31. *La matrice suivante est-elle diagonalisable ?*

$$\begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$$

▷ *La matrice n'est diagonalisable ni dans $\mathbb{R}$, ni dans $\mathbb{C}$ car son polynôme minimal est $(X - 2)^2$ qui n'est pas à racines simples.*

Pour illustrer la puissance de ce théorème énonçons deux corollaires dont la démonstration est triviale lorsque l'on fait appel à ce critère de diagonalisation. Nous allons redémontrer le corollaire 1.25 :

Corollaire 2.11 (►). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. Si A a N valeurs propres distinctes dans $\mathbb{K}$ alors A est diagonalisable dans $\mathbb{K}$.*

► *Démonstration.* Si A a N valeurs propres distinctes alors χ_A est scindée à racine simple et le polynôme minimal est aussi scindé à racines simples. Donc A est diagonalisable. □

Corollaire 2.12 (►). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. Si A possède une seule valeur propre alors A n'est diagonalisable que si A est une matrice scalaire.*

► *Démonstration.* Si A a une seule valeur propre λ_0 , A est diagonalisable si et seulement si le polynôme minimal est $X - \lambda_0$. Or $A - \lambda_0 I = 0$ si et seulement si $A = \lambda_0 I$. □

2.5 Technique de trigonalisation : méthode de Jordan dans $\mathcal{M}_3$

Si nous avons vu ci-dessus que toute matrice dont le polynôme caractéristique est scindé (ce qui est toujours le cas dans un corps algébriquement clos comme $\mathbb{C}$) est trigonalisable, il a déjà été signalé que ceci n'est pas tout à fait suffisant pour calculer la puissance d'une matrice. En effet, toute matrice triangulaire peut s'écrire comme la somme d'une matrice diagonale et d'une matrice nilpotente mais rien n'assure qu'elles commutent et donc que l'on puisse utiliser la formule du binôme de Newton.

La méthode de Dunford permet de déterminer une base dans laquelle la matrice est la somme d'une matrice diagonalisable et d'une matrice nilpotente qui commutent. On peut ainsi calculer la puissance de la matrice en utilisant cette méthode.

Nous avons choisi ici, de donner une introduction rigoureuse mais partielle à une autre méthode : la trigonalisation de Jordan. La théorie complète de la méthode de Jordan nous emmènerait au delà de ce que nous souhaitons aborder dans ce cours, mais nous allons traiter tous les cas possibles dans $\mathcal{M}_3$. Dans $\mathcal{M}_3$, nous allons distinguer plusieurs cas suivant, la forme du polynôme caractéristique et du polynôme minimal.

Cas 1 : $\chi_A = -(X - \lambda_1)^3$

Sous-cas 1 : $m_A = (X - \lambda_1)$

Alors $A = \lambda_1 I$ donc la matrice est déjà diagonale.

Sous-cas 2 : $m_A = (X - \lambda_1)^2$

La matrice A n'est pas diagonalisable. Le principe général est de commencer à construire la base "par la droite", autrement dit par le dernier vecteur de base. La troisième colonne est ainsi une sorte de cinquième colonne de la réduction d'endomorphismes.

Construction de la base : On prend $\varepsilon_3 \in \mathbb{R}^3 \setminus \ker(A - \lambda_1 I)$. Donc on peut choisir ε_3 à peu près au hasard.

On pose $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$.

Il s'avère, c'est la théorie de la trigonalisation qui nous le justifierait théoriquement mais nous nous bornerons à le constater en pratique, que même si ε_2 est dans E_{λ_1} (ce qui est en fait le cas comme nous allons le voir), il y a au moins un autre vecteur propre dans E_{λ_1} . On choisit ce dernier vecteur propre pour ε_3 . Miracle 1.

Construction de la matrice : La première étape consiste à vérifier que $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une base. On le fera à notre niveau en calculant le déterminant par exemple mais la théorie de trigonalisation de Jordan nous assurerait que c'est toujours le cas. Le seul point crucial est de s'assurer que ε_2 n'est pas nul. Ceci est assuré par le choix de ε_3 qui n'est pas dans $\ker(A - \lambda_1 I)$ ce qui assure que $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3 \neq 0$.

Dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, on a :

- Troisième colonne : la dernière colonne est donnée par les coordonnées de $A\varepsilon_3$ dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Or, la définition de $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$ est équivalente à $A\varepsilon_3 = \varepsilon_2 + \lambda_1\varepsilon_3$. La dernière colonne est donc

$$\begin{pmatrix} 0 \\ 1 \\ \lambda_1 \end{pmatrix}$$

Notez que cette troisième colonne ne dépend pas du choix de ε_3 mais juste de la définition de ε_2 .

- Deuxième colonne : On calcule $(A - \lambda_1 I)\varepsilon_2$ qui est égal par définition de ε_2 à $(A - \lambda_1 I)^2\varepsilon_3$ mais comme $m_A = (X - \lambda_1)^2$, on a $(A - \lambda_1 I)^2 = 0$ donc

$$(A - \lambda_1 I)\varepsilon_2 = \underbrace{(A - \lambda_1 I)^2}_{=m_A(A)=0} \varepsilon_3 = 0$$

Ainsi ε_2 est en fait un vecteur propre. Et ce quelque soit le choix de ε_3 ! La deuxième colonne est donc

$$\begin{pmatrix} 0 \\ \lambda_1 \\ 0 \end{pmatrix}$$

- Première colonne : comme ε_1 est un vecteur propre de E_{λ_1} on a $A\varepsilon_1 = \lambda_1\varepsilon_1$ et la première colonne est

$$\begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix}$$

Donc la représentation matricielle de l'application linéaire associée à A dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

Sous-cas 3 : $m_A = (X - \lambda_1)^3$

La matrice A n'est pas diagonalisable.

Construction de la base : On prend $\varepsilon_3 \in \mathbb{R}^3 \setminus \ker[(A - \lambda_1 I)^2]$. Donc on peut encore choisir pour ε_3 à peu près n'importe quel vecteur.

On pose $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$.

On pose $\varepsilon_1 = (A - \lambda_1 I)\varepsilon_2 = (A - \lambda_1 I)^2\varepsilon_3$.

Construction de la matrice : La famille $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une base. Remarquons que ε_2 et ε_1 ne sont pas nuls car ε_3 qui n'est pas dans $\ker[(A - \lambda_1 I)^2]$ ce qui assure que $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3 \neq 0$ et $\varepsilon_1 = (A - \lambda_1 I)^2\varepsilon_3 \neq 0$.

Dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, on a :

- Troisième colonne : la troisième colonne est donnée par les coordonnées de $A\varepsilon_3$ dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Or, la définition de $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$ est équivalente à $A\varepsilon_3 = \varepsilon_2 + \lambda_1\varepsilon_3$. La troisième colonne est donc

$$\begin{pmatrix} 0 \\ 1 \\ \lambda_1 \end{pmatrix}$$

- Deuxième colonne : la deuxième colonne est donnée par les coordonnées de $A\varepsilon_2$ dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Or, la définition de $\varepsilon_1 = (A - \lambda_1 I)\varepsilon_2$ est équivalente à $A\varepsilon_2 = \varepsilon_1 + \lambda_1\varepsilon_2$. La deuxième colonne est donc

$$\begin{pmatrix} 1 \\ \lambda_1 \\ 0 \end{pmatrix}$$

— Première colonne : On calcule $(A - \lambda_1 I)\varepsilon_1$ qui, par définition de ε_2 et ε_1 , vaut

$$(A - \lambda_1 I)\varepsilon_1 = (A - \lambda_1 I)^2\varepsilon_2 = \underbrace{(A - \lambda_1 I)^3}_{m_A(A)=0}\varepsilon_3 = 0$$

La première colonne est donc

$$\begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix}$$

Remarquons encore une fois que quelque soit le choix de ε_3 , ε_1 est en fait un vecteur propre. Donc la représentation matricielle de l'application linéaire associée à A dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

Cas 2 : $\chi_A = (X - \lambda_1)^2(X - \lambda_2)$, $(\lambda_i)_{i \in \{1,2\}}$ **tous distincts**

Sous-cas 1 : $m_A = (X - \lambda_1)(X - \lambda_2)$

Dans ce cas A est diagonalisable et on peut utiliser la méthode du chapitre 1 sur la diagonalisation.

Sous-cas 2 : $m_A = (X - \lambda_1)^2(X - \lambda_2)$

Construction de la base : La construction de la base dans ce cas est moins immédiate que dans les cas précédents. En effet dans ce cas il faut choisir

ε_3 dans $\ker(A - \lambda_1 I)^2 \setminus \ker(A - \lambda_1 I)$.

On pose ensuite $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$.

Pour ε_1 on prend un vecteur propre associé à la valeur propre λ_2 .

Construction de la matrice :

— Troisième colonne : la troisième colonne est donnée par les coordonnées de $A\varepsilon_3$ dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Or, la définition de $\varepsilon_2 = (A - \lambda_1 I)\varepsilon_3$ est équivalente à $A\varepsilon_3 = \varepsilon_2 + \lambda_1\varepsilon_3$.

La troisième colonne est donc

$$\begin{pmatrix} 0 \\ 1 \\ \lambda_1 \end{pmatrix}$$

— Deuxième colonne : Comme on a choisi ε_3 dans $\ker(A - \lambda_1 I)^2$ on a

$$(A - \lambda_1 I)\varepsilon_2 = (A - \lambda_1 I)^2\varepsilon_3 = 0$$

Donc ε_2 est automatiquement un vecteur propre et la deuxième colonne est

$$\begin{pmatrix} 0 \\ \lambda_1 \\ 0 \end{pmatrix}$$

— Première colonne : comme ε_1 est dans E_{λ_2} on a évidemment $A\varepsilon_1 = \lambda_2\varepsilon_1$ donc la première colonne est

$$\begin{pmatrix} \lambda_2 \\ 0 \\ 0 \end{pmatrix}$$

Finalement la représentation matricielle de l'application linéaire associée à A dans la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est

$$\begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

Cas 3 : $\chi_A = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$, $(\lambda_i)_{i \in \{1,2,3\}}$ **tous distincts**

On a évidemment $\chi_A = m_A$. Dans ce cas A est diagonalisable et on peut utiliser la méthode du chapitre 1 sur la diagonalisation.

Discussion

L'intérêt de cette section est d'illustrer que la description de techniques sans s'appuyer sur de la théorie est à la fois barbant comme dirait mon grand-père et compliquée à retenir comme dirait mon fils de 8 ans. Mais derrière ces jugements bien blanchesque, se cache pour celles et ceux qui aiment les mystères, un monde intrigant à explorer comme les affectionnent tant les Blanchet. Je vais essayer d'illustrer ici quelques tours de passe passe mais la théorie restera à notre niveau hors de portée. Un bon rappel pour les arrogants, et pour une fois je n'utilise pas à dessein l'écriture inclusive, qui pensaient avoir tout compris de l'algèbre linéaire.

Avant d'écrire le théorème admis de trigonalisation de Jordan, introduisons :

Définition 2.13 (Bloc de Jordan). *Soient $p \in \mathbb{N}$ et $\lambda \in \mathbb{K}$, on appelle bloc de Jordan de taille p associée à λ la matrice de $\mathcal{M}_p$ $\lambda I + N$ où N est la matrice $J_{\lambda,p}$ de terme général*

$$\begin{cases} n_{ij} = 1 \text{ si } j = i + 1, \text{ pour tout } j \in \{1, \dots, p-1\} \\ 0 \text{ sinon} \end{cases}$$

Un bloc de Jordan $J_{\lambda,p}$ a donc des λ sur la diagonale, des 1 juste au dessus de cette diagonale et des 0 partout ailleurs :

$$J_{\lambda,p} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \dots \\ \vdots & \ddots & \ddots & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$$

On a donc $J_{\lambda,p} = D + N$ où $D = \lambda I$ et N est une matrice nilpotente. Ce qu'il est important de remarquer c'est que D et N commutent et donc que l'on peut appliquer la formule du binôme de Newton pour calculer $J_{\lambda,p}$. Comme pour tout $n \in \mathbb{N}^*$, la formule $A^n = P J_{\lambda,p}^n P^{-1}$ reste vraie, on peut calculer la puissance d'une matrice si elle est semblable à un bloc de Jordan.

Définition 2.14 (Matrice de Jordan). *Soient $p \in \mathbb{N}$ et $\lambda \in \mathbb{K}$, on appelle matrice de Jordan à $\{\lambda_i\}_{i \in \{1, \dots, r\}}$ la matrice de la forme*

$$J = \begin{pmatrix} J_1(\lambda_1) & 0 & 0 & \dots & 0 \\ 0 & J_2(\lambda_2) & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \dots \\ \vdots & \ddots & \ddots & J_{r-1}(\lambda_{r-1}) & 0 \\ 0 & \dots & \dots & 0 & J_r(\lambda_r) \end{pmatrix}$$

où pour tout $i \in \{1, \dots, r\}$, $J_i(\lambda_i)$ est un bloc de Jordan associé à λ_i .

Les blocs de Jordan peuvent être de taille différente et les λ_i ne sont pas nécessairement distincts.

Exemple 2.4.

$$J = \left(\begin{array}{cc|ccc} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array} \right)$$

Nous énonçons ici le théorème qui restera dans ce cours l'unique résultat non-démontré¹ :

Théorème 2.15 (Réduite de Jordan). *Soit $A \in \mathcal{M}_N(\mathbb{K})$. Si χ_A est scindé dans $\mathbb{K}$ alors A est semblable à une matrice de Jordan.*

Remarque. *On peut même voir que*

- *Les λ_i sont les valeurs propres de A .*
- *Le nombre de blocs associés à la valeur propre λ est égal à la dimension de E_λ .*
- *La taille du plus grand bloc de Jordan associé à la valeur propre λ est égal à la multiplicité de λ dans la polynôme minimal.*
- *La somme des taille des blocs de Jordan associés à la valeur propre λ est égal à la multiplicité de λ dans la polynôme caractéristique.*

2.6 Exercices

Exercice 2.1. *Déterminer l'inverse de*

$$A_1 := \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

Exercice 2.2. *Soient*

$$A_1 := \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \quad A_2 := \begin{pmatrix} 2 & 1 & -2 \\ 2 & 3 & -4 \\ 1 & 1 & -1 \end{pmatrix} \quad A_3 := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

1. *Déterminer le polynôme minimal des matrices A_1 , A_2 et A_3 .*
2. *Déterminer l'inverse des matrices A_1 , A_2 et A_3 .*
3. *Les matrices A_1 , A_2 et A_3 sont-elles diagonalisables dans $\mathbb{R}$?*
4. *Les matrices A_1 , A_2 et A_3 sont-elles diagonalisables dans $\mathbb{C}$?*
5. *Les matrices A_1 , A_2 et A_3 sont-elles trigonalisables dans $\mathbb{R}$?*
6. *Les matrices A_1 , A_2 et A_3 sont-elles trigonalisables dans $\mathbb{C}$?*

Exercice 2.3. *Déterminer la réduite de Jordan de*

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

Exercice 2.4. *Déterminer la réduite de Jordan de*

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

Exercice 2.5. *Déterminer la réduite de Jordan de*

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

Exercice 2.6. *Soient*

$$A := \begin{pmatrix} 13 & -5 & -2 \\ -2 & 7 & -8 \\ -5 & 4 & 7 \end{pmatrix}, \quad B := \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}, \quad C := \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}, \quad E := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

1. voir http://exo7.emath.fr/cours/ch_jordan.pdf pour une démonstration complète

1. Ces matrices sont-elles diagonalisables ?
2. Ces matrices sont-elles trigonalisables dans $\mathbb{R}$?
3. Lorsqu'elles sont trigonalisables dans $\mathbb{R}$ déterminer une base dans laquelle l'application linéaire est triangulaire supérieure.

Exercice 2.7. Soit le système défini par récurrence par

$$\begin{cases} x_{n+1} = 3x_n + 2y_n - 2z_n \\ y_{n+1} = -x_n + z_n \\ z_{n+1} = x_n + y_n \end{cases}$$

et où x_0, y_0 et z_0 sont donnés. Exprimer x_n, y_n et z_n en fonction de x_0, y_0, z_0 et n .

2.7 Corrigés des exercices

Corrigé Exercice 2.1. Le polynôme caractéristique est $\chi_{A_1} = -X^2 + 4X - 3$. D'après le théorème de Cayley-Hamilton $\chi_{A_1}(A_1) = 0$ c-à-d

$$-A_1^2 + 4A_1 - 3I = 0 \Leftrightarrow -A_1^2 + 4A_1 = 3I \Leftrightarrow A_1(-A_1 + 4I) = 3I$$

Donc

$$A_1^{-1} = \frac{1}{3}(-A_1 + 4I) = \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ -2 & 4 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

Corrigé Exercice 2.2. 1. (A_1) - Comme le polynôme minimal est un diviseur de χ_{A_1} et qu'il a les mêmes racines, le polynôme minimal est à chercher parmi

$$(X - 3)(X - 1)^2 \quad (X - 3)(X - 1)$$

On calcule

$$(A_1 - 3I)(A_1 - I) \neq 0$$

Donc le polynôme minimal de A_1 est $(X - 3)(X - 1)^2$. On sait que ce dernier est annulateur par le théorème de Cayley-Hamilton.

(A_2) - Le polynôme caractéristique est $-(X - 2)(X - 1)^2$. Le polynôme minimal est à chercher parmi

$$(X - 2)(X - 1)^2 \quad (X - 2)(X - 1)$$

On calcule

$$(A_2 - 2I)(A_2 - I) = 0$$

Donc le polynôme minimal de A_2 est $(X - 2)(X - 1)$.

(A_3) - Le polynôme caractéristique est $\chi_{A_3} = X^2 + 1$. Dans $\mathbb{R}$, le polynôme minimal est donc $m_{A_3} = X^2 + 1$. Dans $\mathbb{C}$, le polynôme minimal est $m_{A_3} = (X - i)(X + i)$.

2. On pourrait utiliser le théorème de Cayley-Hamilton et la méthode du premier exercice mais ici comme on a déterminé le polynôme minimal on peut l'utiliser :

(A_1) - On a vu que $\chi_{A_1} = m_{A_1}$ donc la remarque ci-dessus n'apporte rien et on obtient l'inverse comme dans le premier exercice.

(A_2) - Comme le polynôme minimal de A_2 est $(X - 2)(X - 1)$ on a

$$(A_2 - 2I)(A_2 - I) = 0 \Leftrightarrow A_2^2 - 3A_2 + 2I = 0 \Leftrightarrow -A_2^2 + 3A_2 = 2I \Leftrightarrow A_2(-A_2 + 3I) = 2I$$

Donc

$$A_2^{-1} = \frac{1}{2}(-A_2 + 3I) = \frac{1}{2} \begin{pmatrix} 1 & -1 & 2 \\ -2 & 0 & 4 \\ -1 & -1 & 4 \end{pmatrix}$$

(A₃) - On a $\chi_{A_3} = X^2 + 1$ donc

$$A_3^2 = -I \Leftrightarrow A_3(-A_3) = I$$

Donc

$$A_3^{-1} = -A_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

3. (A₁) - La matrice A_1 n'est pas diagonalisable dans $\mathbb{R}$ car son polynôme minimal est $(X - 3)(X - 1)^2$ qui n'est pas à racines simples.
 (A₂) - Le polynôme minimal de A_2 est $(X - 2)(X - 1)$ qui est scindé à racines simples dans $\mathbb{R}$ donc la matrice A_2 est diagonalisable dans $\mathbb{R}$.
 (A₃) - Le polynôme minimal de A_3 dans $\mathbb{R}$ est $X^2 + 1$ qui n'est pas scindé dans $\mathbb{R}$ donc la matrice A_3 n'est pas diagonalisable dans $\mathbb{R}$.
4. (A₁) - La matrice A_1 n'est pas diagonalisable dans $\mathbb{C}$ car son polynôme minimal est $(X - 3)(X - 1)^2$ qui n'est pas à racines simples.
 (A₂) - La matrice A_2 est diagonalisable dans $\mathbb{R}$ donc aussi dans $\mathbb{C}$.
 (A₃) - Le polynôme minimal de A_3 dans $\mathbb{C}$ est $(X + i)(X - i)$ qui est scindé à racines simples dans $\mathbb{C}$ donc la matrice A_3 est diagonalisable dans $\mathbb{C}$.
5. (A₁) - La matrice A_1 est trigonalisable dans $\mathbb{R}$ car son polynôme caractéristique est $(X - 3)(X - 1)^2$ qui est scindé.
 (A₂) - La matrice A_2 est diagonalisable dans $\mathbb{R}$ donc aussi trigonalisable dans $\mathbb{R}$.
 (A₃) - Le polynôme minimal de A_3 dans $\mathbb{R}$ est $X^2 - 1$ qui n'est pas scindé dans $\mathbb{R}$ donc la matrice A_3 n'est pas trigonalisable dans $\mathbb{C}$.
6. Toute matrice est trigonalisable dans $\mathbb{C}$.

Corrigé Exercice 2.3. On calcule $\chi_A = -(X - 2)^3$. On calcule $m_A = (X - 2)^3$ et A n'est pas diagonalisable. On détermine $E_2 = \text{Vect}\{(0, 1, 1)\}$. Théorème 2.15 stipule que A est semblable à une matrice ayant un bloc de Jordan associé à la valeur propre 2 de taille 3×3 (car 2 est de multiplicité 3 dans χ_A) :

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

N'appliquons pas ce théorème et appliquons la technique décrite plus haut. Déterminons la base, en commençant par la droite, :

- On prend v_3 (presque) quelconque. Par exemple essayons $v_3 = (0, 0, 1)$
- On pose $v_2 = (A - 2I)v_3 = (-1, 1, 0)$. Comme $v_2 \neq 0$ ceci implique que v_3 n'était effectivement pas dans $\ker(A - 2I)$ sinon il aurait fallu changer notre choix de v_3 .
- On pose $v_1 = (A - 2I)v_2 = (0, 1, 1)$. Comme $v_1 \neq 0$ ceci implique que v_3 n'était effectivement pas dans $\ker(A - 2I)^2$ sinon il aurait fallu changer notre choix de v_3 .

Construction de la matrice : (v_1, v_2, v_3) est une base de $\mathbb{R}^3$ et dans cette base on a :

- Troisième colonne : Comme $v_2 = (A - 2I)v_3$, on a $Av_3 = v_2 + 2v_3$.
- Deuxième colonne : On a $(A - 2I)v_1 = (A - 2I)^2v_2 = (A - 2I)^3v_3 = 0$. Donc $Av_2 = v_2$.
- Première colonne : On a $(A - 2I)v_1 = (A - 2I)^2v_2 = (A - 2I)^3v_3 = 0$. Donc $Av_1 = v_1$.

Donc dans la base $\{(0, 1, 1), (-1, 1, 0), (0, 0, 1)\}$ on a effectivement, comme l'aurait énoncé le théorème 2.15,

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

Corrigé Exercice 2.4. On calcule $\chi_A = -(X - 1)^3$. On calcule $m_A = (X - 1)^2$ et A n'est pas diagonalisable. On détermine $E_1 = \text{Vect}\{(1, 0, 1), (-1, 1, 0)\}$. Théorème 2.15 stipule que A est semblable à une matrice ayant un bloc de Jordan associé à la valeur propre 1 de taille 2×2 et un

bloc de Jordan associé à la valeur propre 1 de taille 1×1 (car 1 est de multiplicité 3 dans χ_A et de multiplicité 2 dans m_A) :

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

N'appliquons pas ce théorème et appliquons la technique décrite plus haut. Déterminons la base, en commençant par la droite, :

- On prend v_3 (presque) quelconque. Par exemple essayons $v_3 = (0, 0, 1)$
- On pose $v_2 = (A - 2I)v_3 = (-2, 1, -2)$.
- On pose $v_1 = (1, 0, 1)$ qui est un autre vecteur propre dans E_1 .

Construction de la matrice : (v_1, v_2, v_3) est une base de $\mathbb{R}^3$ et dans cette base on a :

- Troisième colonne : Comme $v_2 = (A - 2I)v_3$, on a $Av_3 = v_2 + 2v_3$.
- Deuxième colonne : Comme $v_1 = (A - 2I)v_2$, on a $Av_2 = v_1 + 2v_2$.
- Première colonne : Comme $v_1 \in E_1$, on a $Av_1 = v_1$.

Donc dans la base $\{(1, 0, 1), (-2, 1, -2), (0, 0, 1)\}$ on a effectivement, comme l'aurait énoncé le théorème 2.15,

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

Corrigé Exercice 2.5. On calcule $\chi_A = -(X + 1)(X - 2)^2$. On calcule $(A + I)(A - 2I) \neq 0$ donc $m_A = (X + 1)(X - 2)^2$ et A n'est pas diagonalisable. On détermine $E_2 = \text{Vect}\{(1, 0, 1)\}$ et $E_{-1} = \text{Vect}\{(-1, 1, -1)\}$. On peut appliquer la méthode décrite plus haut mais remarquons que le Théorème 2.15 stipule que A est semblable à une matrice ayant un bloc de Jordan associé à la valeur propre -1 de taille 1×1 (car -1 est de multiplicité 1 dans χ_A) et un bloc de Jordan associé à la valeur propre 2 de taille 2 (car 2 est de multiplicité 2 dans χ_A) :

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

Déterminons la base, en commençant par la droite, :

- On détermine $\ker(A - 2I)^2 = \text{Vect}\{(1, 0, 1), (-1, 1, 0)\}$. On prend v_3 dans $\ker(A - 2I)^2 \setminus \ker(A - 2I)$. Par exemple $v_3 = (-1, 1, 0)$
- On pose $v_2 = (A - 2I)v_3 = (1, 0, 1)$.
- On pose $v_1 = (-1, 1, -1) \in \ker(A + I)$.

Construction de la matrice : (v_1, v_2, v_3) est une base de $\mathbb{R}^3$ et dans cette base on a :

- Troisième colonne : Comme $v_2 = (A - 2I)v_3$, on a $Av_3 = v_2 + 2v_3$.
- Deuxième colonne : On a $(A - 2I)v_2 = (A - 2I)^2v_3 = 0$ car $v_3 \in \ker(A - 2I)^2$. Donc $Av_2 = v_2$.
- Première colonne : Comme $v_1 \in \ker(A + I)$ on a $Av_1 = v_1$.

Donc dans la base $\{(-1, 1, -1), (1, 0, 1), (-1, 1, 0)\}$ on a effectivement, comme l'aurait énoncé le théorème 2.15,

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

Corrigé Exercice 2.6. (A) - On a $\chi_A = -(X - 9)^3$. On détermine $m_A = (X - 9)^3$. On a $A = PJP^{-1}$ avec

$$P = \begin{pmatrix} 36 & -2 & 0 \\ 36 & -8 & 0 \\ -18 & -2 & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 9 & 1 & 0 \\ 0 & 9 & 1 \\ 0 & 0 & 9 \end{pmatrix}$$

(B) - On a $\chi_B = -(X - 2)^3$. On détermine $m_B = (X - 2)^2$. On a $B = PJP^{-1}$ avec

$$P = \begin{pmatrix} 3 & -3 & 0 \\ 0 & -12 & 0 \\ 1 & -9 & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

(C) - On a $\chi_C = -(X-2)^2(X-1)$. On détermine $m_C = (X-2)^2(X-1)$. On a $C = PJP^{-1}$ avec

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

(E) - On a $\chi_E = -(X-2)(X-1)^2$. On détermine $m_E = (X-2)(X-1)^2$. On a $E = PJP^{-1}$ avec

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

Corrigé Exercice 2.7. Le système est équivalent à $X_{n+1} = AX_n$ où

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

On a vu plus haut que $A = PJP^{-1}$ avec

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

Or on a $X_n = A^n X_0$ où on peut calculer A^n en utilisant $A^n = PJ^n P^{-1}$.

Deuxième partie

Algèbre bilinéaire

Chapitre 3

Formes bilinéaires et formes quadratiques

Objectif pratique du chapitre :

- Déterminer une base préduale d'une base du dual d'un espace vectoriel,
- Montrer qu'une forme est une forme bilinéaire,
- Montrer qu'une forme est une forme quadratique,
- Déterminer la représentation matricielle d'une forme bilinéaire dans une base,
- Déterminer la représentation matricielle d'une forme quadratique dans une base,
- Déterminer le signe d'une forme bilinéaire/quadratique,
- Déterminer la forme normalisée d'une forme quadratique.

Les formes quadratiques sont des polynômes quadratiques homogènes en n variables.

Exemple 32.

$$\begin{aligned} \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\mapsto ax^2 + by^2 + cz^2 + dxy + exz + fyz \end{aligned}$$

où (a, b, c, d, e, f) sont des scalaires donnés.

La théorie des formes quadratiques et les méthodes utilisées pour les étudier dépendent dans une large mesure de la nature des coefficients, qui peuvent être des nombres réels ou complexes, des nombres rationnels ou des entiers. Dans la théorie algébrique des formes quadratiques, les coefficients sont des éléments d'un certain champ. Dans la théorie arithmétique des formes quadratiques, les coefficients appartiennent à un anneau commutatif fixe, souvent les entiers $\mathbb{Z}$ ou les entiers p -adiques. Les formes quadratiques binaires ont été largement étudiées en théorie des nombres, en particulier dans la théorie des champs quadratiques, des fractions continues et des formes modulaires. La théorie des formes quadratiques intégrales en n variables a des applications importantes en topologie algébrique. Dans ce chapitre, comme c'est le cas dans la plupart des applications en algèbre linéaire et en géométrie analytique, les coefficients sont des nombres réels ou complexes.

Exemple 33. Vous avez déjà rencontré une forme quadratique qui jouera un rôle central dans le cours qui va suivre : la forme $(x, y, z) \mapsto x^2 + y^2 + z^2$ sur $\mathbb{R}^3$, qui définit la structure euclidienne et dont la racine carrée permet de calculer la norme d'un vecteur. Nous étendrons ces concepts à toutes dimensions et à d'autres espaces vectoriels.

Les formes quadratiques interviennent dans de nombreux domaines des mathématiques : différents résultats de classification des coniques et plus généralement des quadriques, introduction de la courbure des surfaces, analyse en composantes principales en statistiques. Les formes quadratiques entières interviennent en théorie des nombres et en topologie algébrique. On trouve également des formes quadratiques dans plusieurs domaines de la physique : pour définir l'ellipsoïde d'inertie en mécanique du solide, en relativité restreinte ou générale, etc.

Exemple 34. La forme quadratique $(x, y, z, t) \mapsto x^2 + y^2 + z^2 - t^2$ sur $\mathbb{R}^4$ permet de définir l'espace de Minkowski utilisé en relativité restreinte.

Dans le cursus suivi à TSE les formes quadratiques seront utilisées dans la recherche de minima ou maxima locaux d'une fonction de plusieurs variables à partir d'un développement limité. Ils permettent notamment d'étendre les conditions d'ordre deux à des fonctions de plusieurs variables.

3.1 Dual

Dans cette section, nous allons donner des noms nouveaux à des concepts que vous étiez déjà habitués à utiliser en première année.

Définition 3.1 (► Espace dual). Soit E un espace vectoriel. On appelle dual de E , noté E^* , l'ensemble des formes linéaires sur E : $E^* := \mathcal{L}(E, \mathbb{K})$

Exemple 35.

$$\begin{aligned} \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\mapsto x - y + 2z \end{aligned}$$

Proposition 3.2 (► E^* est un espace vectoriel). Soit E un espace vectoriel. E^* est un espace vectoriel.

► *Démonstration.* E^* est un sous-espace vectoriel de l'ensemble des applications de E à valeurs dans $\mathbb{R}$ donc c'est un espace vectoriel. $\square$

Exemple 36. *Considérons*

$$\begin{aligned} e_1^* : \quad \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\mapsto x \\ e_2^* : \quad \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\mapsto y \\ e_3^* : \quad \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\mapsto z \end{aligned}$$

Montrer que $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une base de $(\mathbb{R}^3)^*$.

► Montrons que $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une famille libre de $(\mathbb{R}^3)^*$: Soit $\{\alpha_i\}_{i \in \{1,2,3\}}$ tel que

$$\sum_{i=1}^3 \alpha_i e_i^* = 0.$$

Pour tout $j \in \{1, 2, 3\}$ on a

$$0 = \sum_{i=1}^3 \alpha_i e_i^*(e_j) = \alpha_j.$$

Donc $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une famille libre de $(\mathbb{R}^3)^*$.

Montrons maintenant que $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une famille génératrice de $(\mathbb{R}^3)^*$: Tout élément de $(\mathbb{R}^3)^*$ s'écrit $f : (x, y, z) \mapsto ax + by + cz$ où $(a, b, c) \in \mathbb{K}$. Donc pour tout (x, y, z) dans $\mathbb{R}^3$, $f(x, y, z) = a e_1^*(x, y, z) + b e_2^*(x, y, z) + c e_3^*(x, y, z)$. D'où $f = a e_1^* + b e_2^* + c e_3^*$. Finalement $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une famille génératrice de $(\mathbb{R}^3)^*$ et $\{e_i^*\}_{i \in \{1,2,3\}}$ forme une base de $(\mathbb{R}^3)^*$.

Définition 3.3 (► Base canonique de $(\mathbb{R}^N)^*$). Pour tout $i \in \{1, \dots, N\}$ on définit

$$\begin{aligned} e_i^* : \quad \mathbb{R}^N &\rightarrow \mathbb{R} \\ (x_i)_{i \in \{1, \dots, N\}} &\mapsto x_i \end{aligned}$$

La famille $\{e_i^*\}_{i \in \{1, \dots, N\}}$ est appelée base canonique de l'espace $(\mathbb{R}^N)^*$.

Définition 3.4 (► Base duale). Soit E un espace vectoriel de dimension N finie. Considérons $\mathcal{E} := \{e_i\}_{i \in \{1, \dots, N\}}$ une base de E . La famille $\{f_j\}_{j \in \{1, \dots, N\}}$ définie par

$$f_j(e_i) = \delta_{i,j} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

est appelée base duale de $\mathcal{E}$.

Proposition 3.5 (► Dimension de E^*). Soit E un espace vectoriel de dimension finie et $\mathcal{E}$ une base de E . La base duale de $\mathcal{E}$ est une base de E^* . En particulier $\dim(E) = \dim(E^*)$.

► *Démonstration.* Soient $\mathcal{E} := \{e_i\}_{i \in \{1, \dots, N\}}$ une base de E et $\mathcal{F} := \{f_j\}_{j \in \{1, \dots, N\}}$ la base duale de $\mathcal{E}$. Soit $(\alpha_j)_{j \in \{1, \dots, N\}} \in \mathbb{K}^N$ tel que

$$\sum_{j=1}^N \alpha_j f_j = 0.$$

Pour tout $i \in \{1, \dots, N\}$ on a

$$0 = \sum_{j=1}^N \alpha_j f_j(e_i) = \alpha_i.$$

Donc la famille $\mathcal{F}$ est libre.

Tout élément de E^* est de la forme

$$\begin{aligned} f &: E \rightarrow \mathbb{R} \\ x = (x_i)_{i \in \{1, \dots, N\}} &\mapsto \sum_{i=1}^N a_i x_i \end{aligned}$$

où $(a_i)_{i \in \{1, \dots, N\}}$ sont des scalaires et $(x_i)_{i \in \{1, \dots, N\}}$ sont les coordonnées de x dans la base $\mathcal{E}$. Par définition de f_j on a pour tout $j \in \{1, \dots, N\}$,

$$f_j(x) = f_j\left(\sum_{i=1}^N x_i e_i\right) = x_j.$$

Donc toute forme sur E^* s'écrit $\sum_{i=1}^N a_i f_i$. Et la famille $\mathcal{F}$ est génératrice de E^* donc c'est une base de E^* . $\square$

En particulier, la base canonique de l'espace $(\mathbb{R}^N)^*$ est une base de $(\mathbb{R}^N)^*$ appelée *base canonique duale*. En conséquence $\dim(\mathbb{R}^N)^* = N$.

Exemple 37. Déterminer la base duale de $\{(1, 1), (0, 1)\}$.

► Soit pour $i \in \{1, 2\}$,

$$\begin{aligned} f_i &: \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) &\mapsto \alpha_i x + \beta_i y \end{aligned}$$

où $(\alpha_i, \beta_i)_{i \in \{1, 2\}}$ sont des réels à déterminer. On résout pour $e_1 = (1, 1)$ et $e_2 = (0, 1)$

$$\begin{cases} f_1(e_1) = 1 \\ f_1(e_2) = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 + \beta_1 = 1 \\ \beta_1 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = 1 \\ \beta_1 = 0 \end{cases}$$

De même

$$\begin{cases} f_2(e_1) = 0 \\ f_2(e_2) = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 + \beta_2 = 0 \\ \beta_2 = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 = -1 \\ \beta_2 = 1 \end{cases}$$

Donc la base duale est $\{f_1, f_2\}$ où

$$\begin{aligned} f_1 &: \mathbb{R}^2 \rightarrow \mathbb{R} & \text{et} & & f_2 &: \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) &\mapsto x & & & (x, y) &\mapsto -x + y \end{aligned}$$

Définition 3.6 (► Base pré-duale). Soit E un espace vectoriel de dimension finie. Considérons $\mathcal{E}^* := \{f_j\}_{j \in \{1, \dots, N\}}$ une base de E^* . La famille $\{e_i\}_{i \in \{1, \dots, N\}}$ définie par

$$f_j(e_i) = \delta_{i,j} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

est appelée base pré-duale de $\mathcal{E}^*$.

Proposition 3.7 (► Base pré-duale est une base de E^*). Soit E un espace vectoriel de dimension finie et E^* son dual. Soit $\mathcal{E}^*$ une base de E^* . La base pré-duale de $\mathcal{E}^*$ est une base de E .

► *Démonstration.* Soit N la dimension de E et de E^* . Soit $(e_j)_{\{1, \dots, N\}}$ la base pré-duale de $\mathcal{E}^* := \{f_i\}_{\{1, \dots, N\}}$. La famille $(e_j)_{\{1, \dots, N\}}$ contient N éléments et $\dim E = N$ donc il suffit de montrer que la famille $(e_j)_{\{1, \dots, N\}}$ est une famille libre. Soit $(\alpha_j)_{j \in \{1, \dots, N\}} \in \mathbb{K}^N$ tel que

$$\sum_{j=1}^N \alpha_j e_j = 0.$$

En composant par f_i pour $i \in \{1, \dots, N\}$ donné on a

$$0 = f_i \left(\sum_{j=1}^N \alpha_j e_j \right).$$

Donc pour $i \in \{1, \dots, N\}$ donné on a

$$0 = \sum_{j=1}^N \alpha_j f_i(e_j) = \alpha_i.$$

Donc la famille $(e_j)_{\{1, \dots, N\}}$ est libre et c'est une base de l'espace E . □

Remarque (►). En fait on peut prouver que la base duale de la base pré-duale d'une base est la base elle-même. Et de même, la base pré-duale de la base duale d'une base est la base elle-même.

Exemple 38. Soit $\{f_1, f_2\}$ où

$$f_1 : \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & x \end{array} \quad \text{et} \quad f_2 : \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & -x + y \end{array}$$

Déterminer la base pré-duale de $\{f_1, f_2\}$.

► Soient $e_1 = (\alpha_1, \beta_1)$ et $e_2 = (\alpha_2, \beta_2)$ des vecteurs de $\mathbb{R}^2$. On a

$$\begin{cases} f_1(e_1) = 1 \\ f_2(e_1) = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = 1 \\ -\alpha_1 + \beta_1 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = 1 \\ \beta_1 = 1 \end{cases}$$

De même

$$\begin{cases} f_1(e_2) = 0 \\ f_2(e_2) = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 = 0 \\ -\alpha_2 + \beta_2 = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 = 0 \\ \beta_2 = 1 \end{cases}$$

Donc la base pré-duale de (f_1, f_2) est $\{(1, 1), (0, 1)\}$.

Remarque (►). L'étudiant.e opiniâtre pourra voir que ces deux propositions impliquent qu'il y a en fait un isomorphisme entre l'espace E et son dual E^* . Malheureusement cet isomorphisme dépend de la base donc ce n'est pas un isomorphisme canonique. On peut prouver par contre qu'il existe un isomorphisme canonique entre E et $(E^*)^*$. Il faudrait alors parler du bi-dual mais cela nous emmènerait sur une chemin de traverse et que nous avons déjà beaucoup de chemin à parcourir.

L'étudiant.e aventureu.se.x pourra utiliser l'identification d'un élément de $(\mathbb{R}^N)^*$ à un vecteur de $\mathbb{R}^N$ pour pouvoir utiliser les déterminants et la méthode de réduction de Gauss mais cela nécessite d'avoir une bonne maîtrise de cet isomorphisme et nous n'encouragerons pas un.e étudiant.e qui se surestime à l'utiliser. Une identification hasardeuse d'un.e étudiant.e pourrait se conclure par un échec et un cinglant classement dans le groupe des étudiants qui se croient meilleurs qu'ils ne le sont vraiment. Défi. C'est ce que nous allons faire dans la suite de cette section.

En fait, on pourra associer à f_1 (respectivement f_2) de l'exercice 38 ci-dessus la représentation matricielle de la famille $\{f_1, f_2\}$ dans la base canonique de $(\mathbb{R}^2)^*$:

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

La liberté de $\{f_1, f_2\}$ découle du caractère non-nul du déterminant de A . On peut voir, si l'idée nous prend, que

$$(A^{-1})^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

a pour colonne les coordonnées de la base pré-duale! C'est en fait un résultat qui est toujours vrai mais qui restera pour le moment un nouveau miracle extraordinaire qu'il nous faut mettre sur la liste des tour de magie à comprendre... Pour ce faire introduisons

Définition 3.8 (► Crochet de dualité). *Soit E un espace vectoriel et E^* son espace dual. On définit le crochet de dualité (ou appariement dual canonique) la forme bilinéaire*

$$\begin{aligned} \langle \cdot, \cdot \rangle_{E^* \times E} & \quad E^* \times E & \rightarrow & \quad \mathbb{R} \\ (\phi, x) & \mapsto & \phi(x) \end{aligned}$$

Il s'agit encore d'une notion que vous utilisez avec aisance et dextérité comme autant de M. Jourdain sans le savoir. Mais ce formalisme permet de définir plus proprement une notion qui vous est aussi tout à fait familière, ce formalisme nous permettra de clarifier certains points par la suite :

Définition 3.9 (► Application transposée). *Soient E et F deux espaces vectoriels et $u \in \mathcal{L}(E, F)$. L'application transposée de u notée u^T est l'application de F^* dans E^* définie pour tout $\eta \in F^*$ et tout $x \in E$ par*

$$\langle u^T(\eta), x \rangle_{E^* \times E} = \langle \eta, u(x) \rangle_{F^* \times F} .$$

Cette définition ne se réduit pas à la dimension finie mais la proposition suivante devrait rassurer celles et ceux qui s'inquièterait d'une définition qui manquerait de cohérence avec ce qui a été vu en L1 :

Proposition 3.10 (► Transposée en dimension finie). *Soit E et F deux espaces de dimension finie. On considère $\mathcal{E}$ une base de E et $\mathcal{F}$ une base de F . On a*

$$\text{mat}(u^T, \mathcal{F}^*, \mathcal{E}) = \text{mat}(u, \mathcal{E}^*, \mathcal{F})^T .$$

► *Démonstration.* Soit $\mathcal{E} := (e_i)_{i \in \{1, \dots, N\}}$ une base de E et $\mathcal{F} := (f_i)_{i \in \{1, \dots, p\}}$ une base de F . On note $\mathcal{E}^* := (e_i^*)_{i \in \{1, \dots, N\}}$ la base duale de $\mathcal{E}$ (qui est une base de E^* d'après Proposition 3.5) et $(f_i^*)_{i \in \{1, \dots, p\}}$ la base duale de $\mathcal{F}$ (qui est une base de F^* d'après Proposition 3.5).

Commençons par rappeler que $f_i^*(x)$ est la i -ème coordonnée de x , en effet :

$$f_i^*(x) = f_i^* \left(\sum x_j e_j \right) = \sum x_j f_i^*(e_j) = x_i .$$

Or, par définition de la représentation matricielle, le terme général d'indice i, j de la matrice $\text{mat}(u, \mathcal{E}, \mathcal{F}^*)$ est la i -ème coordonnée de $u(e_j)$. Donc le terme général d'indice i, j de la matrice $\text{mat}(u, \mathcal{E}, \mathcal{F}^*)$ est donné par

$$f_i^*(u(e_j)) = \langle f_i^*, u(e_j) \rangle_{F^* \times F} .$$

Ce qui implique, par définition de la transposée, que le terme général d'indice i, j de la matrice $\text{mat}(u, \mathcal{E}, \mathcal{F}^*)^T$ est $\langle f_j^*, u(e_i) \rangle_{F^* \times F}$.

De même le terme général d'indice i, j de la matrice $\text{mat}(u^T, \mathcal{F}^*, \mathcal{E})$ est la i -ème coordonnée de la j -ème colonne de $\text{mat}(u^T, \mathcal{F}^*, \mathcal{E})$. Or celle-ci est donnée par les coordonnées de $u^T(f_j^*)$ dans la base $(e_i)_{i \in \{1, \dots, p\}}$. Donc est $\langle u^T(f_j^*), e_i \rangle_{E^* \times E}$.

Définition 3.9, assure que $\langle u^T(f_j^*), e_i \rangle_{E^* \times E} = \langle f_j^*, u(e_i) \rangle_{F^* \times F}$ ce qui montre le résultat annoncé. $\square$

Finalement on obtient

Proposition 3.11 (► $(P^T)^{-1}$). *Soit E un espace vectoriel de dimension finie et $\mathcal{F}^*$ une base de E^* . Soit P la matrice de passage de la base canonique de E^* à la base $\mathcal{F}^*$. Alors $(P^T)^{-1}$ est la matrice de passage de la base canonique de E à la base pré-duale de $\mathcal{F}^*$.*

Autrement dit les colonnes de $(P^T)^{-1}$ représentent les vecteurs de la base pré-duale.

► *Démonstration.* P la matrice de passage de la base canonique de E^* à la base $\mathcal{F}^*$ donc $P = \mathcal{M}_{\mathcal{C}^*}^{\mathcal{F}^*}(\text{id}) = \text{mat}(\text{id}, \mathcal{F}^*, \mathcal{C}^*)$ où $\mathcal{C}$ est la base canonique de E . Donc d'après Définition 3.9, $P^T = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\text{id}) = \text{mat}(\text{id}, \mathcal{C}, \mathcal{F})$. Donc $(P^T)^{-1} = \text{mat}(\text{id}, \mathcal{F}, \mathcal{C}) = \mathcal{M}_{\mathcal{C}}^{\mathcal{F}}(\text{id})$. $\square$

3.2 Forme bilinéaire

Définition 3.12 (► Forme bilinéaire). *Soit E un espace vectoriel. On considère une application $b : E \times E \rightarrow \mathbb{K}$. b est une application bilinéaire si*

— Pour tout $y \in E$

$$\begin{aligned} b(\cdot, y) : E &\rightarrow \mathbb{K} \\ x &\mapsto b(x, y) \end{aligned}$$

est une application linéaire sur E .

— Pour tout $x \in E$

$$\begin{aligned} b(x, \cdot) : E &\rightarrow \mathbb{K} \\ y &\mapsto b(x, y) \end{aligned}$$

est une application linéaire sur E .

On note $\mathcal{B}(E)$ l'ensemble des formes bilinéaires sur E .

Remarque (►). *On pourrait aussi définir les applications bilinéaires : soit trois espaces vectoriels E_1 et E_2 et F , on définit l'application $b : E_1 \times E_2 \rightarrow F$ qui vérifient les propriétés ci-dessus.*

On pourrait aussi définir une forme multi-linéaire $\phi : E^n \rightarrow \mathbb{R}$ qui est linéaire par rapport à chacune de ses variables. C'est d'ailleurs la définition la plus propre du déterminant : une forme multi-linéaire alternée (= dont le signe change lorsque l'on intervertit deux variables).

Exemple 39. *Montrer que*

$$b : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \mapsto 2x_1y_1 + x_2y_2 + 3x_1y_2 - x_1y_3 + 5x_2y_1 + 2x_2y_3 + x_3y_1 - 6x_3y_2 - 3x_3y_3$$

est une application bilinéaire.

► y étant fixé l'application

$$\begin{aligned} \mathbb{R}^3 &\rightarrow \mathbb{K} \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\mapsto x_1(2y_1 + 3y_2 - y_3) + x_2(5y_1 + y_2 + 2y_3) + x_3(y_1 - 6y_2 - 3y_3) \end{aligned}$$

est une forme linéaire puisqu'elle est de la forme $(x_1, x_2, x_3) \mapsto \alpha x_1 + \beta x_2 + \gamma x_3$ où α, β et γ sont donnés.

De même

$$\begin{aligned} \mathbb{R}^3 &\rightarrow \mathbb{K} \\ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &\mapsto y_1(2x_1 + 5x_2 + x_3) + y_2(3x_1 + x_2 - 6x_3) + y_3(-x_1 + 2x_2 - 3y_3) \end{aligned}$$

est une forme linéaire puisqu'elle est de la forme $(y_1, y_2, y_3) \mapsto \alpha y_1 + \beta y_2 + \gamma y_3$ où α , β et γ sont donnés.

Donc b est une forme bilinéaire sur $\mathbb{R}^3$.

Définition 3.13 (► Forme bilinéaire symétrique). Soit E un espace vectoriel et $b \in \mathcal{B}(E)$. b est symétrique si

$$\forall (x, y) \in E^2, \quad b(x, y) = b(y, x).$$

Exemple 40. On considère

$$\begin{aligned} \phi : \mathcal{C}([0, 1])^2 &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \int_0^1 f(x)g(x)dx \end{aligned}$$

1. Montrer que ϕ est bien définie.
2. Montrer que ϕ est une forme bilinéaire.
3. Montrer que ϕ est symétrique.

▷

1. Toute fonction continue sur le compact $[0, 1]$ est intégrable sur $[0, 1]$ donc ϕ est bien définie sur $[0, 1]$.
2. On va montrer d'abord que ϕ est symétrique car

$$\phi(f, g) = \int_0^1 f(x)g(x)dx = \int_0^1 g(x)f(x)dx = \phi(g, f).$$

Donc il suffit de montrer que $f \mapsto \int_0^1 f(x)g(x)dx$ est linéaire. Ce qui est une conséquence directe de la linéarité de l'intégrale.

► Soient E un espace vectoriel de dimension finie N et $\{e_i\}_{i \in \{1, \dots, N\}}$ une base de E . Tout vecteur x de E s'écrit $x = \sum_{i=1}^N x_i e_i$ et tout vecteur y de E s'écrit $y = \sum_{j=1}^N y_j e_j$. Soit $b \in \mathcal{B}(E)$. Par bilinéarité de b on a

$$b(x, y) = b\left(\sum_{i=1}^N x_i e_i, \sum_{j=1}^N y_j e_j\right) = \sum_{i=1}^N \sum_{j=1}^N x_i y_j b(e_i, e_j).$$

On peut noter que toute forme bilinéaire sur E s'écrit

$$\sum_{i=1}^N \sum_{j=1}^N a_{ij} x_i y_j.$$

où a_{ij} sont des réels et que b est entièrement déterminée par les données de $a_{ij} = b(e_i, e_j)$ pour tout $(i, j) \in \{1, \dots, N\}$. Ceci conduit à définir :

Définition 3.14 (► Représentation matricielle). Soient E un espace vectoriel de dimension finie et $\mathcal{E} := \{e_i\}_{i \in \{1, \dots, N\}}$ une base de E . Soit $b \in \mathcal{B}(E)$. La représentation matricielle de b dans la base $\mathcal{E}$ est la matrice de terme général $b(e_i, e_j)$.

Exemple 41. Déterminer la représentation matricielle dans la base canonique de $\mathbb{R}^3$ de la forme bilinéaire b de l'exercice 39.

▷ On a

$$\begin{pmatrix} 2 & 3 & -1 \\ 5 & 1 & 2 \\ 1 & -6 & -3 \end{pmatrix}$$

► Réciproquement à toute matrice $A \in \mathcal{M}_N(\mathbb{K})$ on peut associer la forme bilinéaire

$$\begin{aligned} \mathcal{M}_N(\mathbb{K})^2 &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto X^T A Y \end{aligned}$$

où X^T est la transposée de X .

Exemple 42. Déterminer l'application bilinéaire associée à la matrice

$$\begin{pmatrix} 2 & 3 & -1 \\ 5 & 1 & 2 \\ 1 & -6 & -3 \end{pmatrix}$$

▷ On calcule

$$\begin{aligned} (x_1 \ x_2 \ x_3) \begin{pmatrix} 2 & 3 & -1 \\ 5 & 1 & 2 \\ 1 & -6 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ = 2x_1y_1 + x_2y_2 + 3x_1y_2 - x_1y_3 + 5x_2y_1 + 2x_2y_3 + x_3y_1 - 6x_3y_2 - 3x_3y_3. \end{aligned}$$

Proposition 3.15 (► Représentation matricielle d'une forme bilinéaire symétrique). Soient E un espace vectoriel de dimension finie et $b \in \mathcal{B}(E)$. Si b est symétrique alors sa représentation matricielle est aussi symétrique.

► *Démonstration.* Soit $\mathcal{E} = \{e_i\}_{i \in \{1, \dots, N\}}$ une base de E . La représentation matricielle de b est la matrice de terme général $a_{ij} = b(e_i, e_j)$. Donc $a_{ij} = b(e_i, e_j) = b(e_j, e_i) = a_{ji}$ si et seulement si b est symétrique. $\square$

Remarque (►). Comme pour le cas des applications linéaires, la base étant donnée, il y a un isomorphisme entre forme bilinéaire et matrice. Si on pousse le raisonnement plus loin on peut donc construire un isomorphisme entre les applications linéaires et les formes bilinéaires en passant par (le transposée de) sa matrice. Cette identification est très fructueuse en théorie mathématiques de l'analyse fonctionnelle. Il n'est pas indispensable à ce niveau de parler davantage de ce théorème de représentation mais nous l'explorerons davantage dans le chapitre 5, voir Théorème 5.3.

La représentation matricielle d'une forme bilinéaire dépend de la base considérée et on l'analogie de la formule de changement de base pour les formes bilinéaires :

Proposition 3.16 (► Formule de changement de base pour une forme bilinéaire). Soit E un espace vectoriel de dimension finie et $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E . Soit $b \in \mathcal{B}(E)$. On a

$$\text{mat}(b, \mathcal{E}) = (P_{\mathcal{E}'}^{\mathcal{E}})^T \text{mat}(b, \mathcal{E}') P_{\mathcal{E}'}^{\mathcal{E}}.$$

► *Démonstration.* Soit $(x, y) \in E^2$. On note $X = \mathcal{M}_{\mathcal{E}}(x)$, $X' = \mathcal{M}_{\mathcal{E}'}(x)$, $Y = \mathcal{M}_{\mathcal{E}}(y)$ et $Y' = \mathcal{M}_{\mathcal{E}'}(y)$. On a d'une part

$$b(x, y) = X^T \text{mat}(b, \mathcal{E}) Y$$

et d'autre part

$$b(x, y) = X'^T \text{mat}(b, \mathcal{E}') Y'. \quad (3.1)$$

D'après la formule de changement de base pour les vecteurs Lemme 1.7 on a $X' = P_{\mathcal{E}'}^{\mathcal{E}} X$ et $Y' = P_{\mathcal{E}'}^{\mathcal{E}} Y$. Donc en remplaçant dans (3.1) on obtient

$$X^T \text{mat}(b, \mathcal{E}) Y = b(x, y) = \underbrace{(P_{\mathcal{E}'}^{\mathcal{E}} X)^T}_{=X'^T} \text{mat}(b, \mathcal{E}') \underbrace{P_{\mathcal{E}'}^{\mathcal{E}} Y}_{=Y'} = X'^T (P_{\mathcal{E}'}^{\mathcal{E}})^T \text{mat}(b, \mathcal{E}') P_{\mathcal{E}'}^{\mathcal{E}} Y$$

ce qui est le résultat annoncé. $\square$

Remarque. Cette formule est à distinguer de la formule de changement de base pour les endomorphismes (1.1) qui peut s'écrire :

$$\mathcal{M}_{\mathcal{E}}(b) = (P_{\mathcal{E}'}^{\mathcal{E}})^{-1} \mathcal{M}_{\mathcal{E}'}(b) P_{\mathcal{E}'}^{\mathcal{E}} .$$

Dans de rares cas, mais fréquents en pratiques, de matrices symétriques ces deux formules coïncident mais c'est assez incroyable. Rappelons que ces deux mondes des applications linéaires et des formes bilinéaires non a priori rien à voir. Cette connection magique fait l'objet du fabuleux Théorème spectral sur lequel nous reviendrons dans la chapitre 5.

3.3 Formes quadratiques

Définition 3.17 (► Forme quadratique). Soit E un espace vectoriel. Une forme quadratique sur $\mathbb{K}$ est une application $q : E \rightarrow \mathbb{K}$ telle que

- pour tout $\lambda \in \mathbb{K}$ et tout $v \in E$, $q(\lambda v) = \lambda^2 q(v)$,
- la fonction $\phi : (u, v) \mapsto [q(u + v) - q(u) - q(v)]/2$ est bilinéaire.

On note $\mathcal{Q}(E)$ l'ensemble des formes quadratiques sur E . Lorsque $q \in \mathcal{Q}(E)$, on appelle ϕ forme polaire de q .

Exemple 43. Considérons

$$q : \begin{array}{l} \mathbb{R}^3 \rightarrow \mathbb{K} \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto 2x_1^2 + x_2^2 + 8x_1x_2 - 4x_3x_2 - 3x_3^2 \end{array}$$

Montrer que q est une forme quadratique.

▷ Déterminons la forme polaire associée à q : pour tout $u = (x_1, x_2, x_3)$ et $v = (y_1, y_2, y_3)$ dans $\mathbb{R}^3$ on a

$$\begin{aligned} q(u+v) - q(u) - q(v) &= 2(x_1+y_1)^2 + (x_2+y_2)^2 + 8(x_1+y_1)(x_2+y_2) - 4(x_3+y_3)(x_2+y_2) - 3(x_3+y_3)^2 \\ &\quad - (2x_1^2 + x_2^2 + 8x_1x_2 - 4x_3x_2 - 3x_3^2) - (2y_1^2 + y_2^2 + 8y_1y_2 - 4y_3y_2 - 3y_3^2) \end{aligned}$$

En développant on obtient

$$b(u, v) := \frac{1}{2} [q(u + v) - q(u) - q(v)] = 2x_1y_1 + x_2y_2 + 4x_1y_2 + 4x_2y_1 - 2x_3y_2 - 2x_2y_3 - 3x_3y_3 .$$

La forme b est bilinéaire car elle est de la forme $\sum a_{ij}x_iy_j$. De plus $b(u, u) = q(u)$ donc pour tout $\lambda \in \mathbb{K}$ et tout $v \in E$, $q(\lambda v) = b(\lambda v, \lambda v) = \lambda^2 b(v, v) = \lambda^2 q(v)$.

Exemple 44. Considérons

$$\psi : \begin{array}{l} \mathcal{C}([0, 1]) \rightarrow \mathbb{R} \\ f \mapsto \int_0^1 f(x)^2 dx \end{array}$$

Montrer que ψ est une forme quadratique.

▷ On a pour tout $f \in \mathcal{C}([0, 1])$ et tout $\lambda \in \mathbb{R}$,

$$\psi(\lambda f) = \int_0^1 [\lambda f(x)]^2 dx = \lambda^2 \int_0^1 f(x)^2 dx = \lambda^2 \psi(f) .$$

Déterminons la forme polaire associée à ψ :

$$\psi(f + g) - \psi(f) - \psi(g) = \int_0^1 [f(x) + g(x)]^2 dx - \int_0^1 f(x)^2 dx - \int_0^1 g(x)^2 dx = 2 \int_0^1 f(x) g(x) dx .$$

Donc la forme polaire associée est

$$b : \mathcal{C}([0, 1])^2 \rightarrow \mathbb{R} \\ (f, g) \mapsto \int_0^1 f(x)g(x)dx$$

Pour g donné l'application

$$\mathcal{C}([0, 1]) \rightarrow \mathbb{R} \\ f \mapsto \int_0^1 f(x)g(x)dx$$

est linéaire par linéarité de l'intégrale. De plus b est symétrique donc b est bilinéaire et ψ est une forme quadratique.

Nous avons déjà évoqué que ce chapitre permet de donner à cadre théorique propre à la géométrie intuitive au collège. Nous y reviendrons ultérieurement dans le chapitre 4 mais nous pouvons déjà constater que l'on retrouve le célèbre

Proposition 3.18 (► Identité du parallélogramme). Soient E un espace vectoriel et $q \in \mathcal{Q}(E)$.

$$q(u + v) + q(u - v) = 2q(u) + 2q(v).$$

Si la démonstration de ce résultat faisait appel à des considérations géométriques au collège, le formalisme développé ici rend la preuve absolument limpide :

► *Démonstration.* Soit b la forme polaire de q . On a

$$\begin{aligned} q(u + v) + q(u - v) &= b(u + v, u + v) + b(u - v, u - v) \\ &= b(u, u) + b(u, v) + b(v, u) + b(v, v) + b(u, u) - b(u, v) - b(v, u) + b(v, v) \\ &= 2q(u, u) + 2q(v, v). \end{aligned}$$

□

Remarque (►). L'intérêt de cette proposition reste relativement limité jusque là mais elle prendra beaucoup plus d'importance dans le cours *Functional Analysis*, notamment parce que nous verrons, pour les quelques élu.e.s d'entre vous, qu'une norme est issue d'un produit scalaire si et seulement si l'identité du parallélogramme est vérifiée. Ce qui permet en particulier de prouver quand un espace normé n'est pas un espace de Hilbert. Par exemple on pourra prouver ainsi que les espaces l^p des suites dont la norme

$$\|(x_i)_i\|_p = \left[\sum x_i^p \right]^{1/p}$$

est bornée est un espace de Hilbert si et seulement si $p = 2$, c'est à dire si c'est la distance Euclidienne.

Définition 3.19 (► Forme quadratique associée à une forme bilinéaire). Soit E un espace vectoriel et $b \in \mathcal{B}(E)$. On appelle forme quadratique associée à b la forme

$$E \rightarrow \mathbb{R} \\ x \mapsto b(x, x)$$

Pour justifier la terminologie il nous faut prouver que c'est bien une forme quadratique :

Proposition 3.20 (► La forme quadratique associée est une forme quadratique). Soient E un espace vectoriel et $b \in \mathcal{B}(E)$. La forme quadratique q associée à b est une forme quadratique.

► *Démonstration.* Par bilinéarité de b on a évidemment

$$q(\lambda x) = b(\lambda x, \lambda x) = \lambda^2 b(x, x) = \lambda^2 q(x) .$$

D'autre part

$$\begin{aligned} 2[q(u+v) - q(u) - q(v)] &= b(u+v, u+v) - b(u, u) - b(v, v) \\ &= b(u, u) + b(u, v) + b(v, u) - b(u, u) - b(v, v) \\ &= b(u, v) + b(v, u) , \end{aligned}$$

qui est bilinéaire par bilinéarité de $(u, v) \mapsto b(u, v) + b(v, u)$. Donc q est bien une forme quadratique $\square$

En fait on a prouvé dans le même temps, mais il est important de le souligner, la proposition suivante :

Proposition 3.21 (► Isomorphisme entre forme quadratique et forme bilinéaire symétrique). *Soit $b \in \mathcal{B}(E)$. On considère la forme quadratique q associée à b . La forme polaire $\tilde{b}$ associée à q est égale à b si et seulement si b est symétrique.*

► *Démonstration.* On a $q(u) = b(u, u)$. La forme polaire $\tilde{b}$ est telle que

$$2\tilde{b}(u, v) = q(u+v) - q(u) - q(v) = b(u+v, u+v) - b(u, u) - b(v, v) = b(u, v) + b(v, u) .$$

Donc $\tilde{b} = b$ si et seulement si b est symétrique. $\square$

Si on compare les résultats de l'exercice 43 à l'énoncé de l'exercice 39 on se rend compte effectivement que la forme polaire d'une forme quadratique associée à une forme bilinéaire n'est pas nécessairement la forme bilinéaire elle-même. Notons par contre que si $q \in \mathcal{Q}(E)$, la forme quadratique associée à la forme polaire de q est q elle-même.

Remarque (►). *Comme la forme polaire d'une forme quadratique est une forme bilinéaire et que toute forme bilinéaire sur $\mathbb{R}^N$ est de la forme*

$$\sum_{i=1}^N \sum_{j=1}^N a_{ij} x_i y_j$$

Toute forme quadratique sur $\mathbb{R}^N$ est de la forme

$$q((x_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N \sum_{j=1}^N a_{ij} x_i x_j .$$

que l'on peut réécrire

$$q((x_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N a_{ii} x_i^2 + \sum_{(i,j)=1, i \neq j}^N a_{ij} x_i x_j$$

Comme $x_i x_j = x_j x_i$ on peut simplifier cette expression en

$$q((x_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N c_{ii} x_i^2 + \sum_{(i,j)=1, i < j}^N c_{ij} x_i x_j$$

avec $c_{ij} = a_{ij} + a_{ji}$ pour tout $i \neq j$ et $c_{ii} = a_{ii}$.

Proposition 3.22 (► Forme polaire dans $\mathbb{R}^N$). Dans $\mathbb{R}^N$, soit la forme quadratique

$$q((x_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N c_{ii} x_i^2 + \sum_{(i,j)=1, i < j}^N c_{ij} x_i x_j$$

la forme polaire associée est

$$b((x_i)_{\{1, \dots, N\}}, (y_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N c_{ii} x_i y_i + \frac{1}{2} \sum_{(i,j)=1, i < j}^N c_{ij} (x_i y_j + x_j y_i) .$$

En fait on pourra se souvenir que pour trouver la forme polaire d'une forme quadratique on peut utiliser la règle mnémotechnique x_i^2 devient $x_i y_i$ et $x_i x_j$ devient $(x_i y_j + x_j y_i)/2$.

► *Démonstration.* Dans la formule de la forme polaire chaque terme en x_i^2 devient :

$$\frac{1}{2} [(x_i + y_i)^2 - x_i^2 - y_i^2] = x_i y_i .$$

Et tout terme de la forme $x_i x_j$ avec $i \neq j$, devient

$$\frac{1}{2} [(x_i + y_i)(x_j + y_j) - x_i x_j - y_i y_j] = \frac{1}{2} [x_i y_j + x_j y_i] ; .$$

On remplace dans la formulation de q pour trouver le résultat annoncé. □

Définition 3.23 (► Représentation matricielle d'une forme quadratique). La représentation matricielle d'une forme quadratique dans une base est la représentation matricielle de sa forme polaire dans cette même base.

Remarquons que si $A \in \mathcal{M}_N$ est une matrice symétrique, la base étant donnée, la forme quadratique associée q est isomorphe à

$$X \mapsto X^T A X .$$

On peut noter ici que l'isomorphisme entre les formes quadratiques et les formes bilinéaires symétrique s'étend donc aussi aux matrices symétriques. Une matrice symétrique peut donc à la fois être identifiée à une forme quadratique et à une application linéaire. Il y a donc aussi un isomorphisme entre applications linéaires symétriques et formes quadratiques.

3.4 Signe d'une forme quadratique

On a vu qu'une des applications importantes de ce chapitre est de pouvoir étendre la théorie de l'optimisation pour des fonctions réelles à des fonctions de plusieurs variables. Lors de l'étude des fonctions d'une variable la première étape consistait à déterminer les points critiques de la fonction (en résolvant $f'(x) = 0$) puis à regarder les conditions d'ordre deux (le signe de la dérivée seconde). L'analogue de la dérivée seconde en dimensions supérieures est une matrice et l'objet de cette section est de donner un sens au "signe" d'une matrice (ou d'une forme quadratique).

Les formes bilinéaires symétriques, définies positives, appelées produits scalaires, voir Chapitre 4, jouent aussi un rôle fondamental dans l'extension de l'intuition développée dans la géométrie dans l'espace à des espaces vectoriels plus généraux, y compris, avec quelques hypothèses supplémentaires, à des espaces vectoriels de dimension infinie.

3.4.1 Définitions

Définition 3.24 (► Forme positive et négative). Soient E un espace vectoriel et $b \in \mathcal{B}(E)$.

- b est semi-définie positive si pour tout $x \in E$, $b(x, x) \geq 0$.
- b est semi-définie négative si pour tout $x \in E$, $b(x, x) \leq 0$.
- b est définie positive si pour tout $x \in E$,

$$\begin{cases} b(x, x) \geq 0 \\ b(x, x) = 0 \Leftrightarrow x = 0 \text{ (on dit que } b \text{ est non-dégénérée)} \end{cases}$$

- b est définie négative si pour tout $x \in E$,

$$\begin{cases} b(x, x) \leq 0 \\ b(x, x) = 0 \Leftrightarrow x = 0 \end{cases}$$

Ces définitions ne font intervenir que la forme quadratique associée à la forme bilinéaire. Ces définitions s'étendent naturellement aux formes quadratiques et aux matrices symétriques. On a par exemple : soit $A \in \mathcal{M}_N(\mathbb{R})$, A est définie positive si

$$X^T A X \geq 0 \quad \text{et} \quad (X^T A X = 0 \Leftrightarrow x = 0).$$

Exemple 45. *Considérons*

1.

$$q_1 : \begin{array}{l} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + x_2^2 \end{array}$$

2.

$$q_2 : \begin{array}{l} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + x_2^2 - 2x_1x_2 \end{array}$$

3.

$$q_3 : \begin{array}{l} \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + 5x_2^2 - 4x_1x_2 \end{array}$$

Déterminer si ces formes quadratiques sont définie positive.

►

1. Pour tout $(x_1, x_2) \in \mathbb{R}^2$ on a

$$q_1(x_1, x_2) \geq 0 \quad \text{et} \quad q_1(x_1, x_2) = 0 \Leftrightarrow x_1^2 + x_2^2 = 0 \Leftrightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

Donc q_1 est définie positive.

2. Pour q_2 on pourrait s'attendre naïvement à ce que le terme en $-x_1x_2$ compense les termes positifs lorsque x_1 et x_2 sont grands. Ça n'est évidemment pas le cas puisque $q_2(x_1, x_2) = (x_1 + x_2)^2 \geq 0$. Par contre

$$q_2(x_1, x_2) = 0 \Leftrightarrow (x_1 + x_2)^2 = 0 \Leftrightarrow x_1 + x_2 = 0$$

mais ceci n'implique pas que $x_1 = x_2 = 0$. En effet, $q_2(1, -1) = 0$ ce qui prouve que q_2 n'est pas définie positive.

3. Pour q_3 , on ne peut pas appliquer directement une identité remarquable. Par contre on peut faire apparaître des carrés. Cette méthode de réduction en carrés est due à Gauss et sera présentée en détail dans la section qui va suivre.

$$q_3(x_1, x_2) = x_1^2 + x_2^2 - 4x_1x_2 = (x_1 - 2x_2)^2 - (-2x_2)^2 + 5x_2^2 = (x_1 - 2x_2)^2 + x_2^2.$$

Comme q_3 est une somme de carrés elle est au moins semi-définie positive. Pour ce qui est de la positivité on a

$$q_3(x_1, x_2) = 0 \Leftrightarrow (x_1 - 2x_2)^2 + x_2^2 = 0 \Leftrightarrow \begin{cases} x_1 - 2x_2 = 0 \\ x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

Donc q_3 est définie positive.

3.4.2 Méthode de réduction en carrés de Gauss

Cette méthode permet toujours de déterminer la signature d'une forme quadratique. Cette méthode ouvre aussi le champs à de multiples développements. Nous en présenterons une dans la section à venir. Cette méthode est donc à maîtriser sans négociation possible.

D'après Remarque 3.3, toute forme quadratique sur $\mathbb{R}^N$ s'écrit

$$q((x_i)_{\{1, \dots, N\}}) = \sum_{i=1}^N c_{ii} x_i^2 + \sum_{(i,j)=1, i < j}^N c_{ij} x_i x_j .$$

On va considérer deux cas :

► **Cas où il y a au moins un terme x_i^2** : Sans perte de généralité, quitte à renommer les variables, on peut supposer que $c_{11} \neq 0$. L'idée est de choisir une variable qui a un terme carré, ici on va prendre x_1 et de faire apparaître le début d'un carré afin de faire disparaître cette variable du reste de la forme quadratique :

$$\begin{aligned} \sum_{i=1}^N c_{ii} x_i^2 + \sum_{(i,j)=1, i < j}^N c_{ij} x_i x_j &= c_{11} \left[\sum_{i=1}^N \frac{c_{ii}}{c_{11}} x_i^2 + \sum_{(i,j)=1, i < j}^N \frac{c_{ij}}{c_{11}} x_i x_j \right] \\ &= c_{11} \left[x_1^2 + x_1 \sum_{j=2}^N \frac{c_{1j}}{c_{11}} x_j + \sum_{i=2}^N \frac{c_{ii}}{c_{11}} x_i^2 + \sum_{(i,j)=2, i < j}^N \frac{c_{ij}}{c_{11}} x_i x_j \right] \\ &= c_{11} \left[x_1 + \frac{1}{2} \sum_{j=2}^N \frac{c_{1j}}{c_{11}} x_j \right]^2 - \frac{c_{11}}{4} \left[\sum_{j=2}^N \frac{c_{1j}}{c_{11}} x_j \right]^2 \\ &\quad + c_{11} \left[\sum_{i=2}^N \frac{c_{ii}}{c_{11}} x_i^2 + \sum_{(i,j)=2, i < j}^N \frac{c_{ij}}{c_{11}} x_i x_j \right] \end{aligned}$$

La première ligne de la dernière égalité correspond exactement à tous les termes qui contiennent x_1 . On voit que la partie

$$c_{11} \left\{ - \left[\frac{1}{2} \sum_{j=2}^N \frac{c_{1j}}{c_{11}} x_j \right]^2 + \left[\sum_{i=2}^N \frac{c_{ii}}{c_{11}} x_i^2 + \sum_{(i,j)=2, i < j}^N \frac{c_{ij}}{c_{11}} x_i x_j \right] \right\}$$

est une forme quadratique en $(x_i)_{\{2, \dots, N\}}$ (donc qui ne contient plus de terme en x_1). Donc à chaque étape un élimine une variable ce qui assure qu'en au plus N étapes la forme quadratique sera décomposée sous forme de somme de carrés.

► **Cas où il n'y a pas de terme x_i^2** : Sans perte de généralité, quitte à renommer les variables, on peut supposer que $c_{12} \neq 0$. L'idée est de choisir deux variables, ici on va prendre x_1 et x_2 , et de faire apparaître deux carrés tout en faisant disparaître ces deux variables du reste de la forme quadratique :

$$\begin{aligned} \sum_{(i,j)=1, i < j}^N c_{ij} x_i x_j &= c_{12} \left\{ x_1 x_2 + \sum_{(i,j)=2, i < j}^N \frac{c_{ij}}{c_{12}} x_i x_j \right\} \\ &= c_{12} \left\{ x_1 x_2 + x_1 \sum_{j=3}^N \frac{c_{1j}}{c_{12}} x_j + x_2 \sum_{j=3}^N \frac{c_{2j}}{c_{12}} x_j + \sum_{(i,j)=3, i < j}^N \frac{c_{ij}}{c_{12}} x_i x_j \right\} \\ &= c_{12} \left\{ \underbrace{\left(x_1 + \sum_{j=3}^N \frac{c_{1j}}{c_{12}} x_j \right)}_{=:A} \underbrace{\left(x_2 + \sum_{j=3}^N \frac{c_{2j}}{c_{12}} x_j \right)}_B + \sum_{(i,j)=3, i < j}^N \frac{c_{ij}}{c_{12}} x_i x_j \right\} \end{aligned}$$

En utilisant l'identité

$$AB = \frac{(A+B)^2}{4} - \frac{(A-B)^2}{4}$$

On obtient

$$q((x_i)_{\{1, \dots, N\}}) = c_{12} \left\{ \frac{1}{4} \left(x_1 + \sum_{j=3}^N \frac{c_{ij}}{c_{12}} x_j + x_2 + \sum_{j=3}^N \frac{c_{ij}}{c_{12}} x_j \right)^2 - \frac{1}{4} \left(x_1 + \sum_{j=3}^N \frac{c_{ij}}{c_{12}} x_j - x_2 - \sum_{j=3}^N \frac{c_{ij}}{c_{12}} x_j \right)^2 + \sum_{(i,j)=3, i < j}^N \frac{c_{ij}}{c_{12}} x_i x_j \right\}.$$

Pour la suite de l'algorithme, ou bien on a un terme carré et on utilise la technique vue dans le cas 1 ou bien il n'y a pas de terme carré et on utilise de nouveau cette technique vue dans ce cas 2.

En au plus N étapes l'algorithme s'arrête puisque la cas 1 fait apparaître un carré en faisant disparaître une variable du reste de la forme quadratique, alors que le cas 2 fait disparaître deux variables en faisant apparaître deux carrés. On a donc au plus N terme carrés.

Finalement cette technique permet d'écrire toute forme quadratique sous la forme

$$q(x) = \sum_{i=1}^p \alpha_i l_i^2(x)$$

où les α_i sont des scalaires l_i sont des formes linéaires. De plus, comme on obtient des formes linéaires échelonnées ou de la forme $\{(A, B) \mapsto A+B, (A, B) \mapsto A-B\}$, la famille $\{l_i\}_{i \in \{1, \dots, N\}}$ est libre. Il est noter qu'il est possible d'utiliser une autre méthode pour faire apparaître des carrés mais cette méthode assure la liberté de la famille des formes linéaires. Et la liberté ça compte. Le libéralisme compte aussi mais sans cœur..

Exemple 46. *Considérons*

$$\begin{aligned} f : \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x, y, z) &\rightarrow x^2 + 2y^2 + 3z^2 + 4xy - 6xz - 4yz \end{aligned}$$

Est-ce que f est définie positive ?

▷ *On peut commencer par choisir x car il est précédé du coefficient 1 :*

$$f(x, y, z) = \underbrace{(x^2 + 2x(2y - 3z))}_{\text{tous les termes en } x} + \underbrace{2y^2 + 3z^2 - 4yz}_{\text{pas de terme en } x} \quad (3.2)$$

On reconnaît le début d'un carré dans le terme comportant tous les termes en x :

$$x^2 + 2x(2y - 3z) = [x + (2y - 3z)]^2 - \underbrace{(2y - 3z)^2}_{\text{terme carré à enlever pour corriger}} = (x + 2y - 3z)^2 - 4y^2 + 12yz - 9z^2.$$

Si on réintroduit dans l'expression 3.2 on obtient

$$f(x, y, z) = (x + 2y - 3z)^2 - 4y^2 + 12yz - 9z^2 + 2y^2 + 3z^2 - 4yz = (x + 2y - 3z)^2 - \underbrace{2y^2 - 6z^2 + 8yz}_{\text{forme quadratique en } y \text{ et } z}.$$

Puis on recommence

$$\begin{aligned} f(x, y, z) &= (x + 2y - 3z)^2 - 2(y^2 + 4z^2 - 4yz) \\ &= (x + 2y - 3z)^2 - 2[(y^2 - 4yz) + 3z^2] \\ &= (x + 2y - 3z)^2 - 2[(y - 2z)^2 - 4z^2 + 3z^2] \\ &= (x + 2y - 3z)^2 - 2(y - 2z)^2 + 2z^2. \end{aligned}$$

Il y a au moins un terme qui est précédé d'un coefficient négatif, nous verrons Section 5.4.2 que cela assure que f n'est pas même semi-définie positive, pour le moment il suffit d'exhiber un vecteur en lequel f n'est pas positive. Par exemple en résolvant le système échelonné :

$$\begin{cases} x + 2y - 3z = 0 \\ y - 2z = 1 \\ z = 0 \end{cases}$$

qui a pour solution $(-2, 1, 0)$. Donc $f(-2, 1, 0) = -2$ et f n'est pas positive. On peut remarquer qu'elle n'est pas non plus semi-définie négative, parce qu'il y a un carré précédé d'un coefficient positif, en choisissant par exemple un vecteur tel que

$$\begin{cases} x + 2y - 3z = 1 \\ y - 2z = 0 \\ z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 1 \\ y = 0 \\ z = 0 \end{cases}$$

en lequel f vaut $+1$.

Exemple 47. Considérons la forme quadratique

$$\begin{aligned} f : \quad \mathbb{R}^4 &\rightarrow \mathbb{R} \\ (x, y, z, t) &\rightarrow xy + xz + 2xt + yz - 4yt \end{aligned}$$

Quelle est la signature de f ?

▷ On a

$$\begin{aligned} f(x, y, z, t) = & (x + \underbrace{z - 4t}_{\text{les termes multipliés par } y}) (y + \underbrace{z + 2t}_{\text{les termes multipliés par } x}) \\ & - \underbrace{(z - 4t)(z + 2t)}_{\text{termes de correction, sans } x \text{ ni } y} \end{aligned}$$

On utilise $AB = (A + B)^2/4 - (A - B)^2/4$ et on obtient

$$\begin{aligned} f(x, y, z, t) &= \frac{(x + z - 4t + y + z + 2t)^2}{4} - \frac{(x + z - 4t - y - z - 2t)^2}{4} - z^2 + 8t^2 + 2zt \\ &= \frac{(x + y + 2z - 2t)^2}{4} - \frac{(x - y - 6t)^2}{4} - z^2 + 8t^2 + 2zt. \end{aligned}$$

Le terme de reste, sans x ni y , comporte des carrés donc on fait apparaître, comme dans le premier cas, le début d'un carré :

$$\begin{aligned} f(x, y, z, t) &= \frac{(x + y + 2z - 2t)^2}{4} - \frac{(x - y - 6t)^2}{4} - [z^2 - 8t^2 - 2zt] \\ &= \frac{(x + y + 2z - 2t)^2}{4} - \frac{(x - y - 6t)^2}{4} - [(z - t)^2 - t^2 - 8t^2] \\ &= \frac{(x + y + 2z - 2t)^2}{4} - \frac{(x - y - 6t)^2}{4} - (z - t)^2 + 9t^2. \end{aligned}$$

Ne pas oublier que le signe $-$ dans la dernière égalité porte sur tout le terme entre parenthèses. C'est la raison pour laquelle cela facilite les calculs de commencer par les termes précédés d'un coefficient le plus simple possible.

► On pourrait s'inquiéter de la possibilité d'une réduction en carré différente pourrait donner un résultat différent. C'est effectivement une remarque perspicace. Nous verrons en fait Section 5.4.2 que le nombre de carrés précédés d'un coefficient positif et le nombre de carrés précédés d'un signe négatif après réduction en carrés de Gauss est toujours le même quelque soit l'ordre des variables choisis. C'est la loi d'inertie de Sylvester que nous prouverons rigoureusement mais

pour ce faire nous avons besoin de davantage d'outils théoriques comme la notion d'orthogonalité. Nous verrons aussi Section 5.4.2 la méthode des valeurs propres et un critère des mineurs principaux pour déterminer quand une forme quadratique est définie positive. La méthode des valeurs propres est particulièrement excitante du point de vue théorique et même si elle est beaucoup utilisée en pratique à notre niveau et en analyse numérique, son intérêt pratique est néanmoins un peu limité. Fin du teaser.

3.5 Diagonalisation de formes quadratiques

Proposition 3.25 (► Diagonalisation de formes quadratiques). *Soient E un espace vectoriel et b une forme bilinéaire symétrique sur E . Il existe une base dans laquelle la représentation matricielle de b est diagonale.*

On peut prouver ce résultat par récurrence sur la dimension mais, suivant la philosophie de ce cours, nous allons en donner une démonstration élégante, constructive et pratique.

► *Démonstration.* Nous allons d'abord construire la base et montrer que la représentation matricielle de b est diagonale dans cette base. Dans un troisième temps nous montrerons comment modifier cette base pour obtenir une matrice normalisée.

- Construction de la base : La méthode de Gauss permet de déterminer une famille $\{\alpha_i\}_{i \in \{1, \dots, k\}}$ de scalaires et une famille $\{l_i\}_{i \in \{1, \dots, k\}}$ libre de formes linéaires telles que

$$q(x) = \sum_{i=1}^k \alpha_i l_i^2(x). \quad (3.3)$$

La famille $\{l_i\}_{i \in \{1, \dots, k\}}$ étant libre le théorème de la base incomplète assure que l'on peut compléter cette famille en une base $\mathcal{E} := \{l_i\}_{i \in \{1, \dots, N\}}$ de E^* . Soit $\{\varepsilon_i\}_{i \in \{1, \dots, N\}}$ la base pré-duale de la base $\mathcal{E}$.

- Matrice diagonale : Nous allons voir que dans cette base la représentation matricielle de b est diagonale : en effet, la forme polaire associée à q est

$$b(x, y) = \sum_{i=1}^k \alpha_i l_i(x) l_i(y).$$

Or par définition de la base pré-duale, voir Définition 3.6, on a pour tout $(m, n) \in \{1, \dots, N\}^2$, $m \neq n$

$$b(\varepsilon_m, \varepsilon_n) = \sum_{i=1}^k \alpha_i l_i(\varepsilon_m) l_i(\varepsilon_n) = 0 \quad (3.4)$$

car au moins un des termes m ou n est différent de i dans chaque terme. □

Corollaire 3.26. *Soient E un espace vectoriel et b une forme bilinéaire symétrique sur E . Il existe une base dans laquelle la représentation matricielle de b est diagonale et ne comporte que des 1, des -1 et des 0. Cette matrice est appelée matrice normalisée de b .*

Démonstration. Sans perte de généralité, les termes avec des coefficients α_i positifs, disons au nombre de p peuvent être placés dans les p premières positions. Les termes possédant des coefficients α_i négatifs, au nombre de $k - p$, peuvent aussi être placés dans les positions de $p + 1$ à k .

On définit :

$$\begin{cases} e_i = \frac{1}{\sqrt{|\alpha_i|}} \varepsilon_i & \text{si } i \in \{1, \dots, k\} \\ e_i = \varepsilon_i & \text{si } i \in \{k + 1, \dots, N\} \end{cases} \quad (3.5)$$

Remarquons que, par définition de la base préduale et (3.5)

$$l_i(e_j) = l_i\left(\frac{1}{\sqrt{|\alpha_j|}}\varepsilon_j\right) = \frac{1}{\sqrt{|\alpha_j|}}l_i(\varepsilon_j) = \begin{cases} \frac{1}{\sqrt{|\alpha_i|}} & \text{si } i = j \in \{1, \dots, k\} \\ 1 & \text{si } i = j \in \{k+1, \dots, N\} \\ 0 & \text{si } i \neq j \end{cases}$$

Donc on a, comme précédemment,

$$b(e_m, e_m) = \begin{cases} \frac{\alpha_m}{|\alpha_m|} = 1 & \text{si } m \in \{1, \dots, p\} \\ \frac{\alpha_m}{|\alpha_m|} = -1 & \text{si } m \in \{p+1, \dots, k\} \\ 0 & \text{si } m \in \{k+1, \dots, p\} \end{cases}$$

Ce qui prouve que la représentation matricielle de q dans cette base est diagonale et de la forme

$$\text{diag}\left\{ \underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_{k-p}, 0, \dots, 0 \right\}.$$

Et finit de démontrer le résultat. $\square$

On a finalement prouvé que "toute forme quadratique est diagonalisable dans la base pré-duale de la complétée de la famille libre des formes linéaires obtenues par réduction en carrés de Gauss". Une phrase à absolument retenir pour faire grande sensation lors de votre prochaine soirée. On omettra à cette occasion de dire que l'on a aussi normalisé pour ne pas paraître pédant.

Exemple 48. *Considérons la forme quadratique f de l'exercice 46*

1. Déterminer une base f -orthogonale de $\mathbb{R}^3$.
2. Déterminer la matrice normalisée de f .

▷ On a prouvé que

$$f(x, y, z) = (x + 2y - 3z)^2 - 2(y - 2z)^2 - z^2.$$

1. Posons

$$\begin{array}{lll} l_1 : \mathbb{R}^3 \rightarrow \mathbb{R} & l_2 : \mathbb{R}^3 \rightarrow \mathbb{R} & l_3 : \mathbb{R}^3 \rightarrow \mathbb{R} \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x + 2y - 3z & \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto y - 2z & \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto z \end{array}$$

La méthode de réduction en carrés de Gauss assure la liberté de la famille $\{l_1, l_2, l_3\}$. On peut d'ailleurs le vérifier en calculant le déterminant de

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

qui est évidemment échelonnée et de déterminant 1. La famille est une base de $(\mathbb{R}^3)^*$. Sa base pré-duale est donnée par les colonnes de

$$(P^T)^{-1} = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

D'après Proposition 3.25, la base

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

est f -orthogonale car dans cette base la représentation matricielle de f est $\text{diag}\{1, -2, -1\}$.

2. D'après la deuxième partie de Proposition 3.25, il suffit de diviser chaque vecteur de la base par $\sqrt{|f(e_i, e_i)|}$: on obtient ainsi que dans la base

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix} \right\}$$

la représentation matricielle de f est $\text{diag}\{1, -1, -1\}$.

3.6 Exercices

Espace dual

Exercice 3.1. Déterminer une base duale de $\{(1, 1, 1), (0, 1, 1), (2, 0, 1)\}$.

Exercice 3.2. Soient $l_1 : (x, y, z) \mapsto x - y + z$, $l_2 : (x, y, z) \mapsto x + y + z$ et $l_3 : (x, y, z) \mapsto x + z$. Est-ce que $\{l_1, l_2, l_3\}$ forme une base de $(\mathbb{R}^3)^*$?

Exercice 3.3. Soient $l_1 : (x, y, z) \mapsto 2x - y + z$, $l_2 : (x, y, z) \mapsto x + y + z$ et $l_3 : (x, y, z) \mapsto x - y - z$.

1. Montrer que $\{l_1, l_2, l_3\}$ forme une base de $(\mathbb{R}^3)^*$
2. Déterminer la base préduale de $\{l_1, l_2, l_3\}$.

Exercice 3.4. Soient $l_1 : (x_1, x_2, x_3) \mapsto x_1 + x_2 - 2x_3$ et $l_2 : (x_1, x_2, x_3) \mapsto x_1 - x_2 + x_3$.

1. Montrer que $\{l_1, l_2\}$ forme une famille libre de $(\mathbb{R}^3)^*$.
2. Compléter $\{l_1, l_2\}$ pour obtenir une base $\mathcal{B}$ de $(\mathbb{R}^3)^*$.
3. Déterminer une base préduale de $\mathcal{B}$.

Formes bilinéaires et formes quadratiques

Exercice 3.5. Soit

$$f : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x_1, x_2, x_3), (y_1, y_2, y_3) \mapsto x_1y_1 + 5x_2y_2 + 6x_3y_3 - x_1y_2 - x_2y_1 + x_3y_1 + x_1y_3 - 3x_3y_2 - 3x_2y_3$$

1. Montrer que f est une forme bilinéaire.
2. Déterminer la représentation matricielle de f dans la base canonique.
3. Déterminer la représentation matricielle de f dans la base $\{(1, 1, 0), (1, 0, 1), (0, 0, 1)\}$.
4. Déterminer la forme quadratique Q associée à f .
5. Déterminer la représentation matricielle de Q dans la base canonique.
6. Déterminer la représentation matricielle de Q dans la base $\{(1, 1, 0), (1, 0, 1), (0, 0, 1)\}$.

Exercice 3.6. Est-ce que

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 - 2x_1x_2 + 2x_1$$

est une forme quadratique ?

Exercice 3.7. Montrer que

$$f : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x_1, x_2, x_3) \mapsto x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + 2x_1x_3 + 4x_2x_3$$

est une forme quadratique

Exercice 3.8. Mêmes questions que dans l'exercice 3.5 avec

$$f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2), (y_1, y_2) \mapsto x_1y_1 + 5x_2y_2 - x_1y_2 + x_2y_1$$

dans la base $\{(1, 1), (1, 0)\}$ pour les questions 3 et 6.

Exercice 3.9. *Mêmes questions que dans l'exercice 3.5 pour*

$$\begin{aligned} f : \mathbb{R}^2 \times \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x_1, x_2), (y_1, y_2) &\mapsto x_1y_1 + 5x_2y_2 - 4x_1y_2 \end{aligned}$$

Exercice 3.10. *Considérons*

$$\begin{aligned} \psi : \mathcal{C}[0, 1] &\rightarrow \mathbb{R} \\ f &\mapsto \int_0^1 f(x)^2 dx \end{aligned}$$

Montrer que ψ est une forme quadratique.

Exercice 3.11. *Montrer que*

$$\begin{aligned} \psi : \mathcal{M}_n \times \mathcal{M}_n &\rightarrow \mathbb{R} \\ (A, B) &\mapsto \text{Tr}(A^T \cdot B) \end{aligned}$$

est bilinéaire.

Signe d'une forme quadratique

Exercice 3.12. *Montrer que la forme bilinéaire de l'exercice 3.5 est définie positive.*

Exercice 3.13. *Les applications suivantes sont-elles définies positives ?*

1. $(x_1, x_2, x_3), (y_1, y_2, y_3) \mapsto 2x_1y_1 + 5x_2y_2 + x_3y_3 - 3x_1y_2 - 3x_2y_1$
2. $(x_1, x_2, x_3), (y_1, y_2, y_3) \mapsto x_1y_1 + x_2y_2 + x_3y_3 - 2x_1y_2 - 2x_2y_1 - 2x_1y_3 - 2x_3y_1 - 2x_2y_3 - 2x_3y_2$
3. $(x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4) \mapsto x_1y_2 + x_2y_1 - 2x_1y_3 - 2x_3y_1 + x_2y_3 + x_3y_2 - 2x_3y_4 - 2x_4y_3$
4. $(x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4) \mapsto x_1y_1 + x_2y_2 + 4x_4y_4 + x_1y_2 + x_2y_1 - 2x_1y_4 - 2x_4y_1 + x_2y_3 + x_3y_2 - 2x_3y_4 - 2x_4y_3$
5. $(x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4) \mapsto x_1y_1 + 3x_2y_2 - 2x_1y_2 - 2x_2y_1$

Diagonalisation des formes quadratiques

Exercice 3.14. *Déterminer la base dans laquelle la représentation matricielle des formes quadratiques suivantes sont normalisées :*

1.

$$\begin{aligned} \Phi : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto x_1^2 + 2x_2^2 - 2x_1x_2 \end{aligned}$$

2.

$$\begin{aligned} \Phi : \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\mapsto x_1^2 + 2x_2^2 - 2x_1x_2 \end{aligned}$$

3.

$$\begin{aligned} \Phi : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto x_1^2 - 2x_1x_2 \end{aligned}$$

4.

$$\begin{aligned} \Phi : \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\mapsto x_1^2 - 2x_1x_2 \end{aligned}$$

5. *Les formes quadratiques de l'exercice 3.13.*

3.7 Corrigés des exercices

Espace dual

Corrigé Exercice 3.1. La famille $\{e_1 := (1, 1, 1), e_2 := (0, 1, 1), e_3 := (2, 0, 1)\}$ est bien une base de $\mathbb{R}^3$. Soit

$$f_i : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto \alpha_i x + \beta_i y + \gamma_i z$$

où $(\alpha_i, \beta_i, \gamma_i)_{i \in \{1, 2, 3\}}$ sont des réels à déterminer. On doit résoudre

$$\begin{cases} f_1(e_1) = 1 \\ f_1(e_2) = 0 \\ f_1(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 + \beta_1 + \gamma_1 = 1 \\ \beta_1 + \gamma_1 = 0 \\ 2\alpha_1 + \gamma_1 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = 1 \\ \beta_1 = 2 \\ \gamma_1 = -2 \end{cases}$$

De même

$$\begin{cases} f_2(e_1) = 0 \\ f_2(e_2) = 1 \\ f_2(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 + \beta_2 + \gamma_2 = 0 \\ \beta_2 + \gamma_2 = 1 \\ 2\alpha_2 + \gamma_2 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_2 = -1 \\ \beta_2 = -1 \\ \gamma_2 = 2 \end{cases}$$

Et pour finir

$$\begin{cases} f_3(e_1) = 0 \\ f_3(e_2) = 0 \\ f_3(e_3) = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_3 + \beta_3 + \gamma_3 = 0 \\ \beta_3 + \gamma_3 = 0 \\ 2\alpha_3 + \gamma_3 = 1 \end{cases} \Leftrightarrow \begin{cases} \alpha_3 = 0 \\ \beta_3 = -1 \\ \gamma_3 = 1 \end{cases}$$

La base duale de $\{(1, 1, 1), (0, 1, 1), (2, 0, 1)\}$ est donc $\{f_1, f_2, f_3\}$ où

$$f_1 : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto x + 2y - 2z$$

$$f_2 : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto -x - y + 2z$$

et

$$f_3 : \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto -y + z$$

On peut vérifier qu'avec ces f_1, f_2 et f_3 on a bien $f_1(1, 1, 1) = 1, f_1(0, 1, 1) = 0$ et $f_1(2, 0, 1) = 0$; $f_2(1, 1, 1) = 0, f_2(0, 1, 1) = 1$ et $f_2(2, 0, 1) = 0$; et $f_3(1, 1, 1) = 0, f_3(0, 1, 1) = 0$ et $f_3(2, 0, 1) = 1$.

Corrigé Exercice 3.2. Soit

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

la représentation matricielle de la famille $\{f_1, f_2, f_3\}$ dans la base canonique duale. On a $\det(P) = 0$ donc la famille est liée.

Corrigé Exercice 3.3. Soit

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

la représentation matricielle de la famille $\{f_1, f_2, f_3\}$ dans la base canonique duale. On calcule

$$(P^T)^{-1} = \frac{1}{4} \begin{pmatrix} 0 & 2 & 2 \\ -2 & 3 & 1 \\ 2 & -1 & -3 \end{pmatrix}$$

Donc la base préduale est $\{(0, -1/2, 1/2), (1/2, 3/4, -1/4), (1/2, 1/4, -3/4)\}$. Or la base duale d'une base préduale est la base elle-même donc $\{f_1, f_2, f_3\}$ est une base de $(\mathbb{R}^3)^*$ ce qui répond aussi à la première question.

On peut vérifier comme dans l'exercice 1.

Corrigé Exercice 3.4. 1. On pose

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

On voit que A est de rang 2 donc la famille $\{f_1, f_2\}$ est une famille libre de $(\mathbb{R}^3)^*$.

2. On pose

$$f_3 : \quad \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x, y, z) \mapsto x$$

La représentation matricielle de la famille $\{f_1, f_2, f_3\}$ dans la base canonique dual est

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

qui est de déterminant 2 donc la famille $\{f_1, f_2, f_3\}$ est une base de $(\mathbb{R}^3)^*$.

3. On calcule

$$(P^T)^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & -2 & 3 \\ -1 & -1 & 2 \end{pmatrix}$$

Donc $\{(0, -1, -1), (0, -2, -1), (1, 3, 2)\}$ est la base préduale de $\{f_1, f_2, f_3\}$. On peut vérifier comme dans l'exercice 1.

Formes bilinéaires et formes quadratiques

Corrigé Exercice 3.5. 1. f est de la forme $\sum_i \sum_j a_{ij} x_i y_j$ avec $a_{ij} = a_{ji}$ donc c'est une forme bilinéaire symétrique.

2. Dans la base canonique $\{e_1, e_2, e_3\}$ c'est la matrice de terme général $b(e_i, e_j) = a_{ij}$

$$A := \begin{pmatrix} 1 & -1 & 1 \\ -1 & 5 & -3 \\ 1 & -3 & 6 \end{pmatrix}$$

3. C'est la matrice de terme général $b(\varepsilon_i, \varepsilon_j)$, comme la forme bilinéaire est symétrique il suffit de déterminer les coefficients pour $i \geq j$:

$$B := \begin{pmatrix} 4 & -2 & -2 \\ -2 & 9 & 7 \\ -2 & 7 & 6 \end{pmatrix}$$

on pouvait aussi utiliser la formule de changement de base.

4. La forme quadratique associée à f est

$$Q : \quad \mathbb{R}^3 \rightarrow \mathbb{R} \\ (x_1, x_2, x_3) \mapsto x_1^2 + 5x_2^2 + 6x_3^2 - 2x_1 x_2 + 2x_1 x_3 - 6x_2 x_3$$

5. Comme f est symétrique la forme polaire de Q est f . Donc la représentation matricielle de Q dans la base canonique est A .

6. la représentation matricielle de Q est B .

Corrigé Exercice 3.6. f est de la forme

$$\sum_i \sum_j a_{ij} x_i x_j$$

donc c'est une forme quadratique.

Corrigé Exercice 3.7. 1. f est de la forme $\sum_i \sum_j a_{ij} x_i y_j$ donc c'est une forme bilinéaire.

2. Dans la base canonique $\{e_1, e_2\}$ c'est la matrice de terme général $b(e_i, e_j) = a_{ij}$

$$A := \begin{pmatrix} 1 & -1 \\ 1 & 5 \end{pmatrix}$$

3. C'est la matrice de terme général $b(\varepsilon_i, \varepsilon_j)$:

$$B := \begin{pmatrix} 6 & 2 \\ 0 & 1 \end{pmatrix}$$

on pouvait aussi utiliser la formule de changement de base.

4. La forme quadratique associée à f est

$$Q : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + 5x_2^2$$

5. Comme f n'est pas symétrique la forme polaire de Q n'est pas f . La forme polaire de Q est

$$\tilde{b} : (\mathbb{R}^2)^2 \rightarrow \mathbb{R} \\ (x_1, x_2), (y_1, y_2) \mapsto x_1 y_1 + 5x_2 y_2$$

Comme la forme polaire est symétrique il suffit de calculer les coefficients pour $i \geq j$, la représentation matricielle dans la base canonique est

$$\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

6. Comme la forme polaire est symétrique il suffit de calculer les coefficients pour $i \geq j$, la représentation matricielle de Q dans la base $\{(1, 1), (1, 0)\}$ est

$$\begin{pmatrix} 6 & 1 \\ 1 & 1 \end{pmatrix}$$

Corrigé Exercice 3.8. 1. f est de la forme $\sum_i \sum_j a_{ij} x_i y_j$ donc c'est une forme bilinéaire.

2. Dans la base canonique $\{e_1, e_2\}$ c'est la matrice de terme général $b(e_i, e_j) = a_{ij}$

$$A := \begin{pmatrix} 1 & -4 \\ 0 & 5 \end{pmatrix}$$

3. C'est la matrice de terme général $b(\varepsilon_i, \varepsilon_j)$:

$$B := \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$$

on pouvait aussi utiliser la formule de changement de base.

4. La forme quadratique associée à f est

$$Q : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + 5x_2^2 - 4x_1 x_2$$

5. Comme f n'est pas symétrique la forme polaire de Q n'est pas f . La forme polaire de Q est

$$\tilde{b} : (\mathbb{R}^2)^2 \rightarrow \mathbb{R} \\ (x_1, x_2), (y_1, y_2) \mapsto x_1 y_1 + 5x_2 y_2 - 2x_1 y_2 - 2x_2 y_1$$

Comme la forme polaire est symétrique il suffit de calculer les coefficients pour $i \geq j$, la représentation matricielle dans la base canonique est

$$\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$$

6. Comme la forme polaire est symétrique il suffit de calculer les coefficients pour $i \geq j$, la représentation matricielle de Q dans la base $\{(1, 1), (1, 0)\}$ est

$$\begin{pmatrix} 6 & -1 \\ -1 & 1 \end{pmatrix}$$

Corrigé Exercice 3.9. Ψ est la forme quadratique associée à

$$\begin{aligned} \mathcal{C}[0, 1] &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \int_0^1 f(x)g(x)dx \end{aligned}$$

qui est une forme bilinéaire d'après Exemple 50.

Corrigé Exercice 3.10. La forme est symétrique car pour tout (A, B) dans $\mathcal{M}_n$, $\Psi(A, B) = \Psi(B, A)$.

Pour B fixé par linéarité de la transposée et de la trace on a pour tout (A, B, C) dans $\mathcal{M}_n$

$$\begin{aligned} \Psi(A + \lambda C, B) &= \text{Tr}((A + \lambda C)^T B) = \text{Tr}(A^T B + \lambda C^T B) \\ &= \text{Tr}(A^T B) + \lambda \text{Tr}(C^T B) = \Psi(A, B) + \lambda \Psi(C, B). \end{aligned}$$

Ce qui prouve la linéarité à gauche et donc la bilinéarité par symétrie.

Signe d'une forme quadratique

Corrigé Exercice 3.11. On va faire la méthode de réduction en carrés de Gauss sur Q : en posant $X = (x_1, x_2, x_3)$ on a

$$\begin{aligned} Q(X) &= \underline{x_1^2} + 5x_2^2 + 6x_3^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3 \\ &= (x_1 \underbrace{-x_2 + x_3}_{\text{double produit}})^2 - \underbrace{(-x_2 + x_3)^2}_{\text{correction}} + \underbrace{5x_2^2 + 6x_3^2 - 6x_2x_3}_{\text{ce qui reste}} \\ &= (x_1 - x_2 + x_3)^2 - \underbrace{x_2^2 - x_3^2 + 2x_2x_3 + 5x_2^2 + 6x_3^2 - 6x_2x_3}_{\text{forme quadratique en } x_2 \text{ et } x_3} \\ &= (x_1 - x_2 + x_3)^2 + 4x_2^2 + 5x_3^2 - 4x_2x_3 \\ &= (x_1 - x_2 + x_3)^2 + (2x_2 - x_3)^2 - x_3^2 + 5x_3^2 \\ &= (x_1 - x_2 + x_3)^2 + (2x_2 - x_3)^2 + 4x_3^2 \end{aligned}$$

On peut vérifier en développant l'expression ci-dessus.

La forme bilinéaire est positive car la forme quadratique est la somme de carrés précédés de coefficients tous positifs. De plus $Q(X) = 0$ implique

$$\begin{cases} x_1 - x_2 + x_3 = 0 \\ 2x_2 - x_3 = 0 \\ x_3 = 0 \end{cases}$$

donc $x_1 = x_2 = x_3 = 0$ et la forme bilinéaire est bien symétrique définie positive.

Corrigé Exercice 3.12.

Corrigé Exercice 3.13. 1. La forme quadratique est pour $X = (x_1, x_2, x_3)$

$$\begin{aligned} q(X) &= 2x_1^2 + 5x_2^2 + x_3^2 - 6x_1x_2 \\ &= x_3^2 + 2[x_1^2 - 3x_1x_2] + 5x_2^2 \\ &= x_3^2 + 2\left[\left(x_1 - \frac{3}{2}x_2\right)^2 - \frac{9}{4}x_2^2\right] + 5x_2^2 \\ &= x_3^2 + 2\left(x_1 - \frac{3}{2}x_2\right)^2 + \left(-\frac{9}{2}x_2^2 + 5\right)x_2^2 \\ &= x_3^2 + 2\left(x_1 - \frac{3}{2}x_2\right)^2 + \frac{1}{2}x_2^2 \end{aligned}$$

2. La forme quadratique est pour $X = (x_1, x_2, x_3)$

$$\begin{aligned} q(X) &= x_1^2 + x_2^2 + x_3^2 - 4x_1x_2 - 4x_1x_3 - 4x_2x_3 \\ &= (x_1 - 2x_2 - 2x_3)^2 - (-2x_2 - 2x_3)^2 + x_2^2 + x_3^2 - 4x_2x_3 \\ &= (x_1 - 2x_2 - 2x_3)^2 - 4x_2^2 - 8x_2x_3 - 4x_3^2 + x_2^2 + x_3^2 - 4x_2x_3 \\ &= (x_1 - 2x_2 - 2x_3)^2 - 3x_2^2 - 3x_3^2 - 12x_2x_3 \\ &= (x_1 - 2x_2 - 2x_3)^2 - 3(x_2 + 2x_3)^2 + 12x_3^2 - 3x_3^2 \\ &= (x_1 - 2x_2 - 2x_3)^2 - 3(x_2 + 2x_3)^2 + 9x_3^2 \end{aligned}$$

3. La forme quadratique est pour $X = (x_1, x_2, x_3, x_4)$

$$\begin{aligned} q(X) &= 2x_1x_2 - 4x_1x_3 + 2x_2x_3 - 4x_3x_4 \\ &= (2x_1 + 2x_3)(x_2 - 2x_3) + 4x_3^2 - 4x_3x_4 \\ &= \frac{1}{4}(2x_1 + x_2)^2 - \frac{1}{4}(2x_1 - x_2 + 4x_3)^2 + 4x_3^2 - 4x_3x_4 \\ &= \frac{1}{4}(2x_1 + x_2)^2 - \frac{1}{4}(2x_1 - x_2 + 4x_3)^2 + (2x_3 - x_4)^2 - x_4^2 \end{aligned}$$

4. La forme quadratique est pour $X = (x_1, x_2, x_3, x_4)$

$$\begin{aligned} q(X) &= x_1^2 + x_2^2 + 4x_4^2 + 2x_1x_2 - 4x_1x_4 + 2x_2x_3 - 4x_3x_4 \\ &= (x_1 + x_2 - 2x_4)^2 + 4x_2x_4 + 2x_2x_3 - 4x_3x_4 \\ &= (x_1 + x_2 - 2x_4)^2 + (2x_2 - 4x_4)(x_3 + 2x_4) + 8x_4^2 \\ &= (x_1 + x_2 - 2x_4)^2 + \frac{1}{4}(2x_2 + x_3 - 2x_4)^2 - \frac{1}{4}(2x_2 - x_3 - 6x_4)^2 + 8x_4^2 \end{aligned}$$

5. La forme quadratique est pour $X = (x_1, x_2, x_3, x_4)$

$$\begin{aligned} q(X) &= x_1^2 + 3x_2^2 - 4x_1x_2 \\ &= (x_1 - 2x_2)^2 - x_2^2 \end{aligned}$$

Diagonalisation des formes quadratiques

Corrigé Exercice 3.14. 1. Pour commencer on applique la méthode de réduction en carrés de Gauss :

$$\Phi(x_1, x_2) = (x_1 - x_2)^2 + x_2^2$$

On pose

$$\begin{aligned} l_1 : \mathbb{R}^2 &\rightarrow \mathbb{R} & \text{et} & & l_2 : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto x_1 - x_2 & & & (x_1, x_2) &\mapsto x_2 \end{aligned}$$

Ce sont deux formes linéaires qui forment une base de $(\mathbb{R}^3)^*$. La base préduale est $\{(1, 0), (1, 1)\}$. On peut vérifier que $l_i(e_j) = \delta_{ij}$. La base $\{(1, 0), (1, 1)\}$ est une base Φ -orthogonale de $\mathbb{R}^2$.

De plus la représentation matricielle de Φ dans la Φ -base orthogonale est

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

qui est aussi la matrice normalisée de Φ .

2. On applique la méthode de réduction en carrés de Gauss :

$$\Phi(x_1, x_2, x_3) = (x_1 - x_2)^2 + x_2^2$$

On pose

$$\begin{array}{lcl} l_1 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - x_2 \end{array} \quad \text{et} \quad \begin{array}{lcl} l_2 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 \end{array}$$

Ce sont deux formes linéaires qui forment une famille libre de $(\mathbb{R}^3)^*$. On la complète en une base de $(\mathbb{R}^3)^*$ par la forme linéaire

$$\begin{array}{lcl} l_3 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_3 \end{array}$$

La base préduale est $\{(1, 0, 0), (1, 1, 0), (0, 0, 1)\}$. C'est une base Φ -orthogonale de $\mathbb{R}^3$. De plus la représentation matricielle de Φ dans la base Φ -orthogonale est

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

qui est aussi la matrice normalisée de Φ .

3. On applique la méthode de réduction en carrés de Gauss :

$$\Phi(x_1, x_2) = (x_1 - x_2)^2 - x_2^2$$

On pose

$$\begin{array}{lcl} l_1 : \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x_1, x_2) & \mapsto & x_1 - x_2 \end{array} \quad \text{et} \quad \begin{array}{lcl} l_2 : \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x_1, x_2) & \mapsto & x_2 \end{array}$$

Ce sont deux formes linéaires qui forment une base de $(\mathbb{R}^2)^*$. La base préduale est $\{(1, 0), (1, 1)\}$. On peut vérifier que $l_i(e_j) = \delta_{ij}$. La base $\{(1, 0), (1, 1)\}$ est une base Φ -orthogonale de $\mathbb{R}^2$.

De plus la représentation matricielle de Φ dans la base Φ -orthogonale est

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

qui est aussi la matrice normalisée de Φ .

4. On applique la méthode de réduction en carrés de Gauss :

$$\Phi(x_1, x_2, x_3) = (x_1 - x_2)^2 - x_2^2$$

On pose

$$\begin{array}{lcl} l_1 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - x_2 \end{array} \quad \text{et} \quad \begin{array}{lcl} l_2 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 \end{array}$$

Ce sont deux formes linéaires qui forment une famille libre de $(\mathbb{R}^3)^*$. On la complète en une base de $(\mathbb{R}^3)^*$ par la forme linéaire

$$\begin{array}{lcl} l_3 : \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_3 \end{array}$$

La base préduale est

$$\{(1, 0, 0), (1, 1, 0), (0, 0, 1)\}$$

C'est une base Φ -orthogonale de $\mathbb{R}^3$. De plus la représentation matricielle de Φ dans la base Φ -orthogonale est

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

qui est aussi la matrice normalisée de Φ .

5. (a) On a montré que

$$q(x_1, x_2, x_3) = x_3^2 + 2 \left(x_1 - \frac{3}{2}x_2 \right)^2 + \frac{1}{2}x_2^2$$

La base b -orthogonale est

$$\left\{ (0, 0, 1), (1, 0, 0), \left(\frac{3}{2}, 1, 0 \right) \right\}$$

La représentation matricielle de b est $\text{diag}(1, 2, 1/2)$. Pour obtenir la matrice normalisée de b on "renormalise" la base en

$$\left\{ (0, 0, 1), \left(\frac{1}{\sqrt{2}}, 0, 0 \right), \left(\frac{3\sqrt{2}}{2}, \sqrt{2}, 0 \right) \right\}$$

la représentation matricielle de b dans cette base est I .

(b) On a

$$q(x_1, x_2, x_3) = (x_1 - 2x_2 - 2x_3)^2 - 3(x_2 + 2x_3)^2 + 9x_3^2$$

La base préduale est

$$\{(1, 0, 0), (2, 1, 0), (-2, -2, 1)\}$$

C'est une base b -orthogonale et la représentation matricielle de b est $\text{diag}(1, -1, 9)$. Pour obtenir la matrice normalisée de b on "renormalise" et réorganise la base en

$$\left\{ (1, 0, 0), \left(-\frac{2}{3}, -\frac{2}{3}, \frac{1}{3} \right), \left(\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0 \right) \right\}$$

la représentation matricielle de b dans cette base est

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

(c) On a

$$q(x_1, x_2, x_3, x_4) = \frac{1}{4}(2x_1 + x_2)^2 - \frac{1}{4}(2x_1 - x_2 + 4x_3)^2 + (2x_3 - x_4)^2 - x_4^2$$

La base b -orthogonale est

$$\left\{ \left(\frac{1}{4}, \frac{1}{2}, 0, 0 \right), \left(\frac{1}{4}, -\frac{1}{2}, 0, 0 \right), \left(-\frac{1}{2}, 1, \frac{1}{2}, 0 \right), \left(-\frac{1}{2}, 1, \frac{1}{2}, 1 \right) \right\}$$

La représentation matricielle de b est

$$\text{diag} \left(\frac{1}{4}, -\frac{1}{4}, 1, -1 \right)$$

Pour obtenir la matrice normalisée de b on "renormalise" et on réorganise les vecteurs de la base pour mettre les 1 en début de matrice, la base devient

$$\left\{ \left(\frac{1}{2}, 1, 0, 0 \right), \left(-\frac{1}{2}, 1, \frac{1}{2}, 0 \right), \left(\frac{1}{2}, -1, 0, 0 \right), \left(-\frac{1}{2}, 1, \frac{1}{2}, 1 \right) \right\}$$

la représentation matricielle de b dans cette base est

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

(d) On a

$$q(x_1, x_2, x_3, x_4) = (x_1 + x_2 - 2x_4)^2 + \frac{1}{4}(2x_2 + x_3 - 2x_4)^2 - \frac{1}{4}(2x_2 - x_3 - 6x_4)^2 + 8x_4^2$$

La base b -orthogonale est

$$\left\{ (1, 0, 0, 0), \left(-\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right), \left(-\frac{1}{4}, \frac{1}{4}, -\frac{1}{2}, 0\right), (0, 2, -2, 1) \right\}$$

La représentation matricielle de b est

$$\text{diag} \left(1, \frac{1}{4}, -\frac{1}{4}, 8 \right)$$

Pour obtenir la matrice normalisée de b on "renormalise" et on réorganise les vecteurs de la base pour mettre les 1 en début de matrice, la base devient

$$\left\{ (1, 0, 0, 0), \left(-\frac{1}{2}, \frac{1}{2}, 1, 0\right), \left(0, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{4}\right), \left(-\frac{1}{2}, \frac{1}{2}, -1, 0\right) \right\}$$

la représentation matricielle de b dans cette base est

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

(e) On a

$$q(x_1, x_2, x_3, x_4) = (x_1 - 2x_2)^2 - x_2^2$$

On complète la famille des formes linéaires pour obtenir une base de $(\mathbb{R}^4)^*$ avec

$$l_3 : \mathbb{R}^4 \rightarrow \mathbb{R} \quad \text{et} \quad l_4 : \mathbb{R}^4 \rightarrow \mathbb{R} \\ (x_1, x_2, x_3, x_4) \mapsto x_3 \quad (x_1, x_2, x_3, x_4) \mapsto x_4$$

La base préduale est

$$\{(1, 0, 0, 0), (2, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$$

C'est une base b -orthogonale et la représentation matricielle de b , qui est aussi sa matrice normalisée, est $\text{diag}(1, -1, 0, 0)$.

Chapitre 4

Projection orthogonale dans un espaces Euclidien

Objectif du chapitre :

- Montrer qu'une application est un produit scalaire,
- Déterminer une base orthonormée d'un sous-espace vectoriel pour un produit scalaire donné,
- Déterminer l'orthogonal d'un sous-espace vectoriel pour un produit scalaire donné,
- Caractériser une projection orthogonale sur un sous-espace de dimension finie pour un produit scalaire donné,
- Déterminer la distance d'un vecteur à un sous-espace vectoriel de dimension finie pour un produit scalaire donné.

4.1 Espaces Euclidiens

Il ne faudrait pas sous-estimer la portée de cette section. Après une année à jouer avec des vecteurs qui flottaient dans l'air, les espaces vectoriels, la notion d'orthogonalité, assez proche de celle de perpendicularité vue au collège, ouvre énormément de nouvelles perspectives. Nous allons introduire la notion de produit scalaire qui permet d'étendre à un grand nombre d'espaces toutes les intuitions que l'on a développées en géométrie de l'espace. Un espace de dimension finie équipé d'un produit scalaire est appelé espace Euclidien. L'espace Euclidien est l'espace fondamental de la géométrie, destiné à représenter l'espace physique.

Les géomètres de la Grèce antique ont introduit l'espace Euclidien pour modéliser l'espace physique. Leurs travaux ont été rassemblés par le mathématicien grec Euclide dans ses *Éléments*, avec la grande innovation de prouver toutes les propriétés de l'espace sous forme de théorèmes, en partant de quelques propriétés fondamentales, appelées postulats, qui étaient soit considérées comme évidentes (par exemple, il existe exactement une ligne droite passant par deux points), soit semblaient impossibles à prouver (postulat de la parallèle). Après l'introduction, à la fin du XIXe siècle, de géométries non euclidiennes, les anciens postulats ont été reformalisés pour définir les espaces Euclidiens par le biais de la théorie axiomatique. C'est ce que nous allons présenter ici.

4.1.1 Produit scalaire

Définition 4.1 (► Produit scalaire). *Soit E un espace vectoriel. Une forme bilinéaire $b : E \rightarrow \mathbb{R}$ qui est*

- *symétrique,*
- *définie positive*

est un produit scalaire sur E .

Lorsque E est de dimension finie le couple (E, b) est alors appelé espace Euclidien. En dimension infinie cette structure s'appelle espace préhilbertien.

L'extension de cette structure à la dimension infinie nécessite un peu plus que d'être un espace vectoriel muni d'un produit scalaire, on aura besoin de la complétude de l'espace pour définir ensuite les espaces de Hilbert qui jouent un très grand rôle dans les applications. Cette théorie sera développée dans les cours à suivre.

Exemple 49. *Les formes suivantes sont des produits scalaires :*

1. sur $\mathbb{R}^N$,

$$b: \quad \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$$

$$(x_i)_{i \in \{1, \dots, N\}}, (y_i)_{i \in \{1, \dots, N\}} \mapsto \sum_{i=1}^N x_i y_i$$

2. sur $\mathcal{M}_N$,

$$b: \quad \mathcal{M}_N^2 \rightarrow \mathbb{R}$$

$$(A, B) \mapsto \text{Tr}(A^T B)$$

3. sur $\mathbb{R}_N[X]$,

$$b: \quad \mathbb{R}_N[X]^2 \rightarrow \mathbb{R}$$

$$(P, Q) \mapsto \int_0^1 P(x) Q(x) dx$$

Remarque (►). *Ce dernier exemple fournit aussi un produit scalaire sur l'ensemble $\mathbb{R}[X]$ qui est de dimension infinie. Par contre il n'est pas un produit scalaire sur l'ensemble des fonctions continues sur $\mathbb{R}$ ni même sur l'ensemble des fonctions définies sur $[0, 1]$ comme le montre les contre-exemples ci-dessous :*

- *La fonction qui vaut identiquement 0 sur $]-\infty, 1]$ puis qui croît linéairement sur $[1, \infty[$ annule le produit scalaire sans être identiquement nulle. Donc la forme quadratique associée à b est dégénérée.*
- *La fonction qui vaut 1 en 1/2 est d'intégrale nulle alors qu'elle n'est pas elle-même identiquement nulle. Donc la forme quadratique associée à b est dégénérée.*

Ces deux dernières remarques montrent qu'il faut faire attention avec les espaces de dimension infinie. Pour clarifier le dernier point, les mathématiciens ont développé la théorie de la mesure : La théorie des espaces L^p de fonctions dont la puissance p est intégrable donne un excellent cadre pour lever ces problèmes, pour peu que l'on redéfinisse l'égalité de deux fonctions par le biais de leur intégrale. Dans le même temps cela permet d'unifier le premier et le dernier produit scalaire en introduisant la mesure de comptage. Nous n'en dirons pas davantage cette année.

Cette nouvelle structure permet d'étendre les concepts intuitifs depuis les collèges sur la géométrie dans l'espace de façon bourbakienne, limpide, rigoureuse et qui est développée dans sa grande généralité et puissance. Pour commencer on peut définir la notion de norme :

Définition 4.2 (► Norme Euclidienne). *Soit (E, b) un espace Euclidien. On appelle norme Euclidienne l'application suivante :*

$$N_b: \quad E \rightarrow E$$

$$x \mapsto \sqrt{b(x, x)}$$

Remarquons que la positivité de la forme quadratique associée à b assure que la norme est bien définie. Pour montrer que c'est effectivement une norme il va nous falloir montrer que N vérifie les axiomes d'une norme à savoir :

- Pour tout $x \in E$, $N(x) \geq 0$,
- Pour tout $x \in E$, $N(x) = 0$ implique $x = 0$,
- Pour tout $(x, y) \in E^2$, $N(x + y) \leq N(x) + N(y)$.

Si les deux premières sont des conséquences directes du fait que b est définie positive, la dernière nécessite de prouver plusieurs inégalités qui seront utiles de connaître par elle-même. C'est l'objet de la section suivante.

4.1.2 Inégalités fonctionnelles

Pour se mettre en jambe nous allons illustrer le fameux adage des mathématiciens pures “quand une définition est bien construite, le théorème vient facilement” avec l’immense :

Théorème 4.3 (► Théorème de Pythagore). *Soit (E, b) un espace préhilbertien et N_b la norme Euclidienne associée. On a*

$$N_b(x + y)^2 = N_b(x)^2 + N_b(y)^2 \Leftrightarrow x \perp_b y$$

La preuve de ce théorème est rendue triviale par le travail bourbakiste d’une définition propre de la structure.

► *Démonstration.* Pour tout $(x, y) \in E^2$ on a

$$N_b(x + y)^2 = b(x + y, x + y) = N_b(x)^2 + 2b(x, y) + N_b(y)^2.$$

Donc on a $N_b(x + y)^2 = N_b(x)^2 + N_b(y)^2$ si et seulement si $b(x, y) = 0$. □

Théorème 4.4 (► Théorème de Cauchy-Schwarz). *Soit (E, b) un espace préhilbertien et N_b la norme Euclidienne associée. On a*

$$|b(x, y)| \leq N_b(x) N_b(y).$$

Il existe des livres entiers avec une multitudes de démonstration de ce théorème. Nous en donnons une ici qui est particulièrement élégante parce qu’indirecte.

► *Démonstration.* Soit $(x, y) \in E^2$, et soit t un réel. On a

$$N_b(x + ty)^2 = N_b(x)^2 + 2tb(x, y) + t^2 N_b(y)^2.$$

Mais cette quantité est toujours positive ou nulle. Donc c’est un polynôme en t de signe constant. Le discriminant de ce polynôme est

$$4b(x, y)^2 - 4N_b(x)^2 N_b(y)^2.$$

Ce discriminant est négatif pour que le polynôme soit de signe constant donc

$$4b(x, y)^2 - 4N_b(x)^2 N_b(y)^2 \leq 0$$

ou

$$4b(x, y)^2 \leq 4N_b(x)^2 N_b(y)^2$$

On obtient le résultat en divisant par 4 et en prenant la racine dans chaque membre de l’inégalité. □

Notez que l’on peut déterminer dans cet démonstration les cas d’égalité : on a égalité dans l’inégalité de Cauchy-Schwarz si et seulement si x et y sont colinéaires.

Exemple 50. *Les inégalités de Cauchy-Schwarz pour les produits scalaires définis dans l’exemple 49 sont respectivement*

1. Pour $\mathbb{R}^N$, considérons le produit scalaire usuel défini par

$$\left| \sum_{i=1}^N x_i y_i \right| \leq \sqrt{\sum_{i=1}^N x_i^2} \sqrt{\sum_{i=1}^N y_i^2}.$$

2. Pour $\mathcal{M}_N$,

$$|\mathrm{Tr}(A^T B)| \leq \mathrm{Tr}(A^T A) \mathrm{Tr}(B^T B).$$

3. Pour

$$\left| \int_0^1 P(x)Q(x)dx \right| \leq \sqrt{\int_0^1 [P(x)]^2 dx} \sqrt{\int_0^1 [Q(x)]^2 dx}$$

Ces inégalités se révèlent particulièrement intéressantes pour montrer la convergence de série ou d'intégrales :

Exemple 51. Soit $(x, y, z) \in \mathbb{R}^3$ un vecteur dans la boule unité c-à-d tel que $x^2 + y^2 + z^2 \leq 1$. Montrer que $(x + y + z)^2 \leq 1/3$.

▷ On considère que $\mathbb{R}^3$ le produit scalaire usuel noté b et N_b la norme Euclidienne associée. On a, par l'inégalité de Cauchy-Schwarz (ICS)

$$(x + y + z)^2 = b((x, y, z), (1, 1, 1))^2 \underset{ICS}{\leq} N_b(x, y, z)^2 N_b(1, 1, 1)^2 \leq N_b(1, 1, 1)^2 = \frac{1}{3}$$

Exemple 52. Soit f une fonction continue strictement positive sur $[0, 1]$. Montrer que

$$\int_0^1 \frac{1}{f(x)} dx \geq \frac{1}{\int_0^1 f(x) dx} .$$

▷ On va montrer que

$$\int_0^1 f(x) dx \int_0^1 \frac{1}{f(x)} dx \geq 1 .$$

On a évidemment

$$1 = \int_0^1 \sqrt{f(x)} \frac{1}{\sqrt{f(x)}} dx .$$

Par l'inégalité de Cauchy-Schwarz on a

$$1 = \int_0^1 \sqrt{f(x)} \frac{1}{\sqrt{f(x)}} dx \leq \int_0^1 f(x) dx \int_0^1 \frac{1}{f(x)} dx .$$

ce qui prouve le résultat.

L'inégalité de Cauchy-Schwarz permet aussi de donner une notion d'angle entre deux vecteurs d'un espace vectoriel quelconque, qu'il s'agisse de vecteurs de $\mathbb{R}^N$, de matrices, de polynômes ou de fonctions continues pour ne citer qu'eux. En effet, si (E, b) est un espace préhilbertien et N_b la norme Euclidienne associée, par l'inégalité de Cauchy-Schwarz on a

$$-N_b(x)N_b(y) \leq b(x, y) \leq N_b(x)N_b(y)$$

donc

$$-1 \leq \frac{b(x, y)}{N_b(x)N_b(y)} \leq 1 .$$

Pour peu que l'on sache définir une notion de cosinus sur l'espace que l'on considère, voir remarque ci-dessous, on peut donc définir $\theta \in [0, \pi]$ l'angle entre deux vecteurs de E comme étant l'unique réel tel que

$$\cos(\theta) = \frac{b(x, y)}{N_b(x)N_b(y)}$$

Autrement dit on a la définition suivante

Définition 4.5 (► Angle entre deux vecteurs). Soient (E, b) un espace préhilbertien et N_b la norme Euclidienne associée. Soit $(x, y) \in E^2$, on définit l'angle entre x et y comme étant l'unique réel θ dans $[0, \pi]$ tel que

$$\theta := \arccos \left[\frac{b(x, y)}{N_b(x)N_b(y)} \right] .$$

Remarque (►). Pour définir le cosinus pour un vecteur A d'un espace vectoriel quelconque on peut utiliser la série

$$\sum_{k=0}^M (-1)^k \frac{A^{2k}}{(2k)!}.$$

Cette série converge normalement sur l'ensemble des matrices (par exemple en utilisant $\|AB\| \leq \|A\|\|B\|$, qui repose sur l'inégalité de Cauchy-Schwarz). On définit ainsi le cosinus de A comme

$$\sum_{k=0}^{\infty} (-1)^k \frac{A^{2k}}{(2k)!}.$$

On peut même définir π comme le premier zéro de ce cosinus et montrer que cosinus forme une bijection de $[0, \pi]$ dans $[-1, 1]$. Tout ceci sort du cadre du cours mais l'étudiant.e déterminé.e pourra sans peine s'en convaincre pour l'honneur.

Avec cette inégalité en main nous pouvons démontrer :

Proposition 4.6 (► Inégalité triangulaire). Soit (E, b) un espace préhilbertien et N_b la norme Euclidienne associée. Pour tout $(x, y) \in E^2$ on a

$$N_b(x + y) \leq N_b(x) + N_b(y).$$

Démonstration. Pour tout $(x, y) \in E^2$ on a par l'inégalité de Cauchy-Schwarz

$$\begin{aligned} N_b(x + y)^2 &= N_b(x)^2 + 2b(x, y) + N_b(y)^2 \leq N_b(x)^2 + 2|b(x, y)| + N_b(y)^2 \\ &\leq N_b(x)^2 + 2N_b(x)N_b(y) + N_b(y)^2 = [N_b(x) + N_b(y)]^2. \end{aligned}$$

□

Nous sommes maintenant armés pour prouver

Proposition 4.7 (►). Soit (E, b) un espace préhilbertien. La norme Euclidienne N_b associée à b est une norme.

► *Démonstration.* Nous avons déjà mentionné que la norme était bien définie car $x \mapsto b(x, x)$ est positive.

Par bilinéarité du produit scalaire on a

$$N_b(\lambda x)^2 = b(\lambda x, \lambda x) = \lambda^2 b(x, x) = \lambda^2 N_b(x)^2.$$

On obtient l'homogénéité en prenant la racine carrée.

Comme b est définie positive on a $x \mapsto b(x, x)$ est positive ou nulle et $b(x, x) = 0$ implique que $x = 0$.

Enfin l'inégalité triangulaire a été prouvée dans la Proposition 4.6. □

4.1.3 Base orthonormée

Définition 4.8 (► Vecteurs orthogonaux). Soit (E, b) un espace préhilbertien. Un vecteur $y \in E$ est orthogonal à un vecteur $x \in E$ si $b(x, y) = 0$.

Évidemment comme b est symétrique, si x est orthogonal à y alors y est orthogonal à x et on dit que x et y sont *orthogonaux*. Cette notion correspond à la notion de perpendicularité vue au collège :

Exemple 53. Soit $\mathbb{R}^2$ muni du produit scalaire usuel. Montrer que $(1, 1)$ et $(-1, 1)$ sont orthogonaux.

► On calcule $1 \times -1 + 1 \times 1 = 0$ donc $(1, 1)$ et $(-1, 1)$ sont orthogonaux.

Exemple 54. Soit $\mathbb{R}^2$ et le forme bilinéaire symétrique b définie pour tout $((x_1, x_2), (y_1, y_2))$ dans $\mathbb{R}^2$ par

$$b((x_1, x_2), (y_1, y_2)) = x_1 y_1 + 2x_2 y_2 - x_1 y_2 - x_2 y_1 .$$

Montrer que $(1, 1)$ et $(1, 0)$ sont orthogonaux.

▷ On calcule $1 * 1 = 2(1 * 0) - 1 * 0 - 1 * 1 = 0$ donc $(1, 1)$ et $(1, 0)$ sont orthogonaux pour ce produit scalaire.

Exemple 55. Dans $\mathcal{C}^0[0, \pi]$ on considère la forme bilinéaire symétrique définie par

$$\begin{aligned} \Phi : \quad \mathcal{C}^0[0, \pi] \times \mathcal{C}^0[0, \pi] &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \int_0^\pi f(x) g(x) dx \end{aligned}$$

Montrer que les fonctions $\cos$ et $\sin$ sont orthogonales pour le produit scalaire Φ .

▷ On calcule

$$\int_0^\pi \cos(x) \sin(x) dx = \int_0^\pi \frac{\sin(2x)}{2} dx = -\frac{1}{4} [\cos(2x)]_0^\pi = 0 ,$$

où on a utilisé la formule de trigonométrie $\cos(x) \sin(x) = \sin(2x)/2$ (ça n'est pas parce que ce cours se concentre sur l'algèbre qu'on peut s'affranchir d'une culture mathématiques au sens large). Donc les fonctions $\cos$ et $\sin$ sont orthogonales pour le produit scalaire Φ .

On peut définir la notion suivante :

Définition 4.9 (► Famille b -orthogonale). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie. Une famille $\mathcal{F} := \{\varepsilon_i\}_{i \in \{1, \dots, N\}}$ de F est une famille b -orthogonale si

$$\forall (i, j) \in \{1, \dots, N\}^2, \quad i \neq j \Rightarrow b(\varepsilon_i, \varepsilon_j) = 0 .$$

Exemple 56. Dans $\mathbb{R}^4$ muni du produit scalaire usuel. Montrer que la famille

$$\{(1, 1, 0, 1), (-1, 1, 0, 0), (1, 1, 0, -2)\}$$

forme une famille orthogonale.

▷ On calcule

$$\begin{aligned} b((1, 1, 0, 1), (-1, 1, 0, 0)) &= 0 \\ b((1, 1, 0, 1), (1, 1, 0, -2)) &= 0 \\ b((-1, 1, 0, 0), (1, 1, 0, -2)) &= 0 \end{aligned}$$

donc la famille $\{(1, 1, 0, 1), (-1, 1, 0, 0), (1, 1, 0, -2)\}$ forme une famille orthogonale.

On a déjà vu apparaître la notion de base orthogonale lors de la diagonalisation de formes quadratiques, Proposition 3.25, cette notion a du sens dès que la forme bilinéaire est symétrique mais lorsque la forme bilinéaire est un produit scalaire on peut aller plus loin et définir :

Définition 4.10 (► Base orthogonale et orthonormée). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie.

Une base $(e_i)_{i \in \{1, \dots, N\}}$ de F est une base b -orthogonale si elle est une famille b -orthogonale c -à- d

$$\forall (i, j) \in \{1, \dots, N\}^2, \quad i \neq j \Rightarrow b(e_i, e_j) = 0 .$$

Une base $(\varepsilon_i)_{i \in \{1, \dots, N\}}$ de F est une base b -orthonormée si

$$\forall (i, j) \in \{1, \dots, N\}^2, \quad b(\varepsilon_i, \varepsilon_j) = \delta_{ij} .$$

Ce type de base est amené à jouer un grand rôle dans l'histoire que nous racontons dans ce cours, dans les mathématiques et la société en général. Le fait de travailler avec une base orthogonale simplifie en effet extraordinairement de nombreux calculs qui se révélaient techniques sans cet outil. Nous y reviendrons en abondance dans la suite mais, par exemple, la construction d'une base orthogonale est le cœur de la détermination d'une projection orthogonale que nous verrons plus tard.

L'existence de bases orthogonales en dimension infinie est tellement cruciale que de nombreux chercheurs ayant réussi à en construire ont laissé leur nom : séries de Fourier, polynôme de Legendre, polynôme de Tchebychev, polynôme de Hermite, base de Schauder, base de Hamel, etc. La découverte d'une base orthogonale de dimension finie qui approxime bien des espaces de dimension infinie est aussi une quête frénétique des analystes numériques pour qui une telle base permet de simuler numériquement des systèmes complexes en grandes dimensions, en ingénierie, météo, aéronautique, etc.

Remarque. Dans un espace H de dimension infinie de Hilbert, l'analogie de ces bases orthonormées sont appelées base de Hilbert. En plus de la propriété énoncée dans la définition la base $\mathcal{F}$ doit être totale c-à-d $\text{Vect}(\mathcal{F}) = 0$.

On peut remarquer que

Proposition 4.11 (► Liberté d'une famille orthogonale). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie. Une famille b -orthogonale de vecteurs non-nuls de F est une famille libre.

► *Démonstration.* Notons $\{\varepsilon_i\}_{i \in \{1, \dots, N\}}$ une famille b -orthogonale de vecteurs non-nuls de F . Soit $\{\alpha_i\}_{i \in \{1, \dots, N\}}$ une famille de scalaires tels que

$$\sum_{i=1}^N \alpha_i \varepsilon_i = 0.$$

Alors pour tout $j \in \{1, \dots, N\}$,

$$0 = b \left(\sum_{i=1}^N \alpha_i \varepsilon_i, \varepsilon_j \right) = \alpha_j.$$

Donc la famille est libre. □

L'étudiant.e vif.ve mais angoissé.e s'inquiétera, à la vue de ce que nous avons dit plus haut sur l'importance des bases orthogonales, de l'existence de telles bases. L'étudiant.e inébranlable rétorquera que la procédure de diagonalisation d'une forme quadratique, Proposition 3.25, donne précisément l'existence d'une base b -orthogonale. Et même que la base dans laquelle la matrice de b est normalisée est une base b -orthonormée. Énonçons donc proprement :

Proposition 4.12 (Existence de bases orthogonales de l'espace). Soit (E, b) un espace Euclidien. Il existe une base b -orthonormée de E .

Mais revenons à notre empathie pour l'étudiant.e vif.ve mais angoissé.e qui arguera que l'on a su construire une base b -orthogonale de l'espace vectoriel tout entier mais pas pour un sous-espace vectoriel quelconque. Ne soyez pas si angoisé.e, voici de quoi rassurer :

Proposition 4.13 (► Théorème fondamental des espaces Euclidiens). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie. Il existe une base b -orthogonale (et une base b -orthonormée) de F .

Une fois rassuré.e on pourra se détendre et profiter de la démonstration suivante. Nous allons vivre une expérience assez rare d'un théorème qui est fondamental mais qui a peu d'importance pratique alors que la preuve qui va suivre est incontournable. On aurait pu prouver ce théorème de diverses façons mais la démonstration que nous allons voir ici permet d'associer théorie à la pratique. On peut même lui donner son nom, c'est le *procédé de Gram-Schmidt* qu'il conviendra de connaître aussi bien dans sa formule finale que dans le procédé lui-même :

► *Démonstration.* Soit $\{u_i\}_{\{1, \dots, k\}}$ une base de F . Nous allons construire une base $\{e_i\}_{\{1, \dots, k\}}$ b -orthogonale progressivement puis la normaliser pour obtenir une base b -orthonormée :

— Construction de la base b -orthogonale :

- On pose $e_1 = u_1$.
- On pose ensuite, pour α à déterminer ultérieurement, $e_2 = u_2 + \alpha e_1$. En fait α peut être choisi de sorte que $\{e_1, e_2\}$ forme une famille b -orthogonale c'est-à-dire tel que $b(e_1, e_2) = 0$. En effet

$$0 = b(e_1, e_2) = b(e_1, u_2 + \alpha e_1) = b(e_1, u_2) + \alpha b(e_1, e_1).$$

Donc il suffit de prendre

$$\alpha := -\frac{b(e_1, u_2)}{b(e_1, e_1)}.$$

On a donc

$$e_2 = u_2 - \frac{b(e_1, u_2)}{b(e_1, e_1)} e_1$$

On remarque que par construction $\text{Vect}\{e_1, u_2\} = \text{Vect}\{e_1, e_2\}$ et que $\{e_1, e_2\}$ est une famille b -orthogonale.

- Et on peut itérer le procédé en posant à l'étape suivante $e_3 = u_3 + \beta e_1 + \gamma e_2$ où β et γ peuvent être choisis de sorte que $\{e_1, e_2, e_3\}$ forme une famille b -orthogonale c'est-à-dire tels que $b(e_1, e_3) = 0$ et $b(e_2, e_3) = 0$. En effet

$$\begin{cases} 0 = b(e_1, e_3) \\ 0 = b(e_2, e_3) \end{cases} \Leftrightarrow \begin{cases} 0 = b(e_1, u_3 + \beta e_1 + \gamma e_2) \\ 0 = b(e_2, u_3 + \beta e_1 + \gamma e_2) \end{cases} \Leftrightarrow \begin{cases} 0 = b(e_1, u_3) + \beta b(e_1, e_1) + \gamma b(e_1, e_2) \\ 0 = b(e_2, u_3) + \beta b(e_2, e_1) + \gamma b(e_2, e_2) \end{cases}$$

On pourrait déjà voir qu'il s'agit d'un système linéaire de deux équations à deux inconnues mais en fait c'est un système encore plus simple puisque l'on sait déjà que $b(e_1, e_2) = 0$ donc le système est équivalent à

$$\begin{cases} 0 = b(e_1, u_3) + \beta b(e_1, e_1) \\ 0 = b(e_2, u_3) + \gamma b(e_2, e_2) \end{cases}$$

qui a pour solution

$$\begin{cases} \beta = -\frac{b(e_1, u_3)}{b(e_1, e_1)} \\ \gamma = -\frac{b(e_2, u_3)}{b(e_2, e_2)} \end{cases}$$

Donc a donc

$$e_3 = u_3 - \frac{b(e_1, u_3)}{b(e_1, e_1)} e_1 - \frac{b(e_2, u_3)}{b(e_2, e_2)} e_2$$

On remarque que par construction $\text{Vect}\{e_1, e_2, u_3\} = \text{Vect}\{e_1, e_2, e_3\}$ et que $\{e_1, e_2, e_3\}$ est une famille b -orthogonale.

- En itérant encore ce procédé on peut construire récursivement une base $\{e_i\}_{\{1, \dots, k\}}$ b -orthogonale qui est génératrice de F en posant pour tout $j \in \{1, \dots, k\}$:

$$e_j = u_j - \sum_{i=1}^{j-1} \frac{b(e_i, u_j)}{b(e_i, e_i)} e_i.$$

— Construction de la base b -orthonormée : Pour obtenir une base b -orthonormée $(\varepsilon_i)_{i \in \{1, \dots, k\}}$ on pose pour tout $i \in \{1, \dots, k\}$

$$\varepsilon_i = \frac{e_i}{N_b(e_i)}$$

□

Remarque (►). Soulignons qu'il est aisé d'obtenir une base b -orthonormée à partir d'une base b -orthogonale : il suffit de normaliser en posant

$$\varepsilon_i = \frac{e_i}{N_b(e_i)}$$

où N_b est la norme Euclidienne associée à b .

En effet,

$$b(\varepsilon_i, \varepsilon_i) = b\left(\frac{e_i}{N_b(e_i)}, \frac{e_i}{N_b(e_i)}\right) = \frac{1}{N_b(e_i)^2} b(e_i, e_i) = \frac{N_b(e_i)^2}{N_b(e_i)^2} = 1.$$

Exemple 57. Soit $\mathbb{R}^3$ muni du produit scalaire usuel et N_b la norme Euclidienne associée. On considère $F := \text{Vect}\{(1, 1, 1), (1, 1, 0)\}$.

1. Déterminer une base b -orthogonale de F .
2. Déterminer une base b -orthonormée de F .

►

1. On applique le procédé de Gram-Schmidt. On trouve

$$\varepsilon_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{N_b\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right)^2} b\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

Si deux vecteurs sont b -orthogonaux tous multiples de ces vecteurs forment aussi b -orthogonaux.

On peut donc prendre $\{(1, 1, 1), (1, 1, -2)\}$ qui forme une base b -orthogonale de F . On peut vérifier que $1 \times 1 + 1 \times 1 + 1 \times (-2) = 0$ donc les vecteurs sont bien b -orthogonaux.

2. En normalisant, on obtient donc que

$$\left\{ \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix} \right\}$$

qui forme une base b -orthonormée de F .

Exemple 58. Soit $\mathbb{R}_2[X]$ et la forme bilinéaire définie par

$$\begin{aligned} \Psi : \quad \mathbb{R}_2[X]^2 &\rightarrow \mathbb{R} \\ (P, Q) &\mapsto \int_0^1 P(x) Q(x) e^x dx \end{aligned}$$

Déterminer une base Ψ -orthogonale de $\text{Vect}\{1, X - 1\}$.

► On va utiliser le procédé de Gram-Schmidt. On pose $\varepsilon_1 = 1$ et

$$\varepsilon_2 = X - 1 - \frac{\Psi(1, X - 1)}{\Psi(1, 1)} 1 = X - 1 - \frac{2 - e}{e - 1} = X - \frac{1}{e - 1}.$$

car

$$\Psi(1, 1) = \int_0^1 e^x dx = e - 1$$

et

$$\begin{aligned} \Psi(1, X - 1) &= \int_0^1 (x - 1)e^x dx = \int_0^1 xe^x dx - \int_0^1 e^x dx = - \int_0^1 xe^x dx + [xe^x]_0^1 - \int_0^1 e^x dx \\ &= -2 \int_0^1 xe^x dx + e = -2(e - 1) + e = 2 - e. \end{aligned}$$

Donc $(1, X - 1/(e - 1))$ est une base Ψ -orthogonale.

On peut noter que du fait de la définition de la représentation matricielle d'une forme bilinéaire on a évidemment :

Proposition 4.14 (► Représentation matricielle de b dans une base b -orthonormée). *Soit (E, b) un espace Euclidien. La représentation matricielle de b dans une base b -orthonormée est la matrice identité.*

► *Démonstration.* D'après Proposition 3.14, la représentation matricielle de b dans la base est la matrice de terme général $b(\varepsilon_i, \varepsilon_j)$, mais comme $\{\varepsilon_i\}_{\{1, \dots, N\}}$ est une base orthonormée on a $b(\varepsilon_i, \varepsilon_j) = \delta_{ij}$. Ce qui montre que la matrice est l'identité. $\square$

En conséquence, on peut représenter dans une base b -orthonormée, pour tout $(x, y) \in E^2$, $b(x, y)$ par $X^T Y$. On retrouve bien l'expression du produit scalaire dans $\mathbb{R}^N$ mais c'est aussi le cas pour n'importe quel produit scalaire dans une base b -orthonormée.

Définition 4.15 (► Matrice orthogonale). *Soit (E, b) un espace Euclidien. On appelle matrice b -orthogonale une matrice de passage entre deux bases b -orthonormées.*

Attention, on parle de matrice orthogonale représentant une matrice de passage entre des bases orthonormées et non seulement orthogonales.

Ces matrices ont une propriété fondamentale :

Proposition 4.16 (► Inverse d'une matrice orthogonale). *Soit (E, b) un espace Euclidien. Si P est une matrice b -orthogonale alors*

$$P^{-1} = P^T .$$

► *Démonstration.* Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées. D'après Proposition 4.14, $\text{mat}(b, \mathcal{B}) = I$ et $\text{mat}(b, \mathcal{B}') = I$. D'après la formule de changement de base dans les formes bilinéaires, Proposition 3.16 on a donc pour $P = P_{\mathcal{B}'}^{\mathcal{B}}$

$$\text{mat}(b, \mathcal{B}) = P^T \text{mat}(b, \mathcal{B}') P .$$

ce qui donne $I = P^T I P$ et permet de conclure. $\square$

Remarque (►). *On peut voir que si P est une matrice orthogonale alors elle est de déterminant ± 1 en effet*

$$1 = \det(I) = \det(P P^T) = \det(P)^2 .$$

La réciproque n'est pas vraie, même avec des colonnes qui forment une base orthogonale, comme le montre le contre-exemple suivant

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

Remarque (►). *On peut montrer que l'ensemble des matrices orthogonales forme un groupe, le groupe orthogonal, noté $\mathcal{O}(n)$. L'étude de ce groupe orthogonal est intéressante parce qu'elle permet de voir géométriquement l'effet de passer de la base canonique, qui est orthonormée pour le produit scalaire usuel, à n'importe quelle autre base orthonormée.*

On peut commencer par voir que les matrices du groupe orthogonal sont des isométries, c-à-d qu'elles préservent la norme Euclidienne :

$$\|PX\| = \|X\| ,$$

ou de façon équivalente (exercice) le produit scalaire

$$\langle PX, PY \rangle = \langle X, Y \rangle .$$

Ce dernier point vient du fait que

$$\langle PX, PY \rangle = \langle X, P^T PY \rangle = \langle X, Y \rangle .$$

Un exercice classique consiste à déterminer tous les applications linéaires associées à une matrice orthogonale. Dans $\mathbb{R}^2$, on trouvera les rotations :

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

et les symétries

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

Il n'est pas beaucoup plus difficile de décrire géométriquement toute matrice orthogonale de $\mathbb{R}^3$. Tout ceci est passionnant mais il nous faut continuer à avancer dans le cours.

4.1.4 Espace orthogonal

On peut aussi définir l'équivalent des deux droites perpendiculaires en géométrie Euclidienne :

Définition 4.17 (► Espace b -orthogonal). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E . Le b -orthogonal de F est

$$F^{\perp b} := \{y \in E : b(x, y) = 0, \forall x \in F\}.$$

Proposition 4.18 (► $F^{\perp b}$ est un sous-espace vectoriel). Soient (E, b) un espace préhilbertien et F un sous-espace vectoriel de E . $F^{\perp b}$ est un sous-espace vectoriel.

- *Démonstration.*
- 0 est dans $F^{\perp b}$ donc $F^{\perp b}$ est non-vide.
 - Soient $(y, y') \in [F^{\perp b}]^2$ et $\lambda \in \mathbb{K}$, on a

$$\forall x \in F, b(x, y + \lambda y') = b(x, y) + \lambda b(x, y') = 0.$$

Donc $F^{\perp b}$ est un sous-espace vectoriel de E . □

Nous reviendrons sur cette notion en s'ébattant davantage avec ces espaces en dimension infinie dans le cours "Functional Analysis" en Master. En dimension finie toute la technologie développée en première année peut prendre son essor ici aussi. Notamment cette idée fondatrice que dans un espace de dimension finie la description de cet ensemble infini $F^{\perp b}$ se réduit à la compréhension de ce qui se passe en un nombre fini de vecteurs : les vecteurs de base. En effet on a

Proposition 4.19 (► Caractérisation de $F^{\perp b}$). Soit (E, b) un espace Euclidien et F un sous-espace vectoriel de E de base $\mathcal{F} := \{\varepsilon_i\}_{i \in \{1, \dots, N\}}$. On a

$$F^{\perp b} = \{y \in E : b(\varepsilon_i, y) = 0, \forall i \in \{1, \dots, N\}, \}.$$

► *Démonstration.* Notons $A = \{y \in E : b(\varepsilon_i, y) = 0, \forall i \in \{1, \dots, N\}, \}$. On va montrer que $F^{\perp b} \subset A$ puis que $A \subset F^{\perp b}$.

- $F^{\perp b} \subset A$: Soit $z \in F^{\perp b}$ alors $b(x, z) = 0$ pour tout vecteur x de F donc pour tout vecteur ε_i de la base $\mathcal{F}$. Donc $z \in A$.
- $A \subset F^{\perp b}$ Considérons maintenant un vecteur z de A . Alors $b(\varepsilon_i, z) = 0$ pour tout $i \in \{1, \dots, N\}$. Soit x un vecteur quelconque de F . Comme $\mathcal{F}$ est une base de F il existe un unique $(x_i)_{i \in \{1, \dots, N\}}$ tel que $x = \sum_{i=1}^N x_i \varepsilon_i$. On a alors

$$b(x, z) = b\left(\sum_{i=1}^N x_i \varepsilon_i, z\right) = \sum_{i=1}^N x_i \underbrace{b(\varepsilon_i, z)}_{=0} = 0.$$

Donc z est dans $F^{\perp b}$. □

Exemple 59. Soit $\mathbb{R}^3$ muni du produit scalaire usuel b . Déterminer l'orthogonal de $F = \{(x, y, z) \in \mathbb{R}^3 : x - y = 0\}$.

▷ $\{(1, 1, 0), (0, 0, 1)\}$ est une base de F . Le vecteur (x, y, z) est dans $F^{\perp_b}$ si

$$\begin{cases} b\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = 0 \\ b\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = 0 \end{cases} \Leftrightarrow \begin{cases} x + y = 0 \\ z = 0 \end{cases}$$

Donc $F^{\perp_b} = \text{Vect}\{(1, -1, 0)\}$.

Proposition 4.20 (► Propriétés des orthogonaux). Soit (E, b) un espace préhilbertien. On considère F et G deux sous-espaces vectoriels de E . On a

1. $E^{\perp_b} = 0$,
2. $F \subset G$ implique $G^{\perp_b} \subset F^{\perp_b}$,
3. $F \subset (F^{\perp_b})^{\perp_b}$.

► *Démonstration.* 1. Si x est dans $E^{\perp_b}$, alors en particulier $b(x, x) = 0$. Ce qui implique $x = 0$ par non-dégénérescence de b .

2. Soit $y \in G^{\perp_b}$. Alors pour tout x dans F , x est aussi dans G . Donc $b(x, y) = 0$. Ce qui prouve que $y \in F^{\perp_b}$.

3. Soit $x \in F$. On prend y dans $F^{\perp_b}$, on a alors $b(x, y) = 0$ donc x est dans $(F^{\perp_b})^{\perp_b}$. □

Théorème 4.21 (► Théorème du supplémentaire orthogonal). Soit (E, b) un espace Euclidien. On considère F un sous-espace vectoriel de E . On a

$$F \oplus F^{\perp_b} = E.$$

En conséquence $\dim(F^{\perp_b}) = \dim E - \dim F$.

► *Démonstration.* Si $F = \{0\}$ alors $F^{\perp_b} = E$ et le résultat est évident.

Supposons maintenant que $\dim(F) = p > 0$. Soit $(e_i)_{i \in \{1, \dots, p\}}$ une base b -orthogonale. On complète cette base en une base b -orthogonale $(e_i)_{i \in \{1, \dots, N\}}$ de E . Notons $G = \text{Vect}\{(e_i)_{i \in \{p+1, \dots, N\}}\}$. On va montrer que $G = F^{\perp_b}$.

$(G \subset F^{\perp_b})$ Comme $(e_i)_{i \in \{1, \dots, N\}}$ est une base orthogonale, les vecteurs de la famille $(e_i)_{i \in \{p+1, \dots, N\}}$ sont orthogonaux à tout vecteur de F . Ces vecteurs sont donc dans $F^{\perp_b}$ ainsi que toute combinaison linéaire. Donc $G \subset F^{\perp_b}$.

$(F^{\perp_b} \subset G)$ Soit x un vecteur de $F^{\perp_b}$. Posons

$$x = \sum_{i=1}^N x_i e_i$$

on va montrer que x est dans G .

Pour tout $j \in \{1, \dots, N\}$ on a

$$b(x, e_j) = b\left(\sum_{i=1}^N x_i e_i, e_j\right) = x_j b(e_j, e_j).$$

On a donc

$$x = \sum_{i=1}^N \frac{b(x, e_i)}{b(e_i, e_i)} e_i.$$

Mais comme x appartient à $F^{\perp b}$, pour tout $j \in \{1, \dots, p\}$, on a $b(x, e_j) = 0$. Donc

$$x = \sum_{i=p+1}^N \frac{b(x, e_i)}{b(e_i, e_i)} e_i .$$

ce qui implique que x est dans l'espace G . □

Cette démonstration est intéressante à plusieurs titres, notamment parce qu'elle indique un moyen de construire l'orthogonal d'un sous-espace vectoriel F : on détermine une base orthogonale de F , on la complète en une base de E (ce qui est possible par le théorème de la base incomplète) et on orthogonalise cette base (en utilisant le procédé de Gram-Schmidt par exemple). La complétion de la base orthogonale est alors une base de $F^{\perp b}$.

En dimension infinie, F et $F^{\perp b}$ sont en somme directe puisque si un élément x appartient à F et à $F^{\perp b}$ alors $b(x, x) = 0$ ce qui implique $x = 0$. Par contre Théorème 4.21 n'est pas vrai en dimension infinie comme le prouve l'exemple suivant :

Exemple 60. Soit $\mathbb{R}[X]$ muni du produit scalaire

$$\begin{aligned} \phi : (\mathbb{R}[X])^2 &\rightarrow \mathbb{R} \\ (P, Q) &\mapsto \int_0^1 P(x) Q(x) dx \end{aligned}$$

On considère

$$H := \{P \in \mathbb{R}[X] : P(0) = 0\} .$$

Déterminer $H^{\perp \phi}$.

▷ Soit Q un élément de $H^{\perp \phi}$, on doit avoir pour tout $P \in H$, $\phi(P, Q) = 0$. Donc cela doit être vrai pour $P = XQ$ qui est dans H . On a pour ce P ,

$$0 = \phi(P, Q) = \phi(XQ, Q) = \int_0^1 x [Q(x)]^2 dx .$$

L'intégrand est une fonction continue positive d'intégrale nulle donc $XQ \equiv 0$ sur $[0, 1]$. D'où l'on déduit $Q \equiv 0$ sur $(0, 1]$. Mais comme Q est un polynôme cela induit que $Q \equiv 0$ sur $\mathbb{R}$. On a donc prouvé que $H^{\perp \phi} = \{0\}$.

Dans cet exemple il n'est pas vrai que $H \oplus H^{\perp \phi} = \mathbb{R}[X]$.

Remarque (►). Dans un espace de Hilbert H en dimension infinie, on peut démontrer que

$$F \oplus F^{\perp} = H \quad \Leftrightarrow \quad F \text{ est fermé} \quad \Leftrightarrow \quad (F^{\perp})^{\perp} = F .$$

On peut démontrer que l'orthogonal d'un hyperplan vectoriel en dimension quelconque est soit $\{0\}$ soit une droite vectorielle. Voir <https://math-os.com/orthogonal-sev/>

L'exemple 60 ci-dessus montre aussi que, en général, l'on ne peut pas espérer pouvoir améliorer le dernier point de Proposition 4.20 et prouver $F = (F^{\perp b})^{\perp b}$. Par contre en dimension finie on a

Proposition 4.22 ($F = (F^{\perp b})^{\perp b}$). Soient (E, b) un espace Euclidien et F un sous-espace vectoriel de E . On a $F = (F^{\perp b})^{\perp b}$.

Démonstration. Par Théorème 4.21, on a

$$\dim(F^{\perp b}) = \dim(E) - \dim(F)$$

mais on a aussi

$$\dim((F^{\perp b})^{\perp b}) = \dim(E) - \dim(F^{\perp b})$$

donc

$$\dim((F^{\perp b})^{\perp b}) = \dim(E) - \dim(F^{\perp b}) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F)$$

Donc $F = (F^{\perp b})^{\perp b}$. □

Terminons cette section par une proposition qui est très pratique dans le cas de $\mathbb{R}^N$ muni du produit scalaire usuel :

Proposition 4.23 (► Orthogonal d'un hyperplan pour le produit scalaire usuel). *Soit $\mathbb{R}^N$ muni du produit scalaire usuel noté b . Soit $\alpha := (\alpha_i)_{i \in \{1, \dots, N\}} \neq (0, \dots, 0)$ et l'hyperplan*

$$H = \{x := (x_i)_{i \in \{1, \dots, N\}} : \sum_{i=1}^N \alpha_i x_i = 0\}.$$

On a $H^\perp = \text{Vect}(\alpha)$.

► *Démonstration.* • Le sous-espace vectoriel H peut être ré-écrit

$$\{x : b(x, \alpha) = 0\}.$$

H est un hyperplan parce qu'il est le noyau de la forme linéaire $g : x \mapsto b(x, \alpha)$. Le théorème du rang assure que

$$\underbrace{\dim(\mathbb{R}^N)}_{=N} = \dim(\underbrace{\ker g}_{=H}) + \underbrace{\text{rg}(g)}_{=1}$$

Donc $\dim(H) = N - 1$.

• On va commencer par montrer que $[\text{Vect}(\alpha)]^\perp \subset H$: Tout élément de $\text{Vect}(\alpha)$ est de la forme $\lambda\alpha$, pour $\lambda \in \mathbb{R}$. Soit z dans $[\text{Vect}(\alpha)]^\perp$. On a, pour tout $\lambda \in \mathbb{R}$

$$0 = b(z, \lambda\alpha) = \lambda b(z, \alpha).$$

D'où l'on déduit, par définition de H , que z est dans H . Ce qui implique $[\text{Vect}(\alpha)]^\perp \subset H$.

• On va maintenant prouver que $[\text{Vect}(\alpha)]^\perp = H$: D'après Théorème 4.21, $\dim[\text{Vect}(\alpha)]^\perp = N - 1$. Or $\dim(H) = N - 1$ donc $[\text{Vect}(\alpha)]^\perp = H$.

• Pour conclure on utilise Proposition 4.22 qui implique que $H^\perp = \{[\text{Vect}(\alpha)]^\perp\}^\perp = \text{Vect}(\alpha)$. $\square$

4.2 Projection orthogonale

La projection orthogonale est un des apprentissages les plus importants de ce cours au niveau pratique en économétrie naturellement, ainsi qu'en statistique, mais aussi en optimisation, en science des grandes données, etc. Cette notion sera étendue aux espaces de Hilbert dans le cours "Functional Analysis".

Dans cette section, pour ne pas alourdir les formules, le produit scalaire sera noté $\langle \cdot, \cdot \rangle$ et on omettra souvent de préciser que l'orthogonal se réfère à ce produit scalaire.

Rappelons tout d'abord quelques résultats généraux sur l'ensemble des projecteurs dont font partie les projections orthogonales :

Définition 4.24 (► Projecteur). *Soient E un espace vectoriel et $p \in \mathcal{L}(E)$. L'endomorphisme p est un projecteur si $p^2 = p$.*

Par Définition 3.9, en terme matricielle si P est une matrice carrée, P est (isomorphe à) un projecteur si $P = P^2$.

Exemple 61. *Montrer que*

$$\begin{aligned} p : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (0, 2x + y) \end{aligned}$$

est un projecteur.

▷ *la représentation matricielle de p est*

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

On calcule

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

Donc p est un projecteur.

Proposition 4.25 (► p et $\text{id} - p$ sont supplémentaires). Soient E un espace vectoriel et p un projecteur.

1. L'endomorphisme $\text{id} - p$ est un projecteur,
2. $\text{Im}(p) = \ker(\text{id} - p)$.
3. $\text{Im}(\text{id} - p) = \ker(p)$.

► *Démonstration.* 1. L'endomorphisme $\text{id} - p$ est un projecteur car

$$(\text{id} - p)^2 = \text{id} - 2p + p^2 = \text{id} - 2p + p = \text{id} - p .$$

2. Soit $y \in \ker(\text{id} - p)$ alors $y - p(y) = 0$. Autrement dit $y = p(y)$. Donc il existe $x = y$ tel que $y = p(x)$. Ce qui implique que y est dans $\text{Im}(p)$.

Réciproquement, soit $y \in \text{Im}(p)$ alors il existe x tel que $y = p(x)$. En composant par p on obtient pour ce x ,

$$p(y) = p[p(x)] = p(x) = y .$$

D'où $p(y) = y$. Ce qui implique que y est dans $\ker(\text{id} - p)$.

3. Comme $\text{id} - p$ est un projecteur, on peut appliquer le point précédent à $\text{id} - p$. On obtient

$$\text{Im}(\text{id} - p) = \ker(\text{id} - (\text{id} - p)) = \ker(p) .$$

□

Proposition 4.26 (► Image et noyau sont supplémentaires). Soient E un espace vectoriel et p un projecteur. On a $E = \ker(p) \oplus \text{Im}(p)$.

► *Démonstration.* — On va montrer que $\ker(p)$ et $\text{Im}(p)$ sont en somme directe. Soit $y \in \ker(p) \cap \text{Im}(p)$. On a

$$\begin{cases} \exists x \in E, y = p(x) \\ p(y) = 0 \end{cases} \Leftrightarrow \begin{cases} \exists x, y = p(x) \\ p(p(x)) = 0 \end{cases} \Leftrightarrow \begin{cases} \exists x, y = p(x) \\ p(x) = 0 \end{cases} \Leftrightarrow \begin{cases} \exists x, y = p(x) \\ y = 0 \end{cases}$$

Donc $\ker(p) \cap \text{Im}(p) = 0$.

— Soit $x \in E$. On peut décomposer x sous la forme

$$x = \underbrace{x - p(x)}_{\in \ker(p)} + \underbrace{p(x)}_{\in \text{Im}(p)}$$

Donc $x = u + v$ où $u \in \ker(p)$ et $v \in \text{Im}(p)$.

□

Un projecteur p est donc caractérisé par la donnée de son noyau et de son image. On dit que p est un projecteur sur $\text{Im}(p)$ de direction $\ker(p)$. D'après Proposition 4.25, on a

$$\begin{cases} \forall x \in \text{Im}(p), p(x) = x \\ \forall x \in \ker(p), p(x) = 0 . \end{cases} \quad (4.1)$$

Dans la cadre préhilbertien on peut aller plus loin et définir

Définition 4.27 (► Projecteur orthogonal). Soient (E, b) un espace préhilbertien et $p \in \mathcal{L}(E)$. L'endomorphisme p est appelé projection orthogonale si

- p est un projecteur,

- Pour tout $(x, y) \in E^2$, $b(p(x), y) = b(x, p(y))$.

la projection dépend du produit scalaire considéré on pourrait donc la noter p_b mais nous ne le ferons pas pour ne pas alourdir les notations.

Par Définition 3.9, en terme matricielle si P est une matrice carrée, P est (isomorphe à) une projection orthogonale si $P = P^2 = P^T$.

Exemple 4.1. Montrer que la fonction

$$f : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto & \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \end{array}$$

est une projection orthogonale.

▷ La représentation matricielle de f dans la base canonique est

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

On calcule

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

Donc P est un projecteur. De plus on voit que $P^T = P$ donc f est une projection orthogonale.

On a vu Proposition 4.26 que $\ker(p)$ et $\text{Im}(p)$ sont supplémentaires. Dans le cadre Euclidien on a même :

Proposition 4.28 (► $\ker(p)$ et $\text{Im}(p)$ sont orthogonaux). Soient (E, b) un espace préhilbertien et p une projection orthogonale. $\ker(p)$ et $\text{Im}(p)$ sont b -orthogonaux.

En terme de projecteur cela signifie que p est une projection sur $\text{Im}(p)$ de direction $\ker(p)$ qui lui est b -orthogonal. Cela justifie la terminologie. On précisera donc juste l'espace sur lequel on projette et la direction est l'orthogonal.

► *Démonstration.* Soit $x \in \ker(p)$ et $y \in \text{Im}(p)$. Il existe $z \in E$, tel que $y = p(z)$. On a donc

$$b(x, y) = b(x, p(z)) = b(p(x), z) = 0.$$

Donc x et y sont orthogonaux. □

Proposition 4.29 (► $p_F + p_{F^\perp} = \text{id}$). Soient (E, b) un espace préhilbertien, soit F un sous-espace vectoriel de dimension finie de E et p_F une projection orthogonale sur F .

L'endomorphisme $\text{id} - p_F$ est une projection orthogonale sur $F^{\perp b}$, notée $p_{F^{\perp b}}$ et on a

$$p_F + p_{F^\perp} = \text{id}. \tag{4.2}$$

L'équation (4.2) a une conséquence pratique hyper-importante puisqu'elle permet de déduire aisément une projection orthogonale de la projection sur son orthogonal. Donc quand il nous faut déterminer une projection sur un sous-espace vectoriel F , il peut être très avantageux de commencer par se demander lequel entre F et $F^{\perp b}$ est l'espace de plus petite dimension.

► *Démonstration.* D'après Proposition 4.25, $p_{F^{\perp b}} = \text{id} - p_F$ est un projecteur. De plus

$$\begin{aligned} b((\text{id} - p_F)x, y) &= b(x - p_F(x), y) = b(x, y) - b(p_F(x), y) = b(x, y) - b(x, p_F(y)) = b(x, y - p_F(y)) \\ &= b(x, (\text{id} - p_F)(y)) \end{aligned}$$

donc $p_{F^{\perp b}}$ est une projection orthogonale.

D'après Proposition 4.25, $p_{F^{\perp b}}$ est une projection sur $\text{Im}(p_{F^{\perp b}}) = \text{Im}(\text{id} - p_F) = \ker(p_F) = F^{\perp b}$. □

4.2.1 Expression de la projection orthogonale

Nous sommes maintenant prêts à énoncer un résultat qui permet de déterminer aisément l'expression d'une projection orthogonale.

Théorème 4.30 (► Expression de la projection orthogonale dans le cas d'une base orthogonale). Soient (E, b) un espace Euclidien, $\|\cdot\|_b$ la norme Euclidienne associée, F un sous-espace vectoriel de E , $\{e_i\}_{i \in \{1, \dots, k\}}$ est une base b -orthogonale de F . La projection orthogonale sur F , p_F est

$$p_F(x) = \sum_{i=1}^k b(x, e_i) \frac{e_i}{\|e_i\|_b^2}. \quad (4.3)$$

Attention ce résultat n'est vrai que pour les bases orthogonales. Cela n'est pas du tout vrai pour une base quelconque. Complétons cette remarque en rappelant que l'on a à notre disposition une technique pour transformer n'importe quelle base en une base orthogonale : le procédé de Gram-Schmidt. Donc la détermination d'une projection orthogonale se réduit à utiliser la procédé de Gram-Schmidt et appliquer simplement (4.3).

En utilisant Remarque 4.1.3, ce théorème se traduit par

Corollaire 4.31 (► Expression de la projection orthogonale dans le cas d'une base orthonormée). Soient (E, b) un espace Euclidien, F un sous-espace vectoriel de E , $\{\varepsilon_i\}_{i \in \{1, \dots, k\}}$ est une base b -orthonormée de F . La projection orthogonale sur F , p_F est

$$p_F(x) = \sum_{i=1}^k b(x, \varepsilon_i) \varepsilon_i. \quad (4.4)$$

Ce résultat repose sur un résultat que nous avons déjà évoqué mais qu'il me semble pertinent de mettre en valeur en l'écrivant sous forme de lemme :

Lemma 4.32 (► Coordonnées dans une base orthogonale). Soient (E, b) un espace Euclidien, $\|\cdot\|_b$ la norme Euclidienne associée et F un sous-espace vectoriel de E . On considère une base b -orthogonale $\mathcal{F} := (e_i)_{i \in \{1, \dots, k\}}$ de F . Un vecteur $x \in F$ a pour coordonnées $(x_i)_{i \in \{1, \dots, k\}}$ où

$$x_i = \frac{b(x, e_i)}{\|e_i\|_b^2} \quad \forall i \in \{1, \dots, k\}$$

► *Démonstration.* D'après Théorème 4.21, F et $F^{\perp b}$ sont supplémentaires donc

$$\exists!(x_F, x_{F^{\perp b}}) \in F \times F^{\perp b}, \quad x = x_F + x_{F^{\perp b}}.$$

Comme x est dans F , les coordonnées de x dans $F^{\perp b}$ sont toutes nulles.

Soit $i \in \{1, \dots, k\}$, on calcule

$$b(x, e_i) = b\left(\sum_{j=1}^k x_j e_j, e_i\right) = \sum_{j=1}^k x_j \underbrace{b(e_j, e_i)}_{=0 \text{ si } i \neq j} = x_i b(e_i, e_i) = x_i \|e_i\|_b^2.$$

D'où le résultat □

Ce lemme est assez remarquable et c'est une des raisons de la puissance des bases orthogonales, puisque il permet de déterminer le i -ème coordonnées d'un vecteur sans qu'il ne soit nécessaire de résoudre un système de N équation à N inconnues (ce qui revient à inverser une matrice et qui est très coûteux).

► *Démonstration de Théorème 4.30.* Soit $x \in E$. D'après (4.1), on a

$$p(x) = p(x_F + x_{F^{\perp b}}) = \underbrace{p(x_F)}_{=x_F} + \underbrace{p(x_{F^{\perp b}})}_{=0} = x_F.$$

Il suffit donc de déterminer les coordonnées de x sur le sous-espace vectoriel F . D'après Lemme 4.32 on a donc

$$p(x) = x_F = \sum_{i=1}^k \frac{b(x, e_i)}{\|e_i\|_b^2} e_i$$

ce qui est le résultat annoncé. $\square$

Exemple 62. Soit $\mathbb{R}^3$, muni du produit scalaire usuel $\langle \cdot, \cdot \rangle$. Déterminer la représentation matricielle dans la base canonique de la projection sur $F := \{(x, y, z) \in \mathbb{R}^3 : x - y = 0\}$.

$\triangleright F$ est l'hyperplan de base $\{(1, 1, 0), (0, 0, 1)\}$. C'est une base orthogonale. On peut donc appliquer Théorème 4.30 :

$$\begin{aligned} p \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{\left\| \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\|^2} \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\left\| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\|^2} \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \frac{1}{2} (x + y) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (x + y)/2 \\ (x + y)/2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \\ &= \begin{pmatrix} (x + y)/2 \\ (x + y)/2 \\ z \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

Donc la représentation matricielle dans la base canonique de la projection sur F est

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

4.2.2 Expression matricielle de la projection orthogonale

L'apprentissage par cœur de formules ne me satisfait pas, peut-être parce que je n'imagine pas que l'on puisse les retenir plus que quelques heures. Conscient que mon égo peut me jouer des tours et que le cerveau de l'étudiant.e toulousain.e puisse être plus performant que le mien je donne dans cette section quelques formules qui permettent de déterminer encore plus rapidement l'expression de la projection orthogonale. L'étudiant.e qui cherche à retenir sur le long terme comment déterminer une projection orthogonale, et ma considération distinguée, aura tout de même grand intérêt à comprendre en profondeur la section précédente et ne se servir bêtement des formules qui vont suivre que pour gagner un peu de temps en contrôle et garder plus de temps sur le reste de la copie pour montrer l'étendue de sa maîtrise du cours.

Proposition 4.33 (► Représentation matricielle de la projection orthogonale : cas général). Soient (E, b) un espace Euclidien, $\mathcal{E}$ une base de E et D la représentation matricielle du produit scalaire b dans la base $\mathcal{E}$. On considère F un sous-espace vectoriel de E et A la matrice dont les colonnes forment une base de F . Si P est la représentation matricielle dans la base $\mathcal{E}$ de la projection orthogonale sur F alors

$$P = A(A^T D A)^{-1} A^T D$$

► *Démonstration.* • Montrons que P est un projecteur :

$$P^2 = A(A^T DA)^{-1} A^T DA (A^T DA)^{-1} A^T D = A(A^T DA)^{-1} A^T D = P .$$

Donc P est un projecteur.

• Montrons que P est une projection orthogonale :

$$\begin{aligned} b(PX, Y) &= b(A(A^T DA)^{-1} A^T DX, Y) \\ &= [A(A^T DA)^{-1} A^T DX]^T DY \\ &= X^T D^T A (A^T D^T A)^{-1} A^T DY \\ &= X^T DA (A^T DA)^{-1} A^T DY \\ &= b(X, A(A^T DA)^{-1} A^T DY) \\ &= b(X, PY) . \end{aligned}$$

Donc P est une projection orthogonale.

• On va maintenant démontrer que $\ker(P) = F^{\perp b}$ ce qui permettra de conclure que P est une projection orthogonale sur F . On va montrer dans un premier temps que $F^{\perp b} \subset \ker P$ puis dans un deuxième temps que $\ker(p) \subset F^{\perp b}$.

($F^{\perp b} \subset \ker P$) : Soit $(e_i)_{i \in \{1, \dots, N\}}$ la base canonique de E . En remarquant que Ae_i est la i -ème colonne de A , et donc un vecteur de F , on obtient pour tout $i \in \{1, \dots, N\}$ et tout $X \in F^{\perp b}$

$$0 = \underbrace{\langle Ae_i, X \rangle}_{\in F} = (Ae_i)^T DX = e_i^T A^T DX .$$

Pour tout $X \in F^{\perp b}$, on a donc montré que l'application linéaire $Y \in E \mapsto Y^T A^T DX$ est nulle pour tout élément e_i de base de E d'où l'on déduit que $A^T DX = 0$ sur E . On a ainsi pour tout $X \in F^{\perp b}$,

$$P(X) = A(A^T DA)^{-1} \underbrace{A^T DX}_{=0} = 0 .$$

Ce qui permet de conclure que $F^{\perp b} \subset \ker P$.

($\ker(p) \subset F^{\perp b}$) : On va commencer par montrer que $F \subset \ker(I - P)$. On a évidemment

$$PA - A = A(A^T DA)^{-1} A^T DA - A = 0 .$$

Ainsi pour tout $i \in \{1, \dots, N\}$

$$PAe_i - Ae_i = 0 .$$

En rappelant que Ae_i est le i -ème vecteur de la base de F , on a montré que l'application linéaire $Y \in E \mapsto PY - Y$ est nulle sur tout vecteur de base de F et donc sur F , autrement dit $F \subset \ker(I - P)$. Or $\ker(I - P) = \text{Im}(P)$, donc ceci implique que $F \subset \ker(I - P) = \text{Im}(P)$. En passant aux orthogonaux, par Proposition 4.20, on en déduit que $\text{Im}(P)^{\perp b} \subset F^{\perp b}$. On conclue en utilisant que $\text{Im}(P)^{\perp b} = \ker(P)$ d'après Proposition 4.28. Ainsi $\ker(p) \subset F^{\perp b}$. □

Dans le cas où l'on veut écrire la représentation matricielle d'une projection orthogonale dans une base orthonormée pour le produit scalaire, ce qui est par exemple le cas lorsque l'on cherche à déterminer la représentation matricielle dans la base canonique d'une projection orthogonale dans $\mathbb{R}^N$ pour le produit scalaire usuel, on a $D = I$ dans la proposition précédente. On a donc

Proposition 4.34 (► Représentation matricielle de la projection orthogonale dans une base orthonormée). *Soient (E, b) un espace Euclidien et $\mathcal{E}$ une base b -orthonormée de E . On considère F un sous-espace vectoriel de E et A est la matrice dont les colonnes forment une base de F . Si P est la représentation matricielle dans la base b -orthonormée $\mathcal{E}$ de la projection orthogonale sur F on a*

$$P = A(A^T A)^{-1} A^T$$

Dans la cas où la base de F est normalisée, au sens où les vecteurs de la base de F sont tous de norme 1, on a $A^T A = \langle A, A \rangle = \|A\|^2 = 1$. L'expression ci-dessus conduit alors à

Proposition 4.35 (► Représentation matricielle de la projection orthogonale dans une base orthonormée et une base normalisée de F). *Soient (E, b) un espace Euclidien et $\mathcal{E}$ une base b -orthonormée de E . On considère F un sous-espace vectoriel de E et A est la matrice dont les colonnes forment une base normalisée de F . Si P est la représentation matricielle dans la base b -orthonormée $\mathcal{E}$ de la projection orthogonale sur F on a*

$$P = AA^T . \quad (4.5)$$

4.2.3 Applications à la distance d'un vecteur à un sous-espace vectoriel

Voici un théorème qui embrasse, pour finir dans un "grande finale" digne des feux d'artifice les plus grandioses, l'algèbre à l'analyse. Il est celui qui justifie d'avoir fait tous ces magnifiques tribulations dans ce monde poétique de l'algèbre pour revenir aux vraies mathématiques utiles : l'analyse. Je dis ça parce que je suis bêtement analyste appliqué mais j'ai évidemment, vous avez pu le voir au travers de ces pages, le plus grand respect pour les algébristes et les mathématiciens purs en général. Une autre espèce menacée par notre volonté d'optimiser un objectif, au lieu de laisser le diversité trouver son chemin. Bref je m'égare, revenons à notre fameux théorème.

Afin de ne pas perdre de vue que les résultats qui suivent dépendent du produit scalaire considéré le produit scalaire sera noté b dans les résultats suivant

Théorème 4.36 (► Théorème de projection orthogonale). *Soient (E, b) un espace Euclidien, $\|\cdot\|_b$ la norme Euclidienne associée et F un sous-espace vectoriel de E . Soit p_F la projection orthogonale sur F . Pour tout $x \in E$, on a*

$$\inf_{y \in F} \|x - y\|_b = \min_{y \in F} \|x - y\|_b = \|x - p_F(x)\|_b$$

Remarque (►). *On pourra étendre ce résultat à des espaces qui sont convexes et fermés dans des espaces de Hilbert de dimension infinie.*

► *Démonstration du Théorème 4.36.* On a

$$\|x - y\|_b^2 = \|x - p_F(x) + p_F(x) - y\|_b^2 = \|x - p_F(x)\|_b^2 + 2b(x - p_F(x), p_F(x) - y) + \|p_F(x) - y\|_b^2$$

D'après Proposition 4.25, $x - p_F(x)$ appartient à $\ker(p_F)$. D'autre part, $p_F(x)$ et donc $p_F(x) - y$ appartiennent à F et donc à $\text{Im}(p_F)$. Or $\ker(p_F)$ et $\text{Im}(p_F)$ sont b -orthogonaux, d'après Proposition 4.28. Donc le théorème de Pythagore, Théorème 4.3, implique

$$\|x - y\|_b^2 = \|x - p_F(x)\|_b^2 + \|p_F(x) - y\|_b^2 .$$

Cette expression est une somme de carré dont le premier terme de dépend pas de y alors que le deuxième est minimal est égal à 0, lorsque $y = p_F(x)$. Donc l'infimum du problème de minimisation est obtenue en $y = p_F(x)$ et vaut dans ce cas $\|x - p_F(x)\|_b$. □

Ceci permet de définir

Définition 4.37 (► Distance d'un vecteur à un sous-espace vectoriel). *Soient (E, b) un espace Euclidien $\|\cdot\|_b$ la norme Euclidienne associée et F un sous-espace vectoriel de E . Soit p_F la projection orthogonale sur F . Soit x dans E . On appelle distance de x à F et on note $d_b(x, F)$ le réel*

$$d_b(x, F) = \min_{y \in F} \|x - y\|_b .$$

Exemple 63. *Soit $\mathbb{R}^3$, muni du produit scalaire usuel. Déterminer la distance de $(1, 0, 0)$ à $F := \{(x, y, z) \in \mathbb{R}^3 : x - y = 0\}$.*

▷ On a vu que la représentation matricielle de la projection orthogonale sur F dans la base canonique est

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

donc

$$p \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}.$$

On calcule alors

$$\left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - p \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1/2 \\ -1/2 \\ 0 \end{pmatrix} \right\| = \frac{1}{2} \left\| \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\| = \frac{\sqrt{2}}{2}.$$

Proposition 4.38 (► Projection orthogonale sur un hyperplan pour le produit scalaire usuel). Soit $\mathbb{R}^N$ muni du produit scalaire usuel $\langle \cdot, \cdot \rangle$. Soit $a \in \mathbb{R}^N$ différent du vecteur nul. Soit l'hyperplan $H = \{x \in \mathbb{R}^n : \langle x, a \rangle = 0\}$.

La projection orthogonale sur H est

$$p_H(x) = x - \frac{\langle x, a \rangle}{\|a\|^2} a. \quad (4.6)$$

► *Démonstration.* D'après Proposition 4.23 sur l'expression, $H^\perp = \text{Vect}(a)$. Donc d'après Théorème 4.30,

$$p_{F^\perp}(x) = \frac{\langle x, a \rangle}{\|a\|^2} a.$$

Par Equation (4.2) $p_H + p_{H^\perp} = \text{id}$ ce qui donne le résultat. $\square$

4.3 Exercices

Produit scalaire

Exercice 4.1. Soit

$$\begin{aligned} \Phi : A &\rightarrow \mathbb{R} \\ f &\mapsto \int_0^1 f(x)^2 dx \end{aligned}$$

1. Déterminer Ψ la forme polaire associée à Φ .
2. Si A est l'ensemble des applications de $[0, 1]$ à valeurs dans $\mathbb{R}$. Est-ce que Ψ définit un produit scalaire sur A ?
3. Si A est l'ensemble des applications de carrés intégrables de $[0, 1]$ à valeurs dans $\mathbb{R}$. Est-ce que Ψ définit un produit scalaire sur A ?
4. Si A est l'ensemble des applications continues de $[0, 1]$ à valeurs dans $\mathbb{R}$. Est-ce que Ψ définit un produit scalaire sur A ?
5. Si A est l'ensemble des applications continues de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Est-ce que Ψ définit un produit scalaire sur A ?
6. Si A est l'ensemble des fonctions polynômes. Est-ce que Ψ définit un produit scalaire sur A ?

Exercice 4.2. Montrer que

$$\begin{aligned} \psi : \mathcal{M}_n \times \mathcal{M}_n &\rightarrow \mathbb{R} \\ (A, B) &\mapsto \text{Tr}(A^T \cdot B) \end{aligned}$$

est un produit scalaire.

Exercice 4.3. Montrer que si une fonction est de carré intégrable sur $[0, 1]$ alors elle est intégrable sur $[0, 1]$.

Procédé d'orthonormalisation de Gram-Schmidt

Exercice 4.4. Soit $\mathbb{R}^3$ muni du produit scalaire usuel :

$$b : \quad \mathbb{R}^3 \times \mathbb{R}^3 \quad \rightarrow \quad \mathbb{R}$$

$$\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right), \left(\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right) \mapsto x_1y_1 + x_2y_2 + x_3y_3$$

1. Déterminer une base orthogonale de $\mathbb{R}^3$.
2. Déterminer une base orthogonale de $\text{Vect}\{(1, 1, 1), (1, 2, -1)\}$.
3. Déterminer une base orthogonale de $\text{Vect}\{(1, 2, -1), (1, 1, 1)\}$.
4. Déterminer une base orthogonale de $\text{Vect}\{(1, 1, 1), (1, 2, -1), (0, 1, -2)\}$.
5. Déterminer une base orthogonale de $\text{Vect}\{(1, 1, 1), (1, 2, -1), (0, 1, 1)\}$.
6. Déterminer une base orthogonale de $\{(x, y, z) \in \mathbb{R}^3 : x + y = 0\}$.

Exercice 4.5. Soit $\mathbb{R}^4$ muni du produit scalaire usuel. Déterminer une base orthogonale de $\text{Vect}\{(1, 1, 1, 0), (0, 2, -1, 1), (2, 1, 0, 1)\}$

Exercice 4.6. Dans $\mathbb{R}^3$ considérons

$$b : \quad \mathbb{R}^3 \times \mathbb{R}^3 \quad \rightarrow \quad \mathbb{R}$$

$$\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right), \left(\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right) \mapsto x_1y_1 + 2x_2y_2 + x_3y_3 - x_1y_2 - x_2y_1$$

1. Déterminer une base b -orthogonale de

$$F := \{(x, y, z) \in \mathbb{R}^3 : x - y = 0\}.$$

2. Déterminer une base b -orthogonale de

$$\text{Vect} \left\{ \left(\begin{array}{c} \sqrt{2} \\ \sqrt{3} \\ 0 \end{array} \right), \left(\begin{array}{c} -19 \\ \sqrt{5} \\ 14 \end{array} \right), \left(\begin{array}{c} 0 \\ 10 \\ \sqrt{7} \end{array} \right) \right\}$$

Projection orthogonale

Exercice 4.7. Dans $\mathbb{R}^3$ muni du produit scalaire usuel.

1. Déterminer l'orthogonal de $H := \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0 \text{ et } x - 2y = 0\}$.
2. Déterminer la représentation matricielle dans la base canonique de la projection orthogonale sur H .
3. Quelle est la distance de $(2, 1, 3)$ à H ?
4. Quelle est la distance de $(1, 1, 1)$ à H ?

Exercice 4.8. Dans $\mathbb{R}^3$ muni du produit scalaire usuel. Déterminer la représentation matricielle dans la base canonique de la projection orthogonale sur $H := \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0\}$.

Exercice 4.9. Dans $\mathbb{R}^4$ muni du produit scalaire usuel.

1. Déterminer l'orthogonal de $H := \{(x, y, z, t) \in \mathbb{R}^4 : x - y + z - t = 0 \text{ et } y - 2t = 0\}$.
2. Déterminer la représentation matricielle dans la base canonique de la projection orthogonale sur H .

Exercice 4.10. Dans $\mathbb{R}^3$ considérons

$$b : \quad \mathbb{R}^3 \times \mathbb{R}^3 \quad \rightarrow \quad \mathbb{R}$$

$$\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right), \left(\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right) \mapsto 10x_1y_1 + x_2y_2 + 2x_3y_3 + 2x_1y_2 + 2x_2y_1 - x_2y_3 - x_3y_2$$

1. Déterminer l'orthogonal pour b de

$$F := \{(x, y, z) \in \mathbb{R}^3 : x - y = 0 \text{ et } y - z = 0\}.$$

2. Déterminer une base b -orthogonale de F .

3. Déterminer la matrice de la projection orthogonale sur F pour le produit scalaire b .

4. Déterminer la distance de $(1, 0, 1, 0)$ à F pour le produit scalaire b .

Exercice 4.11. Dans $\mathbb{R}^4$ considérons

$$b : \quad \mathbb{R}^4 \times \mathbb{R}^4 \quad \rightarrow \quad \mathbb{R}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \mapsto x_1y_1 + x_2y_2 + 4x_4y_4 + x_1y_2 + x_2y_1 - 2x_1y_4 - 2x_4y_1 + x_2y_3 + x_3y_2 - 2x_3y_4 - 2x_4y_3$$

1. Déterminer l'orthogonal pour b de

$$F := \{(x, y, z, t) \in \mathbb{R}^4 : x - y = 0 \text{ et } y - z = 0\}.$$

2. Déterminer une base b -orthogonale de F .

3. Déterminer la matrice de la projection orthogonale sur F pour le produit scalaire b .

4. Déterminer la distance de $(1, 0, 1, 0)$ à F pour le produit scalaire b .

4.4 Corrigés des exercices

Produit scalaire

Corrigé Exercice 4.1. 1. La forme polaire associée à Φ est :

$$\Pi : \quad A \rightarrow \mathbb{R}$$

$$(f, g) \mapsto \int_0^1 f(x)g(x)dx.$$

2. Non car Ψ peut n'être pas même définie.

3. Non car l'application identiquement nulle et qui vaut 1 en $1/2$ annule la forme quadratique sans être l'application identiquement nulle. Donc la forme quadratique est dégénérée.

4. Oui. La bilinéarité est une conséquence immédiate de la linéarité de l'intégrale. La positivité est une conséquence de la positivité de l'intégrale d'une fonction positive. La non-dégénérescence vient du fait que si une fonction est positive un un point alors elle est positive dans un interval autour de ce point dont l'intégrale ne peut pas être nulle. On peut l'écrire avec des ϵ .

5. Non car la fonction identiquement nulle sur $[0, 1]$ mais qui croît linéairement en dehors de cet intervalle annule la forme quadratique mais n'est pas identiquement nulle.

6. Oui car si un polynôme est nul sur un intervalle il est nul partout.

Corrigé Exercice 4.2. La bilinéarité est une conséquence immédiate de la linéarité de la transposée et de la trace. Le caractère définie positive se voit lorsque l'on utilise la définition du produit matriciel.

Procédé d'orthonormalisation de Gram-Schmidt

Corrigé Exercice 4.3. 1. La base canonique est une base orthonormée de $\mathbb{R}^3$.

2. On utilise le procédé d'orthonormalisation de Gram-Schmidt :

— On pose $\varepsilon_1 = (1, 1, 1)$.

— On pose $\varepsilon_2 = (1, 2, -1) + \alpha\varepsilon_1$ où α est tel que $\langle \varepsilon_1, \varepsilon_2 \rangle = 0$ c'est à dire

$$0 = \langle \varepsilon_1, \varepsilon_2 \rangle = \left\langle \varepsilon_1, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \alpha\varepsilon_1 \right\rangle = \left\langle \varepsilon_1, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\rangle + \alpha \langle \varepsilon_1, \varepsilon_1 \rangle$$

On prend donc

$$\alpha = -\frac{\left\langle \varepsilon_1, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\rangle}{\langle \varepsilon_1, \varepsilon_1 \rangle} = -\frac{\left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle} = \frac{2}{3}$$

On obtient donc

$$\varepsilon_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} - \frac{2}{3}\varepsilon_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix}$$

Comme si deux vecteurs sont orthogonaux, on peut multiplier chacun d'eux par un réel non-nul et toujours obtenir une famille orthogonale, on peut choisir

$$\varepsilon_2 = \begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix}$$

On vérifie d'ailleurs que

$$\langle \varepsilon_1, \varepsilon_2 \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix} \right\rangle = 0$$

donc $\{(1, 1, 1), (1, 4, -5)\}$ est bien une base orthogonale.

3. On a $\text{Vect}\{(1, 1, 1), (1, 2, -1)\} = \text{Vect}\{(1, 2, -1), (1, 1, 1)\}$. Donc $\{(1, 1, 1), (1, 4, -5)\}$ est une base orthogonale de cet espace.
4. On remarque que $(0, 1, -2) = -(1, 1, 1) + (1, 2, -1)$ donc $\text{Vect}\{(1, 1, 1), (1, 2, -1), (0, 1, -2)\} = \text{Vect}\{(1, 1, 1), (1, 2, -1)\}$. Ainsi $\{(1, 1, 1), (1, 4, -5)\}$ est une base orthogonale de cet espace.
5. La famille $\{(1, 1, 1), (1, 2, -1), (0, 1, 1)\}$ est une famille libre de $\mathbb{R}^3$ donc l'espace est $\mathbb{R}^3$. Une base orthogonale de cet espace est donné par la base canonique.
6. On a $\{(x, y, z) \in \mathbb{R}^3 : x + y = 0\} = \text{Vect}\{(1, -1, 0), (0, 0, 1)\}$. Or $\{(1, -1, 0), (0, 0, 1)\}$ est une famille orthogonale donc $\{(1, -1, 0), (0, 0, 1)\}$ est une base orthogonale de l'espace.

Corrigé Exercice 4.4. On applique le procédé de Gram-Schmidt on obtient :

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

— On a

$$\varepsilon_2 = \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\|^2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix}$$

On peut choisir

$$\varepsilon_2 = \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix}$$

On vérifie que

$$\left\langle \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 0$$

— On a

$$\varepsilon_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\|^2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix} \right\|^2} \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix} = \frac{1}{51} \begin{pmatrix} 57 \\ -30 \\ -27 \\ 33 \end{pmatrix}$$

On peut choisir

$$\varepsilon_3 = \begin{pmatrix} 57 \\ -30 \\ -27 \\ 33 \end{pmatrix}$$

On vérifie que

$$\left\langle \begin{pmatrix} 57 \\ -30 \\ -27 \\ 33 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 0$$

et

$$\left\langle \begin{pmatrix} 57 \\ -30 \\ -27 \\ 33 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -4 \\ 3 \end{pmatrix} \right\rangle = 0$$

Projection orthogonale

Corrigé Exercice 4.5. 1. On détermine $H = \text{Vect}\{2, 1, 3\}$. Donc (x, y, z) est dans l'orthogonal de H si et seulement si

$$\left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\rangle = 0$$

Donc $H^\perp = \{(x, y, z) \in \mathbb{R}^3 : 2x + y + 3z = 0\}$.

2. Soit p la projection orthogonale sur H , on a

$$\begin{aligned} p \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{\left\| \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\|^2} \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\rangle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \\ &= \frac{2x + y + 3z}{14} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \\ &= \frac{1}{14} \begin{pmatrix} 4x + 2y + 6z \\ 2x + y + 3z \\ 6x + 3y + 9z \end{pmatrix} \\ &= \frac{1}{14} \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

Donc la représentation matricielle de p dans la base canonique est

$$\frac{1}{14} \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix}$$

On vérifie que $P^T = P$ ainsi que

$$\frac{1}{14} \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

On pourrait aussi vérifier que $P^2 = P$ et que pour les vecteurs x de $H^\perp$ on a $p(x) = 0$.

3. $(2, 1, 3)$ est dans H donc $d((2, 1, 3), H) = 0$.

4. On calcule

$$\frac{1}{14} \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 6 \\ 3 \\ 9 \end{pmatrix}$$

donc

$$d \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, H \right) = \left\| \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{7} \begin{pmatrix} 6 \\ 3 \\ 9 \end{pmatrix} \right\| = \frac{1}{7} \left\| \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix} \right\| = \frac{\sqrt{21}}{7}$$

Corrigé Exercice 4.6. D'après le cours, $H^\perp = \text{Vect}\{(1, 1, -1)\}$. On applique la formule de projection orthogonale $p_{H^\perp}$ sur $H^\perp$:

$$\begin{aligned} p_{H^\perp} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{\left\| \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\|^2} \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \\ &= \frac{x + y - z}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} x + y - z \\ x + y - z \\ -x - y + z \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

Donc la représentation matricielle de p dans la base canonique est

$$\frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

On obtient la représentation matricielle de la projection sur H en utilisant $p_H + p_{H^\perp} = \text{id}$ donc

$$p_H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Corrigé Exercice 4.7. 1. On a $H = \text{Vect}\{(-1, 0, 1, 0), (3, 2, 0, 1)\}$. Le vecteur (x, y, z, t) appartient à $H^\perp$ si et seulement si

$$\begin{cases} \left\langle \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 0 \\ \left\langle \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 0 \end{cases} \Leftrightarrow \begin{cases} -x + z = 0 \\ 3x + 2y + t = 0 \end{cases}$$

On a donc $H^\perp = \{(1, 0, 1, -3), (0, 1, 0, -2)\}$.

2. On applique le procédé d'orthogonalisation de Gram-Schmidt à $\{(-1, 0, 1, 0), (3, 2, 0, 1)\}$. On obtient que $\{(-1, 0, 1, 0), (3, 4, 3, 2)\}$ est une base orthogonale de H . On applique la formule de projection orthogonale et on obtient que la représentation matricielle de la projection orthogonale sur H est donnée par

$$\frac{1}{19} \begin{pmatrix} 14 & 6 & -5 & 3 \\ 6 & 8 & 6 & 4 \\ -5 & 6 & 14 & 3 \\ 3 & 4 & 3 & 2 \end{pmatrix}$$

Corrigé Exercice 4.8. 1. On a $F = \text{Vect}\{(1, 1, 1)\}$. Donc (x, y, z) appartient à $F^{\perp_b}$ si et seulement si

$$0 = b \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = 10x + y + 2z + 2y + 2x - z - y = 12x + 2y + z.$$

Donc $F^{\perp_b} := \{(x, y, z) \in \mathbb{R}^3 : 12x + 2y + z = 0\}$.

2. La famille $\{(1, 1, 1)\}$ est une base b -orthogonale de F .

3. La formule de projection orthogonale donne

$$\begin{aligned} p \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{\left\| \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\|_b^2} b \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{15} (12x + 2y + z) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} 12x + 2y + z \\ 12x + 2y + z \\ 12x + 2y + z \end{pmatrix} \end{aligned}$$

Donc la représentation matricielle de la projection orthogonale sur F pour le produit scalaire b est donnée par

$$P = \frac{1}{15} \begin{pmatrix} 12 & 2 & 1 \\ 12 & 2 & 1 \\ 12 & 2 & 1 \end{pmatrix}$$

On peut vérifier que $P(1, 1, 1) = (1, 1, 1)$. Par contre la propriété que $P^T = P$ n'est vraie que pour le produit scalaire usuel. Dans la cas général si on appelle A la représentation matricielle de b on doit avoir, avec un léger abus de notation,

$$b(PX, Y) = b(X, PY) \Leftrightarrow (PX)^T AY = X^T APY \Leftrightarrow X^T P^T AY = X^T APY$$

donc $P^T A = AP$. Ce qu'il est bien moins aisé de vérifier.

Corrigé Exercice 4.9. Il y avait une erreur dans l'énoncé car b n'est pas un produit scalaire.

Si b avait été un produit scalaire on aurait pu utiliser la même méthode que pour l'exercice 4, il faut juste un peu faire attention à bien appliquer la formule de la projection orthogonale avec le produit scalaire b et utiliser pour la projection la norme Euclidienne associée au produit scalaire b . Il faut aussi appliquer le procédé de Gram-Schmidt avec le produit scalaire b .

Corrigé Exercice 4.10. 1. On calcule $\chi_A = -(X+2)(X-2)(X-4)$ et $E_{-2} = \text{Vect}\{(1, -2, 1)\}$, $E_2 = \text{Vect}\{(-1, 0, 1)\}$, $E_4 = \text{Vect}\{(1, 1, 1)\}$. La base $\{(1, -2, 1), (-1, 0, 1), (1, 1, 1)\}$ est une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\{(1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6}), (-1/\sqrt{2}, 0, 1/\sqrt{2}), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{-2, 2, 4\}$ et

$$\begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$$

2. On calcule $\chi_A = -X^2(X-3)$ et $E_0 = \text{Vect}\{(-1, 0, 1), (-1, 1, 0)\}$, $E_3 = \text{Vect}\{(1, 1, 1)\}$. La base de E_0 n'est pas orthogonale. On l'orthogonalise en utilisant le procédé de Gram-Schmidt pour obtenir une base orthogonale de E_0 : $\{(-1, 0, 1), (-1, 2, -1)\}$. La base $\{(-1, 0, 1), (-1, 2, -1), (1, 1, 1)\}$ est bien une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\{(-1/\sqrt{2}, 0, 1/\sqrt{2}), (-1/\sqrt{6}, 2/\sqrt{6}, -1/\sqrt{6}), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{0, 0, 3\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

3. On calcule $\chi_A = -(X+3)^2(X-3)$ et $E_{-3} = \text{Vect}\{(-1, 0, 1), (2, 1, 0)\}$, $E_3 = \text{Vect}\{(1, -2, 1)\}$. La base de E_{-3} n'est pas orthogonale. On l'orthogonalise en utilisant le procédé de Gram-Schmidt pour obtenir une base orthogonale de E_{-3} : $\{(-1, 0, 1), (1, 1, 1)\}$. La base $\{(-1, 0, 1), (1, 1, 1), (1, -2, 1)\}$ est bien une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\{(-1/\sqrt{2}, 0, 1/\sqrt{2}), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), (1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{0, 0, 3\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{pmatrix}$$

4. On calcule $\chi_A = -(X + 1)X(X - 3)$ et $E_{-1} = \text{Vect}\{(-1, 0, 1)\}$, $E_0 = \text{Vect}\{(0, 1, 0)\}$, $E_3 = \text{Vect}\{(3, 4, 3)\}$. La base $\{(-1, 0, 1), (0, 1, 0), (3, 4, 3)\}$ est une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\left\{(-1/\sqrt{2}, 0, 1/\sqrt{2}), (0, 1, 0), (3/\sqrt{34}, 4/\sqrt{34}, 3/\sqrt{34})\right\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{-2, 2, 4\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & 0 & 3/\sqrt{34} \\ 0 & 1 & 4/\sqrt{34} \\ 1/\sqrt{2} & 0 & 3/\sqrt{34} \end{pmatrix}$$

Corrigé Exercice 4.11. On utilise la méthode habituelle et $A^n = PD^nP^{-1}$. Comme la matrice est symétrique on sait que A est diagonalisable et que l'on peut diagonaliser A dans une base orthonormée de vecteurs propres. C'est ce que nous avons fait dans l'exercice ci-dessus. À dire vrai entre le temps que 'on gagne dans le calcul de P^{-1} et le risque d'erreur dans le calcul de $A^n = PD^nP^T$, il n'est pas clair que la normalisation apporte grand chose.

Troisième partie
Pour aller plus haut

Chapitre 5

Connections entre applications linéaires et formes bilinéaires et applications

Le programme que nous nous étions fixé est maintenant achevé. Les résultats énoncés jusqu'ici permettent en effet de diagonaliser une matrice (et même d'en trigonaliser certaines), de déterminer une projection orthogonale et de caractériser si une forme quadratique est définie positive/négative ou pas. Ce nouveau chapitre a pour ambition d'explorer une connexion que nous avons souvent évoquée entre applications linéaires et formes bilinéaires.

Nous allons tout d'abord énoncer et démontrer le théorème spectral qui indique que l'on peut diagonaliser toute matrice symétrique. Ensuite nous énoncerons puis démontrerons le théorème de Riesz-Fréchet qui exhibe la connexion entre applications linéaires et formes bilinéaires dans $\mathbb{R}^N$. Enfin nous énoncerons puis démontrerons la loi d'inertie de Sylvester qui indique, essentiellement, que le nombre de carrés précédés d'un coefficient positif ne dépend pas de la méthode choisie dans la réduction en carrés de Gauss.

Cette extension de la théorie donnera de nouvelles méthodes pour répondre à des questions que nous savons déjà traiter autrement. Ces nouvelles méthodes seront parfois plus élégantes, parfois plus rapides mais surtout elles seront toujours la révélation d'une connexion magique entre ces deux mondes. C'est un passage secret que je souhaiterais emprunter avec vous maintenant. Nous n'en explorerons pas tous les trésors mais nous en ramasserons ensemble quelques bijoux. Alors pour celles et ceux qui cherchent juste des méthodes pour suivre dans les années à venir la lecture de ce chapitre peut être repoussée à plus tard, mais pour les plus courageux, accrochez vous, respirez et profitez du paysage.

5.1 Théorème spectral

Théorème 5.1 (► Théorème spectral). *Soit $A \in \mathcal{M}_N(\mathbb{R})$ une matrice symétrique. A est diagonalisable dans une base orthonormée de vecteurs propres.*

La notion de base orthonormée est à comprendre au sens du produit scalaire usuel sur $\mathbb{R}^n$:

$$\langle (x_i)_{i \in \{1, \dots, N\}}, (y_i)_{i \in \{1, \dots, N\}} \rangle = \sum_{i=1}^N x_i y_i .$$

Il nous faut tout d'abord montrer le lemme suivant :

Lemma 5.2 (►). *Soit $A \in \mathcal{M}_N(\mathbb{R})$. Si A est une matrice symétrique alors le polynôme caractéristique de A est scindé dans $\mathbb{R}$.*

► *Démonstration.* Comme $\mathbb{C}$ est algébriquement clôt la polynôme caractéristique de A , χ_A , est scindé dans $\mathbb{C}$. On va montrer que toute valeur propre de A est en fait réelle. Soit $\lambda \in \text{Sp}(A)$ alors d'après Proposition 1.18 on sait déjà que $\bar{\lambda}$ est aussi une valeur propre de A . Soit X_λ un vecteur propre de λ , et donc $\overline{X_\lambda}$ un vecteur propre de $\bar{\lambda}$. On a

$$\langle AX_\lambda, \overline{X_\lambda} \rangle = \langle \lambda X_\lambda, \overline{X_\lambda} \rangle = \lambda \langle X_\lambda, \overline{X_\lambda} \rangle .$$

Mais comme A est symétrique on a d'autre part, par Définition 3.10 de la transposée,

$$\langle AX_\lambda, \overline{X_\lambda} \rangle = \langle X_\lambda, A^T \overline{X_\lambda} \rangle = \langle X_\lambda, A \overline{X_\lambda} \rangle = \langle X_\lambda, \overline{\lambda X_\lambda} \rangle = \overline{\lambda} \langle X_\lambda, \overline{X_\lambda} \rangle.$$

Mais comme $\langle X_\lambda, \overline{X_\lambda} \rangle \neq 0$, c'est en fait $\|X_\lambda\|^2$, ceci implique que $\lambda = \overline{\lambda}$. Donc que toute valeur propre est réelle. $\square$

► *Démonstration du Theorem 5.1.* On va démontrer ce théorème par récurrence sur la dimension de l'espace vectoriel. Soit pour $n \in \mathbb{N}$, la propriété $\mathcal{P}(n)$: "tout endomorphisme symétrique d'un espace de dimension n est diagonalisable dans une base orthonormée de vecteurs propre".

$\mathcal{P}(1)$ est vraie.

Supposons que $\mathcal{P}(n)$ soit vraie à un rang n donné. On considère un espace vectoriel E de dimension $n + 1$. D'après Lemme 5.2, le polynôme caractéristique χ_A a au moins une racine : a . Considérons $F := [\text{Vect}\{a\}]^\perp$ l'orthogonal de $\text{Vect}\{a\}$. Comme $\text{Vect}\{a\}$ est de dimension 1, Théorème 4.21 assure que F est de dimension n . Considérons l'application linéaire $f|_F$ qui est la restriction de l'application linéaire associée à A à l'espace F , c-à-d

$$f|_F : \begin{array}{ccc} F & \rightarrow & E \\ x & \mapsto & f(x) \end{array}$$

Cette application hérite de f les propriétés de linéarité. On va prouver que $f|_F$ est un endomorphisme de F : soit $x \in F$ on a

$$\langle f|_F(x), a \rangle = \langle f(x), a \rangle = \langle x, f^T(a) \rangle = \langle x, f(a) \rangle = \langle x, a \rangle = 0$$

où on a utilisé successivement la définition de la transposée, le fait que f est symétrique puis que a est un vecteur propre. Donc l'image de x par $f|_F$ est orthogonal à a et est donc dans F . L'application $f|_F$ est donc un endomorphisme de F qui est de dimension n . Par hypothèse de récurrence $f|_F$ est diagonalisable dans une base orthonormée de vecteurs propres. Notons $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$ la base orthonormée de F dans laquelle $f|_F$ est diagonale $\text{diag}\{(\lambda_i)_{i \in \{1, \dots, n\}}\}$. Dans la base $\{a, \varepsilon_1, \dots, \varepsilon_n\}$ la représentation matricielle de f est $\text{diag}(a, \lambda_1, \dots, \lambda_n)$. Donc l'endomorphisme est bien diagonalisable dans une base orthonormée de vecteurs propres de l'espace de dimension $n + 1$. Ce qui assure l'hérédité de $\mathcal{P}$ et finit la démonstration. $\square$

5.2 Théorème de représentation de Riez-Fréchet

Nous allons ici donner une introduction à la représentation de Riez-Fréchet. Il va s'agir ici simplement de montrer qu'il y a un isomorphisme et de l'expliciter entre les applications linéaires et les formes linéaires.

Théorème 5.3 (► Théorème de représentation de Riez-Fréchet, dans $\mathbb{R}^N$). *Soit $(E, \langle \cdot, \cdot \rangle)$ un espace Euclidien et a une forme bilinéaire symétrique sur E . Il existe un unique endomorphisme $f \in \mathcal{L}(E)$ tel que*

$$\forall (x, y) \in E^2 \quad a(x, y) = \langle f(x), y \rangle.$$

Ce résultat est un cas simple du théorème de représentation de Riesz-Fréchet dans un espace de Hilbert qui dépasse le cadre de ce cours : Si a est une forme bilinéaire continue sur un espace de Hilbert réel H (ou une forme sesquilinéaire complexe continue sur un Hilbert complexe), alors il existe une unique application A de H dans H telle que, pour tout $(u, v) \in H \times H$, on ait $a(u, v) = \langle Au, v \rangle$. De plus, A est linéaire et continue, de norme égale à celle de a .

► *Démonstration de Théorème.* Rappelons en préliminaire que d'après Proposition 4.14 dans une base orthonormée pour le produit scalaire $\langle \cdot, \cdot \rangle$ on a $\langle x, y \rangle = X^T Y$.

Soient A la représentation matricielle de a . On a pour tout $(x, y) \in E^2$

$$a(x, y) = X^T A Y = (A^T X)^T Y = \langle f(x), y \rangle$$

si f est l'endomorphisme associé à A^T et où on a noté X la représentation matricielle de x et Y celle de y . $\square$

On a prouvé qu'il y a un isomorphisme entre formes linéaires et endomorphisme et on voit que cette identification se fait à travers les représentation matricielle après transposition.

5.3 Loi d'inertie de Sylvester

Définition 5.4 (► Signature d'une forme quadratique). Soient E un espace vectoriel et $q \in \mathcal{Q}(E)$. On appelle signature de q et on note σ_q le couple (α, β) où α est la dimension maximale des sous-espaces F tels que

$$\forall v \in F \setminus \{0\}, \quad q(v) > 0$$

et β est la dimension maximale des sous-espaces G tels que

$$\forall v \in G \setminus \{0\}, \quad q(v) < 0.$$

la méthode de Gauss permet de déterminer aisément la dimension de ces sous-espaces : en effet, α est le nombre de carrés précédés d'un coefficient positif alors que β est le nombre de carrés précédés d'un coefficient négatif dans la réduite de Gauss. Mais l'étudiant.e anxieux.se pourrait s'inquiéter que le nombre de coefficient qui sont positifs, et le nombre de coefficients qui sont négatifs puisse dépendre pas de l'ordre dans lequel on choisit d'éliminer les variables dans la méthode de Gauss. C'est une très bonne remarque mais il n'en est rien comme l'énonce la loi d'inertie de Sylvester :

Proposition 5.5 (► Loi d'inertie de Sylvester, 1852). Soient E une espace vectoriel et $q \in \mathcal{Q}(E)$ de signature (α, β) . Pour toute base q -orthogonale on a

$$\alpha = \text{Card}(\{i : q(\varepsilon_i) > 0\}) \quad \text{et} \quad \beta = \text{Card}(\{i : q(\varepsilon_i) < 0\}).$$

Rappelons que si un théorème porte le nom de son découvreur c'est souvent signe qu'il a profondément marqué les mathématiques. Rappelons aussi le théorème d'Arnold "si un théorème porte un nom ce n'est jamais le nom de celui qui l'a découvert".

► *Démonstration.* On considère deux bases orthogonales $\mathcal{E} = \{\varepsilon_i\}_{i \in \{1, \dots, N\}}$ et $\mathcal{E}' = \{\varepsilon'_i\}_{i \in \{1, \dots, N\}}$. Notons r le nombre de vecteurs de $\mathcal{E}$ sur lesquels q est strictement positive. On appelle I les indices de ces vecteurs :

$$I := \{i : q(\varepsilon_i) > 0\}.$$

Notons r' le nombre de vecteurs de $\mathcal{E}'$ sur lesquels q est strictement positive. Notons s le nombre de vecteurs de $\mathcal{E}$ sur lesquels q est strictement négative. Notons s' le nombre de vecteurs de $\mathcal{E}'$ sur lesquels q est strictement négative. On appelle J les indices de ces vecteurs :

$$J := \{i : q(\varepsilon'_i) < 0\}.$$

On va montrer que la concaténation des familles $\{\varepsilon_i\}_{i \in I}$ et $\{\varepsilon'_i\}_{i \in J}$ forme une famille libre. En effet soit $(\alpha_i)_{i \in \{1, \dots, r+s'\}}$ tel que

$$\sum_{i \in I} \alpha_i \varepsilon_i + \sum_{i \in J} \alpha_i \varepsilon'_i = 0.$$

On a donc

$$\sum_{i \in I} \alpha_i \varepsilon_i = - \sum_{i \in J} \alpha_i \varepsilon'_i.$$

Si on applique la forme quadratique aux deux membres de cette égalité on obtient

$$\sum_{i \in I} \alpha_i^2 q(\varepsilon_i) = \sum_{i \in J} (-\alpha_i)^2 q(\varepsilon'_i), \quad (5.1)$$

où on a utilisé que pour une famille orthogonale $\{\varepsilon_i\}_{i \in I}$,

$$q\left(\sum_{i \in I} \alpha_i \varepsilon_i\right) = \sum_{i \in I} \alpha_i^2 q(\varepsilon_i).$$

Par définition de I et J , le membre de gauche de (5.1) est strictement positif s'il existe un α_i non-nul dans $i \in \{1, \dots, r\}$ alors que son membre de droite est strictement négatif s'il existe un α_i non-nul dans $i \in \{r+1, \dots, r+s\}$. Comme le cardinal d'une famille libre est toujours inférieur à la dimension de l'espace vectoriel on a $\text{Card}(I) + \text{Card}(J) \leq N$. D'où $r + (N - r') \leq N$ ce qui implique $r \leq r'$.

En échangeant le rôle de ε_i et ε'_i on obtient $r' \leq r$ et donc $r = r'$. La démonstration de $s = s'$ est une répétition de ces arguments. $\square$

Exemple 64. *Quelle est la signature de f dans l'exemple 46 ?*

▷ On a prouvé

$$f(x, y, z) = (x + 2y - 3z)^2 - 2(y^2 - 2z)^2 - z^2.$$

donc $\sigma(f) = \{1, 2\}$.

Exemple 65. *Quelle est la signature de f dans l'exemple 47 ?*

▷ On a prouvé

$$f(x, y, z, t) = \frac{(x - 4z - 2t)^2}{4} - \frac{(-4z - 6t)^2}{4} - (z + t)^2 - 7t^2.$$

donc $\sigma(f) = \{1, 3\}$.

Proposition 5.6 (► Signature et positivité d'une forme quadratique). *Soient E un espace vectoriel de dimension N et $q \in \mathcal{Q}(E)$. On a :*

- Si $\sigma_q = (\alpha, 0)$ avec $\alpha \in \{1, \dots, N\}$ alors q est semi-définie positive,
- Si $\sigma_q = (N, 0)$ alors q est définie positive,
- Si $\sigma_q = (0, \beta)$ avec $\beta \in \{1, \dots, N\}$ alors q est semi-définie négative,
- Si $\sigma_q = (N, 0)$ alors q est définie négative.

► *Démonstration.* • Si $\sigma_q = (\alpha, 0)$ alors

$$q(x) = \sum_{i=1}^p \alpha_i l_i^2(x)$$

avec $\alpha_i \geq 0$ donc q est semi-définie positive.

- Si $\sigma_q = (N, 0)$ alors

$$q(x) = \sum_{i=1}^N \alpha_i l_i^2(x)$$

avec $\alpha_i \geq 0$, alors d'une part q est semi-définie positive mais on a aussi

$$q(x) = 0 \Leftrightarrow \forall i \in \{1, \dots, N\}, l_i^2(x) = 0$$

ce qui implique que $x = 0$ car la méthode de Gauss assure que la famille $\{l_i\}_{i \in \{1, \dots, N\}}$ est échelonnée.

Les deux derniers points sont immédiats puisque q est définie négative si et seulement si $-q$ est semi-définie positive. $\square$

5.4 Applications

5.4.1 Application du théorème spectral à la diagonalisation d'une matrice symétrique

Ce théorème permet en effet de simplifier la pratique de la diagonalisation pour les matrices symétrique puisque le calcul de l'inverse sera réduit au calcul de la transposée :

Proposition 5.7 (► Diagonalisation d'une matrice symétrique). *Soit $A \in \mathcal{M}_N(\mathbb{R})$ une matrice symétrique. Il existe une matrice D diagonale et une matrice P orthogonale telle que*

$$A = PDP^T.$$

► *Démonstration.* D'après le théorème spectral, Théorème 5.1, A est diagonalisable dans une base orthonormée de vecteurs propres. Donc il existe D diagonale et P inversible telles que $A = PDP^{-1}$. Mais P est la matrice de passage de la base canonique à la base orthonormée de vecteurs propres, c'est donc une matrice orthogonale. Donc d'après Proposition 4.16, $P^{-1} = P^T$. $\square$

L'étudiant.e perspicace verra dans cette proposition le signe de ce passage entre le monde des applications linéaires, dans lesquelles le changement de base utilise l'inverse de la matrice de passage, et celui des formes bilinéaires dans lequel c'est la transposée de la matrice de passage qui intervient lors d'un changement de base.

5.4.2 Application du théorème de représentation de Riesz-Fréchet à la signature d'une forme quadratique

Signature : critère des valeurs propres

On considère une forme quadratique q et f l'endomorphisme associé par le théorème de représentation de Riesz-Fréchet, soit une base $(\varepsilon_i)_{i \in \{1, \dots, N\}}$ de vecteurs propres de A , on a pour tout x de coordonnées $(x_i)_{i \in \{1, \dots, N\}}$ dans la base $(\varepsilon_i)_{i \in \{1, \dots, N\}}$

$$\begin{aligned} q(x) = \langle f(x), x \rangle &= \left\langle f \left(\sum_{i=1}^N x_i \varepsilon_i \right), \sum_{i=1}^N x_i \varepsilon_i \right\rangle = \left\langle \sum_{i=1}^N x_i f(\varepsilon_i), \sum_{i=1}^N x_i \varepsilon_i \right\rangle \\ &= \left\langle \sum_{i=1}^N \lambda_i x_i \varepsilon_i, \sum_{i=1}^N x_i \varepsilon_i \right\rangle = \sum_{i=1}^N \lambda_i x_i^2. \end{aligned}$$

Comme la nombre de coefficient positifs et négatifs ne dépend pas de la décomposition choisie on voit que le signe des valeurs propres renseigne effectivement sur la signature de la forme quadratique. On obtient ainsi

Proposition 5.8 (► Signature : critère des valeurs propres). *Soit $A \in \mathcal{M}_N$. Considérons la forme quadratique q associée à A . Soit α le nombre de valeurs propres positives et β le nombre de valeurs propres négatives. On a $\sigma(f) = (\alpha, \beta)$.*

Cette méthode est relativement facile en petite dimension. En dimension 2 par exemple elle se révèle très aisée à utiliser. Par contre, comme nous l'avons déjà signalé dans la chapitre sur la diagonalisation, cette méthode peut se révéler extrêmement problématique en dimension plus grande que 5 (pour lesquelles on ne peut pas toujours déterminer les racines du polynôme caractéristique et donc les valeurs propres).

Exemple 66. *Considérons la forme quadratique*

$$\begin{aligned} q : \quad \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow x^2 + y^2 - xy \end{aligned}$$

Quelle est la signature de q ?

► *La représentation matricielle de q dans la base canonique est donnée par*

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

Le spectre de cette matrice est $\{0, 2\}$. Donc $\sigma(q) = \{1, 0\}$.

Cette méthode fonctionne bien en dimension 2. On a même le résultat suivant, qui est spécifique à la dimension 2 et donc largement inutile mais qui servira plus que de nécessaire dans les bébés exemples que l'on prend en économie par ex.

Proposition 5.9 (► Critère des valeurs propres en dimension 2). *Soient $A \in \mathcal{M}_2$ et q sa forme quadratique associée.*

- La forme quadratique q est définie positive si et seulement si $\text{Tr}(A) > 0$ et $\det(A) > 0$.
- La forme quadratique q est définie négative si et seulement si $\text{Tr}(A) < 0$ et $\det(A) > 0$.

Ce résultat tient au fait que, par le théorème spectral, toute matrice symétrique est diagonalisable et donc qu'il existe une base de vecteurs propres dans laquelle la représentation matricielle de f , l'application linéaire associée à A , est $\text{diag}\{\lambda_1, \lambda_2\}$. Or la trace et le déterminant sont invariants par changement de base donc $\text{Tr}(A) = \lambda_1 + \lambda_2$ et $\det(A) = \lambda_1 \lambda_2$. Si le déterminant est positif cela assure que les valeurs propres sont de même signe, la trace permet de déterminer la signe de ces valeurs propres et de conclure.

Attention ce critère ne marche qu'en dimension 2 ! Le contre-exemple suivant l'illustre facilement :

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

qui est de déterminant 3 et de trace 1 sans que les valeurs propres ne soient toutes positives.

Exemple 67. Déterminer la signature de

$$q : \quad \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto x^2 + 2xy + 2y^2$$

1. En utilisant la réduction en carrés de Gauss.
2. En étudiant le signe des valeurs propres.

▷

1. La réduction en carré de Gauss est $q(x, y) = (x + y)^2 + 2y^2$. Donc $\sigma_q = (2, 0)$.
2. On commence par traduire l'expression en q en remplaçant x par x_1 et y par x_2 :

$$q : \quad \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto x_1^2 + 2x_1x_2 + 2x_2^2$$

On peut alors appliquer Proposition 3.22, pour déterminer que la forme polaire associée à q est

$$b : \quad \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x_1, x_2), (y_1, y_2) \mapsto x_1x_2 + x_1y_2 + x_2y_1 + 2x_2y_2$$

D'après Définition 3.14 la représentation matricielle de q est

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

Le polynôme caractéristique de A est $\chi_A = X^2 - 3X + 1$ qui a pour discriminant $\Delta = 5$ et pour racine $(3 + \sqrt{5})/2$ et $(3 - \sqrt{5})/2$. Or $3^2 = 9 > 5$ donc les deux valeurs propres sont positives et la signature est bien $(2, 0)$.

La comparaison entre travail nécessaire à l'application de la méthode des valeurs propres et la méthode de réduction en carrés est sans appel. J'ai peut-être un peu exagéré le trait mais l'étudiant.e perspicace aura entendu le message que je voulais faire passer : il faut connaître la méthode de réduction en carré. Quitte pour appuyer mes propos à jouer avec l'impartialité. Mais n'est ce pas ce que vous êtes aussi censé.e.s apprendre : de savoir identifier ces stratégies qui permettra aux citoyen.ne.s que vous êtes d'être vigilant.e.s devant les dires des expert.e.s. Les mathématiques peuvent tout faire et servent à tout.

En dimensions supérieures il est toujours possible d'appliquer ce résultat mais, de façon générale, c'est bien plus complexe.

Exemple 68. Si on reprend l'exemple de l'exercice 46 et que l'on essaie d'appliquer la méthode des valeurs propres. On construit la matrice associée à f :

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

Son polynôme caractéristique est $\chi_A = -X^3 + 6X^2 + 6X - 4$. On ne voit pas de racine évidente et l'étudiant moyen peut rester perplexe. Mais l'étudiant diligent pourra remarquer que $\lim_{\lambda \rightarrow -\infty} \chi_A(\lambda) = +\infty$ et $\chi_A(0) = -4$, il y a donc un zéro entre $-\infty$ et 0 donc une valeur propre négative. On sait déjà que la forme quadratique n'est pas définie positive (ce qui peut être suffisant si on se demande juste si le point critique d'une fonction est un minimum par exemple). Si on veut aller plus loin on pourra voir que $\chi_A(1) = 7$ donc il y a une valeur propre positive entre 0 et 1. Et comme χ_A est de degré impair on sait qu'il tend vers $-\infty$ en $+\infty$ et donc qu'il y a une deuxième valeur propre positive. Donc $\sigma(A) = (2, 1)$. Méthode à comparer à celle de l'exercice 46.

On pourra, de façon plus générale utiliser la règle des signes de Descartes mais cela reste souvent plus compliqué en pratique que d'utiliser la méthode de réduction en carrés de Gauss, au moins en petite dimension. Numériquement, c'est cependant la méthode des valeurs propres qui est la plus utilisée, parce que des algorithmes numériques rapides utilisent la symétrie de la matrice.

Signature : Méthode des mineurs principaux

Cette méthode est efficace pour déterminer lorsqu'une forme quadratique est définie positive mais elle ne permet même pas de déterminer la signature d'une forme quadratique en général. Nous indiquons néanmoins cette recette de cuisine parce qu'elle est utilisée massivement chez mes collègues et qu'elle marche pour les quelques cas hypers spécifiques de formes quadratiques définies positives sur $\mathbb{R}^n$ que l'on rencontre abondamment pour les exercices bêtement académiques.

Proposition 5.10 (► Signature : Méthode des mineurs principaux). Soit $A \in \mathcal{M}_N$ la matrice de terme général $[a_{ij}]_{(i,j) \in \{1, \dots, N\}^2}$. Pour tout $p \in \{1, \dots, N\}$, on note A_p la matrice de $\mathcal{M}_p$ de terme général $[a_{ij}]_{(i,j) \in \{1, \dots, p\}^2}$. La matrice A est définie positive si et seulement si tous les mineurs principaux : $\det(A_p)$ sont strictement positifs.

L'adaptation de ce résultat à la détermination du caractère définie négatif d'une matrice est déjà en lui-même compliqué : A est définie négative si tous les mineurs principaux A_p pour p pair sont positif et tous ceux pour p impairs sont négatifs. Ceci vient du fait que l'on peut normaliser une forme quadratique, voir Proposition 3.25, une matrice définie négative ayant que des termes -1 sur la diagonale ceci permet de retrouver le résultat.

L'extension de ce résultat à la détermination de la signature d'une forme quadratique est plus complexe et marque les limites de ce résultat.

► *Démonstration de Proposition 5.10.* Soit q la forme quadratique associée à A et b sa forme polaire.

($\Rightarrow$) Si q est définie positive alors nous avons vu Proposition 3.25 qu'il existe une base b -orthonormée dans laquelle la représentation matricielle de b est normalisée avec des 1 sur la diagonale, donc l'identité. D'après la loi d'intertie de Sylvester, la représentation matricielle de q est l'identité dans n'importe quelle base b -orthonormée. D'après la formule de changement de base, Proposition 3.16, on a donc

$$A = P^T I P$$

où P est la matrice de passage de la base b -orthonormée à la base canonique. Ceci implique que $\det(A) = \det(P)^2 > 0$.

Soit $\text{Vect}\{(e_i)_{i \in \{1, \dots, N\}}\}$ la base canonique. On applique alors le même raisonnement au sous-espace vectoriel $E_k = \text{Vect}\{(e_i)_{i \in \{1, \dots, N\}}\}$. La représentation matricielle de la

restriction de q à E_k est la matrice A_p de terme général $[a_{ij}]_{(i,j) \in \{1, \dots, p\}^2}$. Pour cette matrice, comme précédemment en utilisant le Proposition 3.25, il existe P_k telle que

$$A_k = P_k^T I P_k.$$

D'où l'on déduit $\det(A_k) = \det(P_k)^2 > 0$.

($\Leftarrow$) Soit $\mathcal{P}$ la propriété $\mathcal{P}(n)$: "si E_n est un espace de dimension n alors si A est définie positive sur E_n alors tous les mineurs principaux de A sont positifs".

La propriété $\mathcal{P}(1)$ est évidemment vraie.

Supposons que la propriété $\mathcal{P}$ soit vraie jusqu'au rang $n - 1$. Soit un espace E_n de dimension n et A une matrice dont tous les mineurs principaux sont strictement positifs. Soit q la forme quadratique associée à A et b sa forme polaire. On considère un hyperplan H de E_n . Comme H est de dimension $n - 1$ l'hypothèse de récurrence permet d'affirmer que la restriction de q à H , $q|_H$, est définie positive sur H . Donc d'après Proposition 3.25, il existe une base orthonormée $(\varepsilon_i)_{i \in \{1, \dots, n-1\}}$ pour la forme polaire associée à $q|_H$.

D'après Théorème 4.21, $H \oplus H^\perp = E_n$, donc $H^\perp$ est de dimension 1. Soit ε_n un vecteur de $H^\perp$ de norme 1. La famille $(\varepsilon_i)_{i \in \{1, \dots, n\}}$ est une base de E_n et la représentation matricielle de q dans cette base est

$$K = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ 0 & \dots & 0 & q(\varepsilon_n) \end{pmatrix}$$

Par hypothèse $\det(K) > 0$. Or $\det(K) = q(\varepsilon_n)$ donc $q(\varepsilon_n) > 0$. Ainsi q (et A) est définie positive sur E_n et $\mathcal{P}$ est héréditaire. Donc $\mathcal{P}$ est vraie pour tout n ce qui termine la preuve. $\square$

Exemple 69. En dimension 2, considérons la forme quadratique

$$q : \mathbb{R}^2 \rightarrow \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto ax^2 + 2bxy + cy^2$$

La représentation matricielle de q dans la base canonique est

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

La décomposition en carrés de Gauss si $a \neq 0$ est

$$q(x, y) = a \left(x + \frac{b}{a} y \right)^2 + \frac{1}{a} (ac - b^2) y^2$$

Cette forme quadratique est définie positive si et seulement si $\det(A_1) = a > 0$ et $\det(A) = ac - b^2 > 0$.

5.5 Exercices

Exercice 5.1. Déterminer D et P telles que $A = PDP^T$ pour les matrices suivantes :

1.

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

2.

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

3.

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

4.

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

Exercice 5.2. Soit le système défini par

$$\begin{cases} x_{n+1} = 2x_n + 2y_n \\ y_{n+1} = 2x_n + 2z_n \\ z_{n+1} = 2y_n + 2z_n \end{cases}$$

où x_0 , y_0 et z_0 sont donnés. Déterminer x_n en fonction de n , x_0 , y_0 et z_0 .

5.6 Corrigés des exercices

Corrigé Exercice 5.1. 1. On calcule $\chi_A = -(X+2)(X-2)(X-4)$ et $E_{-2} = \text{Vect}\{(1, -2, 1)\}$, $E_2 = \text{Vect}\{(-1, 0, 1)\}$, $E_4 = \text{Vect}\{(1, 1, 1)\}$. La base $\{(1, -2, 1), (-1, 0, 1), (1, 1, 1)\}$ est une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\left\{ \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{-2, 2, 4\}$ et

$$\begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$$

2. On calcule $\chi_A = -X^2(X-3)$ et $E_0 = \text{Vect}\{(-1, 0, 1), (-1, 1, 0)\}$, $E_3 = \text{Vect}\{(1, 1, 1)\}$. La base de E_0 n'est pas orthogonale. On l'orthogonalise en utilisant le procédé de Gram-Schmidt pour obtenir une base orthogonale de E_0 : $\{(-1, 0, 1), (-1, 2, -1)\}$.

La base $\{(-1, 0, 1), (-1, 2, -1), (1, 1, 1)\}$ est bien une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\left\{ \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{0, 0, 3\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

3. On calcule $\chi_A = -(X+3)^2(X-3)$ et $E_{-3} = \text{Vect}\{(-1, 0, 1), (2, 1, 0)\}$, $E_3 = \text{Vect}\{(1, -2, 1)\}$. La base de E_{-3} n'est pas orthogonale. On l'orthogonalise en utilisant le procédé de Gram-Schmidt pour obtenir une base orthogonale de E_{-3} : $\{(-1, 0, 1), (1, 1, 1)\}$.

La base $\{(-1, 0, 1), (1, 1, 1), (1, -2, 1)\}$ est bien une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\left\{ \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix} \right\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{0, 0, 3\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{pmatrix}$$

4. On calcule $\chi_A = -(X+1)X(X-3)$ et $E_{-1} = \text{Vect}\{(-1, 0, 1)\}$, $E_0 = \text{Vect}\{(0, 1, 0)\}$, $E_3 = \text{Vect}\{(3, 4, 3)\}$. La base $\{(-1, 0, 1), (0, 1, 0), (3, 4, 3)\}$ est une base orthogonale de vecteurs propres.

On normalise cette base pour obtenir la base orthonormée

$$\left\{ \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3/\sqrt{34} \\ 4/\sqrt{34} \\ 3/\sqrt{34} \end{pmatrix} \right\}.$$

On a donc $A = PDP^T$ avec $D = \text{diag}\{-2, 2, 4\}$ et

$$\begin{pmatrix} -1/\sqrt{2} & 0 & 3/\sqrt{34} \\ 0 & 1 & 4/\sqrt{34} \\ 1/\sqrt{2} & 0 & 3/\sqrt{34} \end{pmatrix}$$

Corrigé Exercice 5.2. On utilise la méthode habituelle et $A^n = PD^nP^{-1}$. Comme la matrice est symétrique on sait que A est diagonalisable et que l'on peut diagonaliser A dans une base orthonormée de vecteurs propres. C'est ce que nous avons fait dans l'exercice ci-dessus. À dire vrai entre le temps que l'on gagne dans le calcul de P^{-1} et le risque d'erreur dans le calcul de $A^n = PD^nP^T$, il n'est pas clair que la normalisation apporte grand chose.