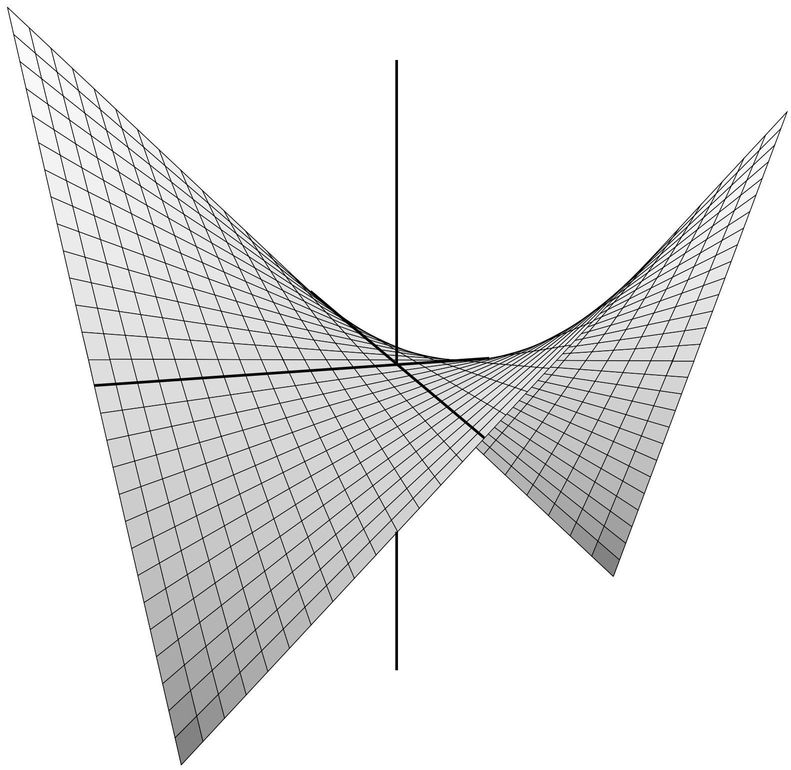
Algèbre linéaire et bilinéaire I

2MA221

Yves Coudène, 1 septembre 2024

Licence de mathématiques, Sorbonne Université

Version 4.2



2024 - 2025

Table des matières

Introduction

Mais malheur à l'auteur qui veut toujours instruire!
Le secret d'ennuyer est celui de tout dire.
Voltaire (1694–1778)

Ces notes accompagnent le cours 2MA221 *Algèbre linéaire et bilinéaire I* donné au premier semestre des années 2020-2023 à Sorbonne Université. Le site internet se trouve à l'adresse

`https://perso.lpsm.paris/~coudene/2MA221-algebre-lineaire-bilineaire-I.html`

Ce cours d'algèbre linéaire suppose connu les notions d'espace vectoriel, de base, d'application linéaire et de matrice ainsi qu'une familiarité avec les notions de déterminants et de valeurs propres. Il a pour objet l'étude des formes quadratiques, des espaces euclidiens et la diagonalisation des applications linéaires. Trois points de vue sont adoptés dans ce texte.

– Les algorithmes permettent un calcul effectif des objets étudiés,

– le point de vue géométrique permet de les visualiser,

– l'algèbre en donne une représentation abstraite simplifiant leur étude.

Les deux résultats principaux de ce texte sont des théorèmes de réduction et de diagonalisation, ils seront abordés sous ces trois points de vue.

Le premier chapitre contient des rappels d'algèbre linéaire. On introduit la notion de matrice échelonnée et on explique comment échelonner une matrice quelconque par l'algorithme du pivot de Gauss. Ce procédé permet entre autre de calculer le noyau et le rang des matrices et de résoudre les systèmes linéaire associés. Le point clef du chapitre est bien sûr l'algorithme du pivot de Gauss.

Le second chapitre porte sur les formes quadratiques. On montre que toute forme quadratique peut s'écrire sous forme de somme de carrés de formes linéaires, grâce à un algorithme de réduction là encore dû à Carl Friedrich Gauss (1777-1855). La notion de signature permet de les classer à équivalence près.

Le troisième chapitre est de nature plus géométrique. L'étude des espaces euclidiens est illustrée par quelques résultats de géométrie plane. On fait le lien entre les notions de produit scalaire, d'aire, de longueur et d'angle et on donne une application à l'étude des coniques du plan et des quadriques de l'espace.

Le dernier chapitre est consacré à l'étude des applications linéaires définies sur un espace vectoriel de dimension finie. Après quelques résultats généraux concernant leur diagonalisation, on fait une étude plus détaillée des applications autoadjointes dans le cadre euclidien.

Ce texte couvre l'intégralité du cours 2MA221. Il est complété par des feuilles d'exercices qui permettent une assimilation progressive de son contenu. La liste des méthodes développées dans ce cours est donnée en annexe. Une bonne connaissance de ces méthodes est exigée à l'examen. On ne saurait aussi trop insister sur la nécessité de s'entraîner à calculer tout au long du semestre pour bien réussir cet examen.

Yves Coudène, le 29 août 2024.

Chapitre 1

Rappels d'algèbre linéaire

On rappelle dans ce chapitre quelques notions d'algèbre linéaire vues en première année de licence. Le corps de base est celui des nombres réels ou complexes.

Le résultat clef est l'algorithme du pivot de Gauss qui permet de mettre les matrices sous forme échelonnée. Une fois sous forme échelonnée, on calcule facilement le rang de la matrice, son noyau ou encore les solutions des systèmes linéaires associés. C'est également cet algorithme qui permet de passer d'une représentation cartésienne d'un sous-espace vectoriel de $\mathbf{R}^n$ à une représentation paramétrique et vice-versa.

1. Généralités sur les matrices

1.1 Définition d'une matrice

Rappelons qu'une matrice A de taille $m \times n$ est une famille de nombres réels ou complexes indexée par deux indices i et j compris respectivement entre 1 et m et entre 1 et n. Elle possède m lignes et n colonnes, on la note

$$A = \{a_{i,j}\}_{\substack{i=1..m \\ j=1..n}}$$

et on la représente sous la forme d'un tableau

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}.$$

L'ensemble des matrices de taille $n \times n$ est noté $M_n(\mathbf{R})$. À toute matrice A de taille $m \times n$ correspond une application linéaire définie de $\mathbf{R}^n$ dans $\mathbf{R}^m$ qui associe au vecteur $x = (x_1, \dots, x_n)$ le vecteur $y = (y_1, \dots, y_m)$ donné par les formules

$$y_i = \sum_{j=1}^{n} a_{i,j} x_j, \quad i \text{ allant de 1 à } m.$$

On utilise l'expression abrégée $y = Ax$. Il est d'usage de représenter x et y par des vecteurs colonnes

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

et de représenter les égalités précédentes sous la forme

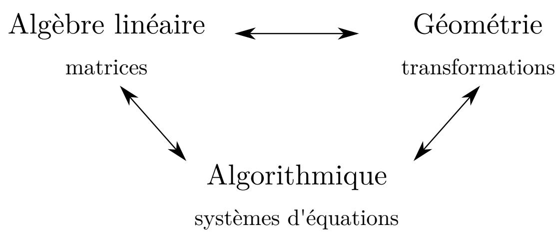
$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

À une matrice donnée $A = \{a_{i,j}\}$ est naturellement associé un système d'équations linéaires

$$\left\{ \begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & y_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & y_m \end{array} \right.$$

C'est le système d'équations qui exprime les coordonnées du vecteur y en fonction de celles de x.

On voit qu'il y a trois types d'objets qui se correspondent : les matrices, les applications linéaires et les systèmes d'équations linéaires. Chaque objet correspond à un point de vue différent.



– En algèbre linéaire, on additionne et on fait des produits de matrices.

– En géométrie, on compose des transformations.

– En algorithmique, on manipule des systèmes d'équations.

Suivant les objectifs qu'on se fixe, on est amené à adopter l'un ou l'autre de ces points de vue. Il est donc important de savoir passer de l'un à l'autre. Par exemple, il faut être capable de traduire un énoncé d'algèbre linéaire en termes géométriques ou de ramener un problème de géométrie à la résolution d'un système d'équations.

1.2 Rang d'une matrice

Rappelons quelques résultats concernant le rang d'une matrice.

Définition 1 *Le* noyau *d'une matrice A de taille $m \times n$ est constitué des vecteurs $x \in \mathbf{R}^n$ pour lesquels $Ax = 0$.*

$$\ker(A) = \{x \in \mathbf{R}^n \mid Ax = 0\}.$$

*L'*image *de A est constituée des vecteurs de $\mathbf{R}^m$ de la forme Ax, $x \in \mathbf{R}^n$.*

$$\mathrm{im}(A) = \{Ax \in \mathbf{R}^m \mid x \in \mathbf{R}^n\}.$$

Le rang *de A est égal à la dimension de l'image de A.*

$$\mathrm{rang}(A) = dim\ \mathrm{im}(A).$$

L'image d'une matrice est égale à l'espace vectoriel engendré par ses colonnes. Le rang est égal à la dimension de cet espace.

Le noyau et l'image d'une matrice sont des espaces vectoriels. Le rang d'une matrice est un entier qui est nul si et seulement si tous les coefficients de la matrice sont nuls.

La proposition suivante montre que la somme du rang d'une matrice et de la dimension de son noyau est égale à la dimension de l'espace sur lequel est définie la matrice.

Proposition 1 (formule du rang) *Soit A une matrice de taille $m \times n$.*

$$\mathrm{rang}(A) + dim\ \ker(A) = n.$$

2. Matrices échelonnées

2.1 Définition d'une matrice échelonnée

Les matrices échelonnées forment une classe de matrices simples à étudier.

Définition 2 *Une matrice est dite* échelonnée *en lignes si le nombre de zéros consécutifs au début de chaque ligne augmente strictement de lignes en lignes jusqu'à ce qu'il ne reste éventuellement plus que des zéros.*

Les *pivots* sont les premiers termes non nuls de chaque ligne.

Exemples
Les matrices suivantes sont échelonnées, les pivots sont en gras :

$$\begin{pmatrix} 0 & \mathbf{3} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 2 & 3 \\ 0 & \mathbf{1} & 5 \end{pmatrix}, \begin{pmatrix} \mathbf{2} & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 2 & 0 \\ 0 & 0 & \mathbf{2} \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 4 & 0 & 0 & 5 \\ 0 & 0 & \mathbf{1} & 1 & 4 \\ 0 & 0 & 0 & \mathbf{1} & 3 \end{pmatrix}.$$

Les matrices suivantes ne sont pas échelonnées :

$$\begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 5 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Remarque : rappelons qu'une matrice est carrée si elle a le même nombre de lignes que de colonnes : $m = n$. Les matrices carrées échelonnées sont triangulaires : $a_{i,j} = 0$ si $i > j$.

Définition 3 *Une matrice échelonnée est* échelonnée réduite *si :*
– ses pivots sont tous égaux à 1,
– dans une colonne contenant un pivot, tous les coefficients sont nuls sauf le pivot.

Exemple
Les matrices suivantes sont échelonnées réduites :

$$\begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 0 & 5 \\ 0 & \mathbf{1} & 2 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 3 & 0 \\ 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \mathbf{1} & 4 & 0 & 0 & 5 \\ 0 & 0 & \mathbf{1} & 0 & 4 \\ 0 & 0 & 0 & \mathbf{1} & 3 \end{pmatrix}.$$

Les matrices suivantes ne sont pas sous forme échelonnée réduite :

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 5 \\ 0 & 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.2 Propriétés

Le rang d'une matrice échelonnée est facile à calculer.

Proposition 2 *Le rang d'une matrice échelonnée est égal à son nombre de pivots. Il est aussi égal au nombre de lignes non identiquement nulles.*
La dimension du noyau d'une matrice échelonnée est égale au nombre de colonnes ne contenant pas de pivots.

Les systèmes linéaires associés à des matrices échelonnées réduites sont aisés à résoudre.

Méthode : *résolution des systèmes linéaires échelonnés réduits*

Soit A une matrice échelonnée réduite. On cherche à résoudre le système décrit plus haut : le vecteur y étant donné, on cherche tous les x pour lesquels $Ax = y$. Les variables $x_1, \dots, x_n$ situées à l'emplacement des pivots sont appelées *variables principales.*

– On introduit un paramètre pour chaque variable non principale.

– On exprime les variables principales en fonction de ces paramètres.

On obtient alors une équation paramétrique pour l'ensemble des solutions du système.

Exemple
Les réels y_1, y_2 étant deux nombres réels donnés, cherchons les solutions $(x_1,x_2,x_3) \in \mathbf{R}^3$ du système

$$\begin{cases} x_1 & + & & + & 3x_3 & = & y_1 \\ & & x_2 & + & x_3 & = & y_2 \end{cases}$$

Il est associé à la matrice échelonnée réduite $\left(\begin{smallmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \end{smallmatrix}\right)$.

– Les variables principales sont x_1 et x_2. Il existe une seule variable non principale, x_3. On introduit donc un paramètre λ et on pose $x_3 = \lambda$.

– On exprime les variables principales x_1 et x_2 en fonction de ce paramètre.

$$\begin{cases} x_1 & = & y_1 & - & 3x_3 & = & y_1 - 3\lambda \\ x_2 & = & y_2 & - & x_3 & = & y_2 - \lambda \end{cases}$$

On obtient l'équation paramétrique désirée :

$$\begin{cases} x_1 & = & y_1 & - & 3\lambda \\ x_2 & = & y_2 & - & \lambda \\ x_3 & = & & & \lambda \end{cases}$$

En notation vectorielle,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}.$$

Pour toute valeur de y_1 et y_2, le système possède une infinité de solutions. L'ensemble des solutions est une droite affine dirigée par le vecteur $(-3,-1,1)$ et passant par le point de coordonnées $(y_1,y_2,0)$.

On en déduit le noyau de A en prenant $y_1 = 0$ et $y_2 = 0$:

$$\ker(A) = \{\lambda(-3,-1,1) \mid \lambda \in \mathbf{R}\} = \operatorname{vect}\left(\left(\begin{smallmatrix} -3 \\ -1 \\ 1 \end{smallmatrix}\right)\right).$$

C'est la droite vectorielle engendrée par le vecteur de coordonnées $(-3,-1,1)$. La dimension du noyau est égale à 1, elle est bien égale au nombre de colonnes de A sans pivots : la troisième colonne de A est la seule sans pivot.

Dans le cas échelonné réduit, on peut aussi trouver une base du noyau en exprimant chaque colonne ne contenant pas de pivot en fonction des colonnes la précédant et contenant des pivots, c'est-à-dire en fonction des éléments de la base canonique. Dans l'exemple précédent, la troisième colonne ne contient pas de pivot, elle est égale à trois fois la première plus la seconde.

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

ou encore

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

3. Algorithme du pivot de Gauss

3.1 Mise sous forme échelonnée

L'algorithme de Gauss (1777-1855) permet de placer une matrice sous forme échelonnée réduite. Il ne modifie pas l'ensemble des solutions du système linéaire associé à la matrice, il permet donc de ramener la résolution d'un système linéaire quelconque à un système échelonné.

Méthode : *mise sous forme échelonnée d'une matrice*

On cherche à transformer une matrice de taille $m \times n$ en une matrice échelonnée. L'algorithme de Gauss opère sur les lignes de la matrice.

– Si le coefficient $a_{1,1}$ est non nul, on soustrait un multiple de la première ligne à la seconde ligne afin d'annuler $a_{2,1}$ puis on répète cette opération pour les lignes suivantes.

– S'il existe une ligne dont le premier coefficient est non nul, on la permute avec la première ligne pour se ramener au cas précédent.

Après ces opérations, seul le premier coefficient de la première colonne peut être non nul. On répète le même procédé pour le bloc obtenu en supprimant la première ligne et la première colonne de la matrice, et ainsi de suite jusqu'à traiter toutes les colonnes. La matrice est alors sous forme échelonnée.

Exemple
Appliquons l'algorithme précédent à la matrice

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 2 & 8 \\ 3 & 3 & 12 \\ 1 & 2 & 5 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & 4 \\ 0 & 6 & 6 \\ 0 & 3 & 3 \end{pmatrix} \quad \begin{array}{l} \\ L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1 \\ L_4 \leftarrow L_4 - L_1 \end{array}$$

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{l} \\ \\ L_3 \leftarrow L_3 - \frac{3}{2}L_2 \\ L_4 \leftarrow L_4 - \frac{3}{4}L_2 \end{array} \qquad \text{Forme échelonnée}$$

Méthode : *mise sous forme échelonnée réduite d'une matrice*

On cherche à transformer une matrice échelonnée en une matrice échelonnée réduite. Pour cela, on effectue les opérations suivantes sur les lignes de la matrice.

– Pour chaque ligne contenant un pivot, on multiplie tous les coefficients de la ligne par l'inverse du pivot. Tous les pivots sont maintenant égaux à 1.

– On se place ensuite au niveau du second pivot. On ajoute aux lignes au dessus du pivot un multiple de la ligne contenant le pivot afin d'annuler tous les coefficients au dessus du pivot.

– On répète l'opération précédente pour tous les pivots, de manière ordonnée, en traitant les pivots de gauche à droite.

Exemple
On reprend l'exemple précédent, avec la matrice

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

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad L_2 \leftarrow \tfrac{1}{4}L_2$$

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad L_1 \leftarrow L_1 + L_2 \qquad \text{Forme échelonnée réduite}$$

3.2 Résolution de systèmes linéaires

L'algorithme précédent fait appel aux opérations suivantes :

– ajouter à une ligne un multiple d'une autre ligne,

– permuter deux lignes,

– multiplier les coefficients d'une ligne par le même facteur.

Ces opérations ne modifient pas les solutions du système linéaire associé à la matrice. Il permet donc de ramener sa résolution à celle d'un système échelonné réduit.

Exemple
On résout le système linéaire associé à la matrice donnée dans les exemples précédents. Soit y_1, y_2, y_3, $y_4 \in \mathbf{R}^4$, on cherche les vecteurs (x_1,x_2,x_3) de $\mathbf{R}^3$ satisfaisant le système

$$\left\{\begin{array}{ccccccccccc} x_1 & - & x_2 & + & 2x_3 & = & y_1 & & & & \\ 2x_1 & + & 2x_2 & + & 8x_3 & = & & & y_2 & & \\ 3x_1 & + & 3x_2 & + & 12x_3 & = & & & & y_3 & \\ 1x_1 & + & 2x_2 & + & 5x_3 & = & & & & & y_4 \end{array}\right.$$

En appliquant les mêmes opérations sur les lignes que dans l'exemple précédent, on obtient le système équivalent suivant.

$$\left\{\begin{array}{ccccccccccccc} x_1 & & & + & 3x_3 & = & \frac{1}{2}\,y_1 & + & \frac{1}{4}\,y_2 & & & & \\ & & x_2 & + & x_3 & = & -\frac{1}{2}\,y_1 & + & \frac{1}{4}\,y_2 & & & & \\ & & & & 0 & = & & - & \frac{3}{2}\,y_2 & + & y_3 & & \\ & & & & 0 & = & \frac{1}{2}\,y_1 & - & \frac{3}{4}\,y_2 & & & + & y_4 \end{array}\right.$$

Il n'y a de solutions que si $-\frac{3}{2}\,y_2 + y_3 = 0$ et $\frac{1}{2}\,y_1 - \frac{3}{4}\,y_2 + y_4 = 0$.

Les solutions sont alors de la forme

$$\left\{\begin{array}{ccccccc} x_1 & = & \frac{1}{2}\,y_1 & + & \frac{1}{4}\,y_2 & - & 3\lambda \\ x_2 & = & -\frac{1}{2}\,y_1 & + & \frac{1}{4}\,y_2 & - & \lambda \\ x_3 & = & & & & & \lambda \end{array}\right.$$

pour $\lambda \in \mathbf{R}$. Il y a une infinité de solutions, une pour chaque valeur de λ.

Pour éviter les erreurs de calcul, il est utile de garder les variables chacune dans sa colonne.

Méthode : *résolution des systèmes linéaires généraux*

On peut procéder comme suit pour alléger la résolution d'un système linéaire de la forme $Ax = y$: on concatène la matrice identité à droite de la matrice A. Dans l'exemple précédent, cela produit la matrice

$$\begin{pmatrix} 1 & -1 & 2 & | & 1 & 0 & 0 & 0 \\ 2 & 2 & 8 & | & 0 & 1 & 0 & 0 \\ 3 & 3 & 12 & | & 0 & 0 & 1 & 0 \\ 1 & 2 & 5 & | & 0 & 0 & 0 & 1 \end{pmatrix}$$

On applique alors l'algorithme de Gauss à cette matrice. On obtient

$$\begin{pmatrix} 1 & 0 & 3 & | & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & 1 & | & -\frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & | & 0 & -\frac{3}{2} & 1 & 0 \\ 0 & 0 & 0 & | & \frac{1}{2} & -\frac{3}{4} & 0 & 1 \end{pmatrix}$$

Les coefficients de cette matrice sont ceux du système réduit associé.

3.3 Interprétation matricielle

Afin d'étudier l'algorithme de Gauss plus en détail, notons une matrice A dont les lignes sont données par les vecteurs lignes $L_1, \ldots, L_n$ comme suit :

$$A = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}$$

Ajouter un multiple de la première ligne à la seconde ligne revient à multiplier la matrice A à gauche par la matrice obtenue en remplaçant le coefficient d'indice 2,1 de l'identité par le coefficient multiplicateur désiré :

$$\begin{pmatrix} L_1 \\ L_2 + \lambda L_1 \\ \vdots \\ L_m \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots \\ \lambda & 1 & 0 & \dots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix}.$$

Échanger les deux premières lignes revient à multiplier la matrice A à gauche par la matrice qui permute les deux premiers vecteurs de la base canonique :

$$\begin{pmatrix} L_2 \\ L_1 \\ \vdots \\ L_m \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix}.$$

Les autres opérations sur les lignes qui interviennent dans l'algorithme se traduisent toutes sous la forme d'un produit à gauche par une matrice inversible. Ceci implique le théorème suivant.

Théorème 1 (forme échelonnée d'une matrice) *Soit A une matrice de taille $m \times n$. Alors il existe une matrice P de taille $m \times m$ inversible et une matrice échelonnée réduite R de taille $m \times n$ telles que*

$$R = PA.$$

Pour calculer la matrice P, on procède comme suit.

– On concatène l'identité à la matrice A par la droite.

– On applique l'algorithme de Gauss à cette nouvelle matrice.

Ceci revient à multiplier à gauche notre matrice $(A \mid id)$ par la matrice P vue plus haut. La matrice obtenue vaut donc $P\,(A \mid id) = (PA \mid P) = (R \mid P)$ ce qui donne R et P.

Exemple
Le calcul de l'exemple précédent donne

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 2 & 8 \\ 3 & 3 & 12 \\ 1 & 2 & 5 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & -\frac{3}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{3}{4} & 0 & 1 \end{pmatrix}.$$

La matrice P est parfois qualifiée de pseudo-inverse. De fait, si A est une matrice carrée inversible, sa forme échelonnée réduite est égale à l'identité et

P est l'inverse de A. Ceci donne une méthode efficace pour calculer l'inverse d'une matrice carrée.

Méthode : *inversion d'une matrice*

Soit A une matrice carrée ($m = n$) à inverser.

– On concatène la matrice identité à droite pour obtenir la matrice $(A \mid id)$.

– On met $(A \mid id)$ sous forme échelonnée réduite par l'algorithme de Gauss.

Si la matrice A est inversible, l'algorithme produira la matrice $(id \mid A^{-1})$.
Si la matrice A n'est pas inversible, la partie gauche de la matrice donnée par l'algorithme sera échelonnée mais avec certains termes diagonaux nuls.

Voici quelques corollaires qui découlent des considérations précédentes.

Corollaire 1
– *Le noyau d'une matrice est égal au noyau de sa forme échelonnée réduite.*
– *Le rang d'une matrice est égal au rang de sa forme échelonnée réduite.*

Attention, l'image d'une matrice n'est pas égale à l'image de sa forme échelonnée réduite. On peut malgré tout se servir de la forme échelonnée pour obtenir une base de l'image de la matrice.

Corollaire 2 *Une base de* $\mathrm{im}(A)$ *est formée des colonnes de A situées aux mêmes positions que les colonnes de R contenant un pivot.*

Méthode : *calcul d'une base du noyau ou de l'image d'une matrice*

Pour le calcul du noyau, il suffit de calculer le noyau de sa forme échelonnée réduite. Pour l'image, on sélectionne les colonnes de la matrice en fonction de la position des pivots dans la forme échelonnée réduite.

Exemple
Revenons sur la matrice A de l'exemple précédent. Sa forme échelonnée réduite R a deux pivots, les matrices R et A sont donc de rang 2. Le noyau de A est égal au noyau de R, que nous avons déjà calculé.

$$\ker(A) = \ker(R) = \mathrm{vect}\left(\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}\right).$$

Les colonnes de R qui contiennent des pivots sont les colonnes 1 et 2. Les colonnes 1 et 2 de A forment donc une base de $\mathrm{im}(A)$.

$$\mathrm{im}(A) = \mathrm{vect}\left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 3 \\ 2 \end{pmatrix}\right).$$

3.4 Applications

Mettre une matrice sous forme échelonnée permet de :

– calculer le rang d'une matrice,
il est égal au nombre de pivots de sa forme échelonnée

– calculer la dimension du noyau d'une matrice,
il est égal au nombre de colonnes de la forme échelonnée sans pivots

– calculer le noyau d'une matrice,
il est égal au noyau de sa forme échelonnée

– calculer une base de l'image d'une matrice,
prendre les colonnes de la matrice associées aux pivots de la forme réduite

– déterminer le rang d'une famille de vecteurs de $\mathbf{R}^n$,
calculer le rang de la matrice dont les colonnes sont les vecteurs donnés

– déterminer si une famille de vecteurs de $\mathbf{R}^n$ est libre,
le rang de la famille doit être égal à son nombre de vecteurs

– résoudre un système d'équations linéaires avec second membre,
mettre le système sous forme échelonnée réduite

– calculer l'inverse d'une matrice carrée,
c'est équivalent à la résolution d'un système linéaire

– passer d'une représentation cartésienne d'un sous-espace vectoriel de $\mathbf{R}^n$ à une représentation paramétrique de cet espace et vice-versa.

Détaillons ce dernier point. Un sous-espace vectoriel E de $\mathbf{R}^n$ peut se représenter par un système d'équations cartésiennes ou par un système d'équations paramétriques.

Équations cartésiennes
Dans le premier cas, on dispose de m équations linéaires pour notre sous-espace E.

$$\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & 0 \\ \vdots & & & & \vdots & & \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & 0 \end{array}\right.$$

Un vecteur de $\mathbf{R}^n$ est dans E si et seulement si ses coordonnées satisfont toutes ces équations. Notons A la matrice donnée par les coefficients $\{a_{i,j}\}$. Un vecteur $x \in \mathbf{R}^n$ appartient à E si et seulement si $Ax = 0$. En d'autres termes,

$$E = \ker(A)$$

et la dimension de E est égale à la dimension du noyau de la matrice A.

Équations paramétriques
Dans le second cas, on dispose de k paramètres λ_1, ..., λ_k. Chaque valeur prise par ces paramètres donne les coordonnées d'un point du sous-espace vectoriel.

$$\begin{cases} x_1 & = & c_{1,1}\lambda_1 & + & c_{1,2}\lambda_2 & + & \dots & + & c_{1,k}\lambda_k \\ x_2 & = & c_{2,1}\lambda_1 & + & c_{2,2}\lambda_2 & + & \dots & + & c_{2,k}\lambda_k \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ x_n & = & c_{n,1}\lambda_1 & + & c_{n,2}\lambda_2 & + & \dots & + & c_{n,k}\lambda_k \end{cases}$$

Notons C la matrice donnée par les coefficients $\{c_{i,j}\}$. Un vecteur est dans E si et seulement si il est de la forme $C\lambda$, où λ est le vecteur de $\mathbf{R}^k$ de coordonnées $\lambda_1, \dots, \lambda_k$. En d'autres termes,

$$E = \mathrm{im}(C)$$

et la dimension de E est égale au rang de la matrice C.

Méthode : *calcul d'équations paramétriques pour un sous-espace de* $\mathbf{R}^n$

Pour passer d'un système d'équations cartésiennes à un système d'équations paramétriques, il suffit de résoudre le système linéaire définissant l'espace E en échelonnant la matrice associée. Lors de la résolution, on introduit autant de paramètres qu'il y a de variables non principales dans le système échelonné.

Exemple
Considérons le sous-espace de $\mathbf{R}^3$ défini par les équations

$$\begin{cases} x_1 & + & 3x_3 & = & 0 \\ x_2 & + & x_3 & = & 0 \end{cases}$$

C'est un cas particulier d'un système que nous avons résolu précédemment, les solutions sont de la forme

$$\begin{cases} x_1 & = & -3\lambda \\ x_2 & = & -\lambda \\ x_3 & = & \lambda \end{cases}$$

L'espace vectoriel considéré est de dimension un, il est engendré par le vecteur $(-3, -1{,}1)$, on parle de droite vectorielle.

Méthode : *calcul d'équations cartésiennes pour un sous-espace de* $\mathbf{R}^n$

Pour passer d'un système d'équations paramétriques à un système d'équations cartésiennes, on résout à nouveau le système en considérant les λ_i comme des variables à exprimer en fonction des $c_{i,j}$ et des x_j. Certaines des solutions obtenues ne comporteront pas les variables λ_i, ce sont les équations cartésiennes recherchées.

Exemple
Considérons le plan de $\mathbf{R}^3$ engendré par les deux vecteurs (1,2,3) et (4,5,6). Un vecteur de coordonnées (x_1,x_2,x_3) est dans ce plan s'il s'exprime comme combinaison linéaire de ces deux vecteurs, c'est-à-dire si on peut trouver λ_1, $\lambda_2 \in \mathbf{R}$ tels que

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

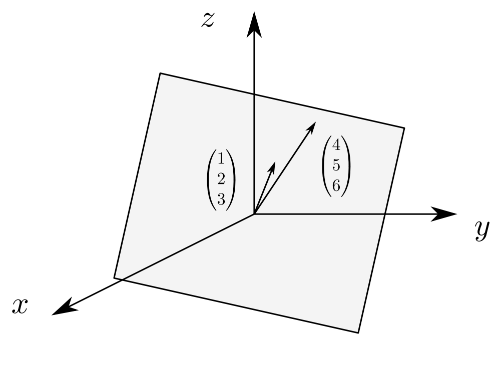
L'équation paramétrique de ce plan est donc

$$\left\{\begin{array}{rcrcr} x_1 & = & \lambda_1 & + & 4\lambda_2 \\ x_2 & = & 2\lambda_1 & + & 5\lambda_2 \\ x_3 & = & 3\lambda_1 & + & 6\lambda_2 \end{array}\right.$$

Appliquons le pivot de Gauss relativement aux variables (λ_1,λ_2) :

$$\left\{\begin{array}{ccrcr} x_1 & = & \lambda_1 & + & 4\lambda_2 \\ x_2 - 2x_1 & = & & - & 3\lambda_2 \\ x_3 - 3x_1 & = & & - & 6\lambda_2 \end{array}\right.$$

$$\left\{\begin{array}{ccrcrl} x_1 & = & \lambda_1 & + & 4\lambda_2 & \\ x_2 - 2x_1 & = & & - & 3\lambda_2 & \\ x_1 - 2x_2 + x_3 & = & & & 0 & \quad \leftarrow \text{ équation cartésienne du plan} \end{array}\right.$$



Expliquons enfin comment obtenir des bases explicites pour les sous-espaces vectoriels de $\mathbf{R}^n$.

Méthode : *calcul d'une base d'un sous-espace vectoriel de* $\mathbf{R}^n$

Si on dispose d'un système d'équations cartésiennes associé à une matrice A, il suffit de calculer une base du noyau de A, comme cela a été expliqué précédemment. Si on dispose d'un système d'équations paramétriques associé à une matrice C, il suffit de calculer une base de l'image de C.

Exemple
Calculons une base d'un sous-espace de $\mathbf{R}^n$ défini par une seule équation cartésienne de la forme

$$a_1x_1 + \ldots + a_nx_n = 0$$

dans le cas où a_1 est non nul. La variable x_1 est la seule variable principale ce qui amène à introduire $n-1$ paramètres $\lambda_1, \ldots, \lambda_{n-1}$. On pose $x_2 = \lambda_1$, $\ldots$, $x_n = \lambda_{n-1}$ et on exprime x_1 en fonction de ces paramètres.

$$\begin{array}{rcl} x_1 & = & -\frac{1}{a_1}(a_2\lambda_1 + \ldots + a_n\lambda_{n-1}) \\ x_2 & = & \lambda_1 \\ \vdots & & \vdots \\ x_n & = & \lambda_{n-1} \end{array}$$

Nous obtenons

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix} = \lambda_1 \begin{pmatrix} -a_2/a_1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -a_3/a_1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \lambda_{n-1} \begin{pmatrix} -a_n/a_1 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

La base recherchée est composée des $n-1$ vecteurs situés dans le second membre de cette égalité.

4. Compléments

4.1 Matrices échelonnées réduites de taille trois

Donnons quelques exemples de matrices échelonnées réduites afin de bien visualiser leur structure. La forme générale de ces matrices est la suivante.

$$\begin{pmatrix}
0 & \dots & 0 & 1 & \square & 0 & \square & 0 & \square & 0 & \square \\
 & & & 0 & \dots\ 0 & 1 & \square & 0 & \square & 0 & \square \\
 & & & 0 & & 0 & \dots\ 0 & 1 & \square & 0 & \square \\
 & & & 0 & & 0 & & 0 & \dots\ 0 & 1 & \square \\
 & & & 0 & & 0 & & \vdots & & 0 & \dots\ 0
\end{pmatrix}$$

Les positions vides sont occupées par des coefficients nuls et les rectangles par des coefficients arbitraires. Faisons la liste des matrices échelonnées réduites de taille 3×3 selon leur rang et donnons pour chacune d'elles une base du noyau.

Rang 0 : seule la matrice nulle est de rang 0, $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, de noyau $\mathbf{R}^3$.

Rang 1 : la matrice possède un seul pivot, ses deux dernières lignes doivent être nulles et le pivot peut prendre trois positions. Suivant sa position, un ou deux coefficients peuvent prendre des valeurs quelconques, notons les a et b.

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

$$\text{Base du noyau :} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ a \\ -1 \end{pmatrix} \qquad \begin{pmatrix} a \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ 0 \\ -1 \end{pmatrix}$$

Rang 2 : la matrice possède deux pivots, donc deux colonnes égales aux deux premiers vecteurs de la base canonique. Ces vecteurs peuvent adopter trois positions différentes.

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

$$\text{Base du noyau :} \quad \begin{pmatrix} a \\ b \\ -1 \end{pmatrix} \qquad \begin{pmatrix} a \\ -1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Rang 3 : la seule matrice échelonnée réduite de rang trois est l'identité car ses trois colonnes sont les trois vecteurs de la base canonique, dans l'ordre.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ le noyau vaut } \{0\}.$$

4.2 Algorithme de Gauss en colonnes

Nous avons vu comment placer une matrice sous forme échelonnée réduite en opérant sur ses lignes. Il est aussi possible d'opérer sur ses colonnes. Cela donne une nouvelle méthode pour calculer le noyau d'une matrice.

Définition 4 *Une matrice est dite* échelonnée en colonnes *si le nombre de zéros consécutifs au début de chaque colonne augmente strictement de colonnes en colonnes jusqu'à ce qu'il ne reste éventuellement plus que des colonnes composées exclusivement de zéros.*

Les pivots *sont les premiers termes non nuls de chaque colonne. Une matrice échelonnée en colonnes est* réduite *si ses pivots sont tous égaux à 1 et dans une ligne contenant un pivot, tous les coefficients sont nuls sauf le pivot.*

En appliquant l'algorithme de Gauss en colonnes plutôt qu'en lignes, on obtient la décomposition suivante.

Théorème 2 (échelonnage en colonnes) *Soit A une matrice de taille $m \times n$. Alors il existe une matrice P de taille $n \times n$ inversible et une matrice R échelonnée en colonnes et réduite, de taille $m \times n$, telles que*

$$R = AP.$$

Pour calculer P, on concatène la matrice identité *en dessous* de A et on applique l'algorithme du pivot de Gauss.

$$\left(\frac{A}{id}\right) P = \left(\frac{AP}{P}\right) = \left(\frac{R}{P}\right).$$

Exemple
Appliquons l'algorithme à la matrice A utilisée dans les exemples précédents.

$$\left(\begin{array}{ccc} 1 & -1 & 2 \\ 2 & 2 & 8 \\ 3 & 3 & 12 \\ 1 & 2 & 5 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & 4 & 4 \\ 3 & 6 & 6 \\ 1 & 3 & 3 \\ \hline 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \\ 1 & 3 & 0 \\ \hline 1 & 1 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & \frac{3}{2} & 0 \\ 1 & \frac{3}{4} & 0 \\ \hline 1 & \frac{1}{4} & -3 \\ 0 & \frac{1}{4} & -1 \\ 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{3}{2} & 0 \\ -\frac{1}{2} & \frac{3}{4} & 0 \\ \hline \frac{1}{2} & \frac{1}{4} & -3 \\ -\frac{1}{2} & \frac{1}{4} & -1 \\ 0 & 0 & 1 \end{array}\right)$$

La partie supérieure de la troisième matrice correspond à la forme échelonnée en colonnes de A, la dernière à sa forme réduite.

La relation $R = AP$ montre que la matrice A envoie les colonnes de P sur les colonnes de R. En d'autres termes, elle envoie les colonnes situées sous la ligne de séparation sur celles situées au dessus de la ligne de séparation. Cela permet de trouver rapidement une base du noyau de A, il suffit de prendre les vecteurs situés sous les colonnes nulles de R. Dans l'exemple, le noyau est de dimension un, engendré par le vecteur $\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}$.

4.3 Unicité de la forme échelonnée réduite

Étudions la question de l'unicité dans le théorème 1.

Proposition 3 *Soit A une matrice de taille $m \times n$, soient P et P' deux matrices inversibles de taille $m \times m$, R et R' deux matrices échelonnées de taille $m \times n$ telles que*

$$R = PA, \quad R' = P'A.$$

Alors $R = R'$.

Cela implique, entre autre, que la forme échelonnée réduite obtenue par l'algorithme de Gauss ne dépend pas de l'ordre des manipulations effectuées au cours de l'échelonnage.

Pour montrer la proposition, expliquons comment associer à la matrice A une matrice échelonnée réduite R sans faire appel à l'algorithme de Gauss. Notons $A_1, \ldots, A_n$ les colonnes de A et A_0 le vecteur colonne nul. On procède par examen des colonnes de la première à la dernière. Prenons un ensemble d'indices $I \subset \{1, \ldots, n\}$ initialement vide, il va contenir les indices des colonnes contenant les pivots. Le vecteur A_0 permet d'initialiser la récurrence. On part de $k = 1$.

– Si A_k est linéairement indépendante des A_i, $i < k$, alors on rajoute l'indice k à I et on prend pour R_k un vecteur de la base canonique, celui dont le coefficient non nul est à la position $Card(I) + 1$.

– Si A_k est linéairement dépendante des A_i, $i < k$, alors il existe des coefficients a_i tels que

$$A_k = \sum_{i=1}^{n} a_i A_i.$$

Ces coefficients sont uniquement déterminés si on impose $a_i = 0$ pour $i \notin I$ car les A_i, $i \in I$, forment une base de $\mathrm{vect}(A_i,\, i < k)$. On prend pour R_k le vecteur colonne $(a_1, \ldots, a_n)$. L'ensemble d'indices I est inchangé.

La matrice R ainsi obtenue est échelonnée réduite. Comme les opérations sur les lignes utilisées dans l'algorithme de Gauss laissent invariantes les relations sur les colonnes, tout matrice obtenue à partir de A en appliquant ces opérations est associée à la même matrice R par la procédure précédente. De plus, si nous appliquons cette procédure à une matrice échelonnée réduite, nous obtenons la matrice elle-même. Toute matrice échelonnée réduite obtenue par l'algorithme de Gauss doit donc coïncider avec R.

Remarquons pour terminer que la matrice inversible P n'est pas nécessairement unique. Si

$$R = PA = P'A,$$

alors $(P - P')A = 0$. On ne peut en déduire l'égalité $P - P' = 0$ que si A est surjective. Dans le cas contraire, rajouter à P une matrice qui s'annule sur l'image de A donne une matrice P' différente de P satisfaisant $R = P'A$.

Chapitre 2

Formes quadratiques

L'étude des formes bilinéaires symétriques passe par un algorithme de réduction qui permet de décomposer toute forme quadratique en carrés de formes linéaires indépendantes.

Ce chapitre commence donc par l'étude du concept de forme linéaire avant de présenter les notions clefs concernant les formes quadratiques : noyau, rang, changement de base, signature. On termine par une application à l'étude des coniques et des quadriques.

1. Formes linéaires

1.1 Définition d'une forme linéaire

Définition 5 *Soit E un espace vectoriel réel. Une* forme linéaire *sur E est une application linéaire définie de E dans* $\mathbf{R}$.

Une forme linéaire est donc une fonction $l : E \to \mathbf{R}$ qui satisfait, pour tout $x,y \in E$ et $\lambda \in \mathbf{R}$,

$$l(\lambda x + y) = \lambda l(x) + l(y).$$

On définit de manière analogue la notion de forme linéaire sur un espace vectoriel complexe en remplaçant $\mathbf{R}$ par $\mathbf{C}$ dans ce qui précède.

Définition 6 *L'ensemble des formes linéaires sur un espace vectoriel E est noté E^*. C'est l'espace vectoriel dual de E.*

Exemples

– Soit E un espace vectoriel quelconque. La fonction nulle, qui associe à tout vecteur de E le réel 0 est une forme linéaire.

– Soit $a_1,a_2 \in \mathbf{R}$. La fonction $l(x_1,x_2) = a_1x_1 + a_2x_2$ est une forme linéaire définie sur $\mathbf{R}^2$.

– Notons $C^0([0,1],\mathbf{R})$ l'espace vectoriel des fonctions continues définies sur l'intervalle $[0,1]$ à valeurs réelles. L'application

$$f \mapsto \int_0^1 f(t)\,dt$$

est une forme linéaire sur $C^0([0,1],\mathbf{R})$.

Proposition 4 *Les formes linéaires définies sur* $\mathbf{R}^n$ *sont les fonctions de la forme*

$$l(x_1, \ldots, x_n) = a_1 x_1 + \ldots + a_n x_n.$$

Les réels $a_1, \ldots, a_n$ *sont les coefficients de la forme linéaire* l.

Vérifions que toute forme linéaire est bien de cette forme. Pour cela, posons $a_1 = l(1,0, \ldots ,0), \ldots, a_n = l(0,0, \ldots ,1)$. Par linéarité,

$$l(x_1, \ldots ,x_n) = x_1\, l(1,0, \ldots ,0) + \ldots + x_n\, l(0,0, \ldots ,1) = a_1 x_1 + \ldots + a_n x_n.$$

1.2 Noyau d'une forme linéaire

Le noyau d'une forme linéaire définie sur E est un sous-espace vectoriel de E. Nous allons caractériser ces sous-espaces par leur dimension.

Définition 7 *Soit* E *un espace vectoriel de dimension finie* n*. Un sous-espace vectoriel de dimension* $n-1$ *est appelé* hyperplan *de* E*.*

Exemple
– Si E est de dimension 2, les hyperplans de E sont des droites.
– Si E est de dimension 3, les hyperplans de E sont des plans.

D'un point de vue géométrique, les vecteurs dirigent des droites vectorielles. Les formes linéaires dirigent des hyperplans.

Proposition 5 *Soit* E *un espace vectoriel de dimension finie* n*.*

– Le noyau d'une forme linéaire non nulle est un hyperplan.

– Tout hyperplan est le noyau d'une forme linéaire non nulle.

– Deux formes linéaires non nulles définissent le même hyperplan si et seulement si elles sont proportionnelles.

Démonstration
Le premier point découle de la formule du rang. Soit l une forme linéaire.

$$dim\ \ker(l) + rang(l) = n.$$

Si l est nulle, son image est égale à $\{0\}$ et $rang(l) = 0$.
Si l est non nulle, son image est égale à $\mathbf{R}$ et $rang(l) = 1$.

Soit H un hyperplan de E et $e_1, \ldots ,e_{n-1}$ une base de H. Soit e_n un vecteur de E qui n'est pas dans H. Il n'est pas combinaison des $e_1, \ldots, e_{n-1}$, la famille $(e_1, \ldots, e_{n-1},e_n)$ est donc libre, c'est une base de E. Un vecteur $x = x_1 e_1 + \ldots + x_n e_n$ de E est dans H si et seulement si il s'exprime en fonction de $e_1, \ldots, e_{n-1}$, c'est-à-dire si et seulement si $x_n = 0$. C'est donc le noyau de la forme linéaire définie par $l(x_1 e_1 + \ldots + x_n e_n) = x_n$.

Soit l_1, $l_2 \in E^*$ non nulles telles que $H = \ker(l_1) = \ker(l_2)$. Dans la base de E qu'on vient de considérer, $l_1(e_i) = l_2(e_i) = 0$ si $i < n$. On en déduit

$$l_1(\sum x_i e_i) = x_n\, l_1(e_n), \quad l_2(\sum x_i e_i) = x_n\, l_2(e_n),$$

d'où $l_1 = \frac{l_1(e_n)}{l_2(e_n)}\, l_2$. Ces deux formes sont bien proportionnelles.

Corollaire 3 *Tout hyperplan de $\mathbf{R}^n$ est de la forme*

$$H = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid a_1 x_1 + \ldots + a_n x_n = 0\}$$

et les coefficients $a_1, \ldots, a_n$ sont définis à un facteur multiplicatif près.

1.3 Base duale

Lorsque E est de dimension finie, on peut associer à toute base de E une base de E^* en procédant comme suit.

Définition 8 *Soit E un espace vectoriel de dimension finie et $(e_1, \ldots, e_n)$ une base de E. Les applications*

$$e_k^*\Big(\sum_{i=1}^n x_i e_i\Big) = x_k$$

sont des formes linéaires sur E. La famille $(e_1^*, \ldots, e_n^*)$ forme une base de E^* appelée base duale de $(e_1, \ldots, e_n)$. Pour tout $x \in E$, nous avons

$$x = \sum_{i=1}^n e_i^*(x)\, e_i.$$

L'application e_i^* associe à un vecteur sa i^{eme} coordonnée. On parle parfois d'application coordonnée. Remarquons que $e_k^*(e_i) = 0$ si $i \neq k$ et $e_k^*(e_k) = 1$.

Démonstration
Vérifions que les (e_k^*) forment une base de E^*.
La famille (e_k^*) est libre: soit $\lambda_1, \ldots, \lambda_n$ tels que $\sum_k \lambda_k e_k^* = 0$.

$$0 = \Big(\sum_k \lambda_k e_k^*\Big)(e_i) = \sum_k \lambda_k e_k^*(e_i) = \lambda_i.$$

Tous les coefficients λ_i sont donc nuls.
La famille (e_k^*) est génératrice: Soit $l \in E^*$, $x = \sum_{k=1}^n x_k e_k$. Par linéarité,

$$l(x) = l\Big(\sum_{k=1}^n x_k e_k\Big) = \sum_{k=1}^n x_k\, l(e_k) = \sum_{k=1}^n l(e_k) e_k^*(x).$$

On a donc $l = \sum l(e_k) e_k^*$, la forme l est combinaison linéaire des e_k^*.

Corollaire 4 *Les coordonnées d'une forme linéaire $l \in E^*$ dans la base $(e_1^*, \ldots, e_n^*)$ valent*

$$l(e_1),\ l(e_2), \ldots, l(e_n).$$

L'espace E^* est de dimension finie et dim E^* = dim E.

On convient de représenter les coordonnées d'un vecteur dans une base donnée (e_i) de E sous forme de vecteur colonne et de représenter les coordonnées d'une forme linéaire dans la base duale (e_i^*) de E^* sous forme de vecteur ligne. Avec cette convention, si $x \in E$ et $l \in E^*$ ont pour coordonnées

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad (\, a_1 \ldots a_n \,),$$

alors la quantité $l(x)$ s'obtient en faisant le produit des deux vecteurs

$$l(x) = (\, a_1 \ldots a_n \,) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \ldots + a_n x_n.$$

Le vecteur $(\, a_1 \ldots a_n \,)$ s'interprète aussi comme la matrice de l'application linéaire $l : E \to \mathbf{R}$ dans les bases $(e_1, \ldots, e_n)$ de E et (1) de $\mathbf{R}$.

1.4 Le cas de $\mathbf{R}^n$

Base canonique

Rappelons ce qu'est la *base canonique* de $\mathbf{R}^n$. Avec la convention qui consiste à représenter les vecteurs de $\mathbf{R}^n$ sous forme de colonnes, il s'agit de la famille de vecteurs

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Tout vecteur de $\mathbf{R}^n$ s'exprime dans cette base comme suit.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

On a convenu de représenter les éléments du dual $\mathbf{R}^{n*}$ sous forme de vecteurs lignes. Les éléments de la base duale de la base canonique sont alors

$$(1\ 0\ \ldots\ 0), \quad (0\ 1\ \ldots\ 0), \ldots, \quad (0\ 0\ \ldots\ 1).$$

Tout vecteur de $\mathbf{R}^{n*}$ s'exprime dans cette base de la façon suivante.

$$\begin{aligned}(\,x_1 \quad x_2 \quad \ldots \quad x_n\,) = \quad & x_1\ (\,1 \quad 0 \quad \ldots \quad 0\,)\\ &+ x_2\ (\,0 \quad 1 \quad \ldots \quad 0\,)\\ &+\ \ldots\\ &+ x_n\ (\,0 \quad 0 \quad \ldots \quad 1\,).\end{aligned}$$

La base duale de la base canonique de $\mathbf{R}^n$ est appelée base canonique de $\mathbf{R}^{n*}$.

Méthode : *calcul de la base duale d'une base de* $\mathbf{R}^n$

Soit $(e_1, \ldots, e_n)$ une base de $\mathbf{R}^n$. Pour calculer les coordonnées de la base duale $(e_1^*, \ldots, e_n^*)$ dans la base canonique de $\mathbf{R}^{n*}$, on considère la matrice P dont les colonnes sont les coordonnées des e_j dans la base canonique. Les coordonnées des e_i^* dans la base canonique du dual sont les lignes de la matrice P^{-1} inverse de P.

De fait, le produit de la matrice dont les lignes sont les coordonnées des e_i^* par la matrice dont les colonnes sont les coordonnées des e_j est une matrice qui a comme coefficients les $e_i^*(e_j)$, elle est donc égale à l'identité.

Exemple
On considère la base de $\mathbf{R}^2$ donnée par

$$e_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix},\ e_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Inversons la matrice dont les colonnes sont données par ces deux vecteurs.

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

On obtient $e_1^*(x_1,x_2) = x_1 - x_2,\quad e_2^*(x_1,x_2) = x_1 - 2x_2$. On peut vérifier directement que

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 - x_2)\begin{pmatrix} 2 \\ 1 \end{pmatrix} + (x_1 - 2x_2)\begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Indépendance

Les formes linéaires sont les éléments d'un espace vectoriel. On peut donc s'intéresser à l'indépendance d'une famille de formes linéaires. Rappelons qu'une famille de vecteurs est linéairement indépendante s'il n'y a pas de relation linéaire non triviale entre ces vecteurs.

Méthode : *indépendance d'une famille de formes linéaires*

Pour déterminer si une famille de formes linéaires $(l_1, \ldots, l_k)$ est linéairement indépendante, on décompose cette famille dans une base et on forme la matrice dont les lignes sont données par les vecteurs lignes associés à chacune des formes linéaires. La famille est linéairement indépendante si et seulement si le rang de cette matrice est égal à k.

Exemple
On considère les trois formes linéaires définies sur $\mathbf{R}^3$ par

$$l_1(x_1,x_2,x_3) = x_1 + x_2, \quad l_2(x_1,x_2,x_3) = x_2 + x_3, \quad l_3(x_1,x_2,x_3) = x_1 - x_3.$$

Les coordonnées de ces formes linéaires dans la base canonique sont

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

La matrice obtenue en empilant ces lignes est égale à $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$.

On détermine son rang en la mettant sous forme échelonnée :

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} L_3 \leftarrow L_3 - L_1 \qquad \begin{pmatrix} \mathbf{1} & 1 & 0 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{pmatrix} L_3 \leftarrow L_3 + L_2$$

Elle est de rang 2. La famille n'est donc pas linéairement indépendante. On aurait pu aller plus vite en remarquant directement qu'il y a une relation linéaire non triviale entre les formes linéaires l_1, l_2 et l_3 : $l_1 = l_2 + l_3$.

Remarquons aussi que deux vecteurs sont linéairement indépendants si et seulement si ils ne sont pas proportionnels. De même, deux formes sont linéairement indépendantes si et seulement si elles ne sont pas proportionnelles.

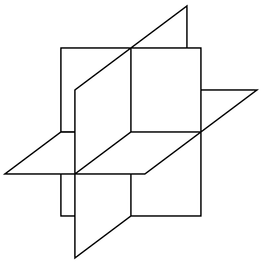
Lien avec les systèmes linéaires
Considérons un système linéaire homogène sur $\mathbf{R}^n$ de la forme

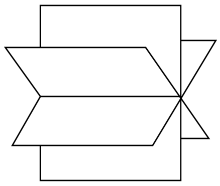
$$\begin{cases} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & 0 \end{cases}$$

Définissons des formes linéaires l_i en posant

$$l_i(x) = a_{i,1}x_1 + \dots + a_{i,n}x_n.$$

Nous voyons qu'un vecteur est solution du système si et seulement s'il annule toutes ces formes linéaires, c'est-à-dire s'il appartient à $\ker(l_i)$ pour tout i. Géométriquement, l'espace vectoriel des solutions du système est une intersection d'hyperplans.





Considérons le cas $m = n = 3$ et supposons les hyperplans tous distincts. Si le système est de rang trois, seul le vecteur nul est solution du système et les trois hyperplans se rencontrent en l'origine seulement. S'il est de rang deux, les trois hyperplans se rencontrent selon une droite.

2. Formes bilinéaires

2.1 Définition d'une forme bilinéaire

Définition 9 *Soit E un espace vectoriel réel. Une* forme bilinéaire *est une fonction définie sur $E \times E$ à valeurs réelles qui est linéaire en chacune de ses variables.*

Il s'agit donc d'une fonction $\varphi : E \times E \to \mathbf{R}$ qui satisfait pour tout x,x',y,y' dans E et $\lambda \in \mathbf{R}$,

- $\varphi(\lambda x + x',y) = \lambda\varphi(x,y) + \varphi(x',y)$,
- $\varphi(x,\lambda y + y') = \lambda\varphi(x,y) + \varphi(x,y')$.

Définition 10 *Une forme bilinéaire est* symétrique *si pour tout $x,y \in E$,*

$$\varphi(x,y) = \varphi(y,x).$$

Exemples

– Sur $\mathbf{R}$, les formes bilinéaires sont de la forme $\varphi(x,y) = axy$ pour $a \in \mathbf{R}$.

– Le produit scalaire usuel sur $\mathbf{R}^n$ est une forme bilinéaire symétrique. Pour $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n) \in \mathbf{R}^n$, il est donné par

$$\varphi(x,y) = \sum_{k=1}^{n} x_k y_k = x_1y_1 + \ldots + x_ny_n.$$

– Soit a, b, c, $d \in \mathbf{R}$ et $x = (x_1,x_2)$, $y = (y_1,y_2) \in \mathbf{R}^2$. La fonction suivante est une forme bilinéaire sur $\mathbf{R}^2$:

$$\varphi(x,y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2.$$

Elle est symétrique si et seulement si $b = c$.

– Sur l'espace vectoriel $C^0([0,1],\mathbf{R})$, la fonction suivante est une forme bilinéaire symétrique.

$$\varphi(f,g) = \int_0^1 f(t)g(t)\,dt.$$

Nous verrons que cet exemple joue un rôle important en analyse.

2.2 Représentation matricielle

On se place maintenant dans le cadre de la dimension finie. On va associer une matrice à toute forme bilinéaire.

Définition 11 *Soit E un espace vectoriel de dimension finie muni d'une base $(e_1,\dots,e_n)$ et φ une forme bilinéaire sur E. La matrice de φ dans la base $(e_1,\dots,e_n)$ est la matrice carrée*

$$B = \{\varphi(e_i,e_j)\}_{\substack{i=1..n\\ j=1..n}} = \begin{pmatrix} \varphi(e_1,e_1) & \dots & \varphi(e_1,e_n) \\ \vdots & & \vdots \\ \varphi(e_n,e_1) & \dots & \varphi(e_n,e_n) \end{pmatrix}.$$

La forme φ est symétrique si et seulement si sa matrice est symétrique.

Exemple

– La matrice associée à la forme bilinéaire sur $\mathbf{R}^2$ donnée par

$$\varphi(x,y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

est la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

– La matrice associée au produit scalaire usuel sur $\mathbf{R}^n$ est la matrice identité.

La forme bilinéaire φ s'exprime en coordonnées à l'aide de la matrice B, notons $b_{i,j}$ ses coefficients. Considérons deux vecteurs $x,y \in E$ que nous décomposons dans la base $(e_1,\dots,e_n)$ et utilisons la bilinéarité de φ.

$$x = \sum x_ie_i, \quad y = \sum y_je_j, \quad b_{i,j} = \varphi(e_i,e_j).$$

$$\varphi(x,y) = \varphi(\sum_i x_ie_i, \sum_j y_je_j) = \sum_{i,j} x_iy_j\varphi(e_i,e_j) = \sum_{i,j} b_{i,j}x_iy_j.$$

Nous avons obtenu une expression pour φ dépendant des $b_{i,j}$, x_i et y_j.

$$\begin{aligned}
\varphi(x,y) &= \sum_{i,j} b_{i,j}x_i\ y_j \\
&= b_{1,1}\,x_1y_1 + b_{1,2}\,x_1y_2 + \dots + b_{n,n}\,x_ny_n \\
&= \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.
\end{aligned}$$

Dans le cas $E = \mathbf{R}^n$ et (e_i) la base canonique, cette expression devient

$$\varphi(x,y) = {}^t x\, B\, y.$$

2.3 Changement de base

La matrice associée à une forme bilinéaire symétrique dépend de la base choisie pour la représenter. Expliquons comment cette matrice est transformée lorsqu'on effectue un changement de base.

Rappelons comment est définie la *matrice de passage* entre deux bases $(e_i)_{i=1..n}$ et $(e'_i)_{i=1..n}$ d'un espace de dimension finie. Il s'agit de la matrice dont les colonnes sont composées des coordonnées des vecteurs de la nouvelle base (e'_i) dans l'ancienne base (e_i). Notons cette matrice $P = \{p_{i,j}\}$.

$$(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)\, P, \qquad e'_j = \sum_{i=1}^{n} p_{i,j} e_i.$$

Soit $x = (x_1, \ldots, x_n)$ les coordonnées d'un vecteur dans l'ancienne base et $x' = (x'_1, \ldots, x'_n)$ les coordonnées dans la nouvelle base. On a alors

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad x_i = \sum_{j=1}^{n} p_{i,j} x'_j.$$

Les matrices A et A' d'une application linéaire $f : E \to E$ dans les bases (e_i) et (e'_i) satisfont la relation

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

On dit que les matrices A et A' sont *conjuguées*. La relation qui lie les matrices associées à une forme bilinéaire est différente.

Proposition 6 *Soit E un espace de dimension finie muni de deux bases $(e_i)_{i=1..n}$, $(e'_i)_{i=1..n}$ et d'une forme bilinéaire $\varphi : E \times E \to \mathbf{R}$. Notons par P la matrice de passage entre ces deux bases et B, B' les matrices de φ dans chacune des bases. Alors*

$$B' = {}^tP\, B\, P.$$

Démonstration
$\varphi(x,y) = {}^t x B y = {}^t(Px') B (Py') = {}^t x' \left({}^t PBP\right) y'$.

Deux matrices carrées B et B' liées par une relation du type $B' = {}^tP\, B\, P$, où P est une matrice inversible, sont dites *congruentes*.

2.4 Rang et noyau

Définition 12 *Le noyau d'une forme bilinéaire est défini par*

$$\ker(\varphi) = \{y \in E \mid \forall\, x \in E,\ \varphi(x,y) = 0\}.$$

Une forme bilinéaire est dite non dégénérée *si son noyau est restreint au vecteur nul :* $\ker(\varphi) = \{0\}$.

Il s'agit des vecteurs $y \in E$ pour lesquels $\varphi(x,y)$ est nul pour tout $x \in E$. C'est un sous-espace vectoriel de E. Dans une base donnée, il coïncide avec le noyau de la matrice associée à la forme bilinéaire.

Exemples
– La forme bilinéaire sur $\mathbf{R}^2$ définie par $\varphi(x,y) = x_1 y_1$ a pour matrice $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Son noyau est constitué des vecteurs dont la première coordonnée s'annule :

$$\ker(\varphi) = \mathrm{vect}\Big(\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\Big).$$

– Un exemple de forme bilinéaire non dégénérée est donné par le produit scalaire usuel sur $\mathbf{R}^n$. De fait, la matrice associée est l'identité, son noyau est bien restreint à $\{0\}$.
– La forme bilinéaire définie sur $C^0([0,1],\mathbf{R})$ par $\varphi(f,g) = \int_0^1 f(t)g(t)\,dt$ est elle aussi non dégénérée. De fait, si $\varphi(f,g) = 0$ pour tout g, on peut prendre $g = f$ et obtenir $\int f(t)^2\,dt = 0$. On conclut en utilisant le résultat suivant d'intégration : une fonction continue positive dont l'intégrale est nulle est identiquement nulle.

Définition 13 *Le* rang *d'une forme bilinéaire définie sur un espace vectoriel de dimension finie est égal au rang de sa matrice. Il ne dépend pas de la base choisie pour exprimer cette matrice.*

La formule du rang pour les matrices entraîne

$$dim\ \ker(\varphi) + rang(\varphi) = dim\ E.$$

Terminons en remarquant que l'ensemble des formes bilinéaires symétriques est un espace vectoriel. On combine deux formes bilinéaires φ_1, φ_2 définies sur un même espace par la formule

$$(\lambda\varphi_1 + \varphi_2)(x,y) = \lambda\varphi_1(x,y) + \varphi_2(x,y).$$

3. Formes quadratiques

3.1 Définition d'une forme quadratique

On se concentre maintenant sur l'étude des formes bilinéaires symétriques. On va associer à une telle forme un polynôme de degré deux à plusieurs variables qu'on pourra ensuite simplifier comme on le faisait avec les polynômes quadratiques à une variable.

Définition 14 *Une* forme quadratique *Q définie sur un espace vectoriel réel E est une fonction de E dans $\mathbf{R}$ de la forme $\varphi(x,x)$, où $\varphi : E \times E \to \mathbf{R}$ est une forme bilinéaire symétrique.*

$$\forall\, x \in E, \quad Q(x) = \varphi(x,x).$$

La forme bilinéaire φ est appelée *forme polaire* de Q. Soit E un espace de dimension finie muni d'une base $(e_i)_{i=1..n}$. Donnons l'expression générale d'une forme quadratique en coordonnées. Soit B la matrice de la forme bilinéaire associée à Q. Alors

$$Q\Big(\sum_{k=1}^{n} x_k e_k\Big) = {}^t x B x = \sum_{i,j} b_{i,j} x_i x_j = \sum_i b_{i,i} x_i^2 + 2 \sum_{\substack{i,j \\ i<j}} b_{i,j} x_i x_j.$$

Cette expression est un polynôme à plusieurs variables, dont tous les termes sont de degré total deux.

Exemples

– La forme quadratique sur $\mathbf{R}^2$ associée à $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ est donnée par

$$Q(x_1,x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

– La forme quadratique associée au produit scalaire usuel de $\mathbf{R}^n$ vaut

$$Q(x_1,\ldots,x_n) = \sum_{k=1}^{n} x_k^2 = x_1^2 + \ldots + x_n^2.$$

C'est le carré de la longueur du vecteur $(x_1,\ldots,x_n)$.

Étant donnée une forme quadratique Q, il n'existe qu'une seule forme bilinéaire symétrique pour laquelle on a $Q(x) = \varphi(x,x)$ pour tout x. On l'obtient par le biais des formules suivantes.

Proposition 7

$$Q(\lambda x) = \lambda^2 Q(x), \quad Q(x+y) = Q(x) + 2\varphi(x,y) + Q(y),$$

$$\begin{array}{rcl} \varphi(x,y) & = & \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big), \\ \varphi(x,y) & = & \frac{1}{2}\Big(Q(x) + Q(y) - Q(x-y)\Big), \\ \varphi(x,y) & = & \frac{1}{4}\Big(Q(x+y) - Q(x-y)\Big). \end{array}$$

Les trois dernières formules s'appellent *identités de polarisation.* Démontrons-les par un calcul direct.

$$\begin{array}{rcl} Q(x+y) & = & \varphi(x+y,x+y) \\ & = & \varphi(x,x) + \varphi(y,x) + \varphi(x,y) + \varphi(y,y) \\ & = & Q(x) + 2\varphi(x,y) + Q(y). \end{array}$$

$$Q(x-y) = Q(x+(-y)) = Q(x) - 2\varphi(x,y) + Q(y).$$

Il suffit de soustraire ces deux égalités pour obtenir la dernière formule.

En pratique, lorsque l'espace est de dimension finie, on calcule facilement la matrice B associée à une forme quadratique à partir des coefficients de son polynôme.

Méthode : *calcul de la matrice associée à une forme quadratique*

Pour calculer la matrice associée à une forme quadratique, on place sur la diagonale les coefficients des carrés x_i^2 apparaissant dans le polynôme de degré deux définissant Q. Pour les termes qui ne sont pas sur la diagonale, il faut diviser par deux les coefficients des termes $x_i x_j$ du polynôme.

Exemple
Soit $Q(x_1,x_2,x_3) = x_1^2 + 2x_2^2 + 6x_1x_2 + 10x_1x_3$. La matrice associée vaut

$$\begin{pmatrix} 1 & 3 & 5 \\ 3 & 2 & 0 \\ 5 & 0 & 0 \end{pmatrix}.$$

On définit également le *rang* et le *noyau* d'une forme quadratique. Ils sont égaux au rang et au noyau de la forme bilinéaire symétrique ou de la matrice associée.

Il reste une dernière notion à introduire.

Définition 15 *Soit Q une forme quadratique sur un espace vectoriel E. Le* cône *des vecteurs isotropes de Q est défini par*

$$C(Q) = \{x \in E \mid Q(x) = 0\}.$$

Un vecteur $x \in E$ satisfaisant $Q(x) = 0$ est dit isotrope.

On verra plus loin pourquoi on parle de cône. Il est important de remarquer que ce cône n'est pas un espace vectoriel en général. Il contient toujours le noyau de la forme quadratique :

$$\ker(Q) \subset C(Q).$$

Exemple
Soit $Q(x_1,x_2) = x_1^2 - x_2^2$. La matrice associée est $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. La forme quadratique est donc de rang deux et $\ker(Q) = \{0\}$. Calculons son cône.

$$Q(x_1,x_2) = 0 \text{ implique } (x_1 + x_2)(x_1 - x_2) = x_1^2 - x_2^2 = 0.$$

c'est-à-dire $x_1 = x_2$ ou $x_1 = -x_2$. Le cône $C(Q)$ est l'union des deux droites vectorielles dirigées par les vecteurs $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

3.2 Signature

On va introduire un invariant important des formes quadratiques, relié au signe que peuvent prendre ses valeurs.

Définition 16 *Une forme quadratique est dite*

- positive *si pour tout* $x \in E$, $Q(x) \geq 0$,
- définie positive *si pour tout* $x \in E$ *non nul,* $Q(x) > 0$,
- négative *si pour tout* $x \in E$, $Q(x) \leq 0$,
- définie négative *si pour tout* $x \in E$ *non nul,* $Q(x) < 0$.

Proposition 8 *Une forme quadratique définie positive est non dégénérée.*

De fait, $\ker(Q) \subset C(Q)$ et le cône est nul dans le cas non dégénéré par définition : $C(Q) = \{x \mid Q(x) = 0\} = \{0\}$.

Exemple
Le produit scalaire usuel sur $\mathbf{R}^n$ est défini positif.
En effet, $Q(x_1, \ldots, x_n) = \sum x_k^2 > 0$ dès que $(x_1, \ldots, x_n)$ est non nul.

On peut toujours restreindre une forme quadratique à un sous-espace vectoriel de E. On notera $Q_{|F}$ la restriction de Q à un sous-espace $F \subset E$. Remarquons qu'une forme quadratique qui est définie positive sur un espace vectoriel E est automatiquement définie positive en restriction à tous les sous-espaces vectoriels de E.

Nous allons nous intéresser à la dimension maximale des sous-espaces en restriction desquels Q est définie positive ou définie négative.

Définition 17 *Soit E un espace vectoriel de dimension finie et Q une forme quadratique sur E. La* signature *de Q est le couple d'entiers $(p,q) \in \mathbf{N}$ donné par*

$$p = max\{dim\, F \mid F \text{ sous-espace de } E \text{ tel que } Q_{|F} \text{ est définie positive}\}$$

$$q = max\{dim\, F \mid F \text{ sous-espace de } E \text{ tel que } Q_{|F} \text{ est définie négative}\}$$

Exemple
Une forme quadratique Q sur un espace E de dimension n qui est définie positive est de signature $(n,0)$. C'est le cas par exemple pour le produit scalaire usuel sur $\mathbf{R}^n$.

De fait, pour une forme quadratique définie positive, le plus grand espace sur lequel Q est définie positive est E lui-même. Si F est un espace sur lequel Q est définie négative, il ne peut pas contenir de vecteur x non nul, car on aurait pour ce vecteur $Q(x) > 0$ et $Q(x) < 0$. Le sous-espace $F = \{0\}$ est le seul sur lequel Q est définie négative.

3.3 Réduction des formes quadratiques

L'étude des formes quadratiques repose sur un algorithme dû à F. Gauss, qui permet de simplifier l'expression d'une forme quadratique en l'écrivant sous forme d'une somme de carrés de formes linéaires.

Théorème 3 (réduction de Gauss) *Soit E un espace de dimension finie, Q une forme quadratique sur E de signature (p,q). Alors il existe $p+q$ formes linéaires indépendantes $l_1, \dots, l_{p+q} \in E^*$ telles que*

$$Q(x) = \sum_{i=1}^{p} l_i(x)^2 - \sum_{j=p+1}^{p+q} l_j(x)^2.$$

Cette décomposition en carrés de formes linéaires n'est pas unique. Cependant, on peut toujours calculer la signature de la forme quadratique à partir d'une telle décomposition, si les formes linéaires sont indépendantes entre elles.

Proposition 9 *Considérons une forme quadratique donnée par l'expression*

$$Q(x) = \pm l_1(x)^2 \pm l_2(x)^2 \pm \dots \pm l_k(x)^2$$

où les l_i sont des formes linéaires indépendantes entre elles. Alors la signature est donnée par le nombre de signes + et le nombre de signes - apparaissant dans cette décomposition.

La démonstration de cette proposition sera donnée plus loin (en 3.6, page 46). Remarquons que la forme bilinéaire symétrique associée à la forme quadratique apparaissant dans le théorème précédent est donnée par

$$\varphi(x,y) = \sum_{i=1}^{p} l_i(x)l_i(y) - \sum_{j=p+1}^{p+q} l_j(x)l_j(y)$$

car φ est bien une forme bilinéaire et $\varphi(x,x) = Q(x)$. Présentons l'algorithme qui permet d'exprimer une forme quadratique sous forme de somme de carrés.

Méthode : *réduction des formes quadratiques*

On se place dans une base arbitraire, Q prend la forme

$$Q(x_1, \dots, x_n) = \sum_i b_{i,i} x_i^2 + 2 \sum_{i<j} b_{i,j} x_i x_j.$$

• Si un des coefficients diagonaux $b_{i,i}$ est non nul, disons $b_{1,1} \neq 0$, on considère le polynôme comme une fonction de la variable x_1 et on le met sous forme canonique, comme suit.

– On place le terme x_1^2 en tête du polynôme et on met x_1 en facteur dans les termes restants, si possible. L'expression de Q prend alors la forme

$$Q(x) = b_{1,1}\Big[\, x_1^2 + x_1(\ldots) + (\ldots)\Big].$$

Les expressions entre parenthèses **ne doivent plus contenir la variable** x_1. Notons les a_1 et a_2 :

$$Q(x) = b_{1,1}\Big[\, x_1^2 + x_1 a_1 + a_2\,\Big].$$

– On factorise l'expression précédente en utilisant l'identité remarquable

$$x_1^2 + x_1 a_1 = \Big(x_1 + \frac{a_1}{2}\Big)^2 - \frac{a_1^2}{4}.$$

– On pose $l_1(x_1, \ldots, x_n) = \sqrt{|b_{1,1}|}(x_1 + \frac{a_1}{2})$. L'expression de Q devient

$$Q(x_1, \ldots, x_n) = \pm l_1(x_1, \ldots, x_n)^2 + b_{1,1}\Big(a_2 - \frac{a_1^2}{4}\Big).$$

– L'expression $a_2 - \frac{a_1^2}{4}$ ne fait plus intervenir la variable x_1. C'est une forme quadratique à laquelle on peut appliquer à nouveau l'algorithme. Le signe devant la forme linéaire l_1 est celui de $b_{1,1}$.

• Si tous les coefficients diagonaux sont nuls, on considère un coefficient $b_{i,j} \neq 0$, disons $b_{1,2}$ pour alléger les notations.

– On place le terme $x_1 x_2$ en tête du polynôme et on met x_1 en facteur dans les termes restants, si possible. On fait de même avec x_2. On obtient

$$Q(x) = 2b_{1,2}\Big[\, x_1 x_2 + x_1(\ldots) + x_2(\ldots) + (\ldots)\Big].$$

Les expressions entre parenthèses **ne doivent plus contenir les variables** x_1, x_2. Notons les a_1, a_2 et a_3 :

$$Q(x) = 2b_{1,2}\Big[\, x_1 x_2 + x_1 a_1 + x_2 a_2 + a_3\,\Big].$$

– On factorise l'expression précédente en utilisant l'identité remarquable

$$x_1 x_2 + x_1 a_1 + x_2 a_2 = (x_1 + a_2)(x_2 + a_1) - a_1 a_2.$$

– On convertit le produit $(x_1 + a_2)(x_2 + a_1)$ en différence de deux carrés à l'aide de l'identité remarquable

$$xy = \frac{1}{4}(x + y)^2 - \frac{1}{4}(x - y)^2.$$

L'expression de Q devient

$$Q(x) = 2b_{1,2}\left[\Big(\frac{x_1 + x_2 + a_1 + a_2}{2}\Big)^2 - \Big(\frac{x_1 - x_2 - a_1 + a_2}{2}\Big)^2 - a_1 a_2 + a_3\right].$$

– On pose

$$\begin{aligned} l_1(x_1,\dots,x_n) &= \sqrt{|2b_{1,2}|}(x_1+x_2+a_1+a_2)/2 \\ l_2(x_1,\dots,x_n) &= \sqrt{|2b_{1,2}|}(x_1-x_2-a_1+a_2)/2 \end{aligned}$$

ce qui donne

$$Q(x_1,\dots,x_n) = \pm l_1(x_1,\dots,x_n)^2 \mp l_2(x_1,\dots,x_n)^2 + 2b_{1,2}(a_3 - a_1a_2).$$

Le dernier terme dans l'expression précédente ne fait plus intervenir les variables x_1 et x_2. C'est une forme quadratique à laquelle on peut appliquer à nouveau l'algorithme.

Au final, l'algorithme repose sur les trois règles suivantes.

Règle 1 : $X^2 + Xa = \left(X + \frac{a}{2}\right)^2 - \frac{a^2}{4}$

Règle 2 : $XY + Xa_1 + Ya_2 = (X+a_2)(Y+a_1) - a_1a_2$

Règle 3 : $XY = \frac{1}{4}(X+Y)^2 - \frac{1}{4}(X-Y)^2$

Remarquons que cet algorithme est une généralisation de la méthode de résolution des polynômes de degré deux à une variable.

Expliquons pourquoi la famille de formes linéaires données par l'algorithme est linéairement indépendante. Considérons la matrice dont les lignes sont données par les coordonnées de ces formes linéaires.

Si dans l'algorithme, nous disposons à chaque étape d'un coefficient diagonal non nul, la matrice obtenue est échelonnée, sans ligne nulle. Son rang est égal à son nombre de ligne, ce qui montre que ces lignes sont bien linéairement indépendantes entre elles.

Si à une certaine étape, il n'y a pas de termes diagonaux non nuls, l'algorithme produit deux formes linéaires l_1 et l_2 dont les expressions sont données plus haut. Il suffit de soustraire la première à la seconde pour mettre la matrice sous forme échelonnée.

Exemples

– $Q(x_1,x_2) = x_1^2 + x_1x_2 + x_2^2$.

$$\begin{aligned} x_1^2 + x_1x_2 + x_2^2 &= \left(x_1 + \frac{x_2}{2}\right)^2 - \frac{x_2^2}{4} + x_2^2 \\ &= \left(x_1 + \frac{x_2}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}x_2\right)^2. \end{aligned}$$

On pose

$$l_1(x_1,x_2) = x_1 + \frac{x_2}{2}, \quad l_2(x_1,x_2) = \frac{\sqrt{3}}{2}x_2$$

ce qui donne $Q = l_1^2 + l_2^2$. La signature vaut (2,0).

– $Q(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_3x_1$.

$$\begin{aligned} x_1x_2 + x_2x_3 + x_3x_1 &= (x_1+x_3)(x_2+x_3) - x_3^2 \\ &= \Big(\frac{x_1+x_2+2x_3}{2}\Big)^2 - \Big(\frac{x_1-x_2}{2}\Big)^2 - x_3^2. \end{aligned}$$

On pose

$$l_1(x_1,x_2,x_3) = \frac{x_1+x_2+2x_3}{2}, \quad l_2(x_1,x_2,x_3) = \frac{x_1-x_2}{2}, \quad l_3(x_1,x_2,x_3) = x_3$$

ce qui donne $Q = l_1^2 - l_2^2 - l_3^2$. La signature vaut (1,2).

3.4 Base orthogonale

À partir du théorème précédent, on peut construire une base de E dans laquelle Q est donnée par une expression particulièrement simple.

Définition 18 *Soit E un espace vectoriel de dimension finie et φ une forme bilinéaire symétrique définie sur E. Deux vecteurs x, $y \in E$ sont orthogonaux entre eux relativement à φ si $\varphi(x,y) = 0$. Une base $(e_1,\dots,e_n)$ de E est orthogonale pour φ ou pour la forme quadratique Q associée à φ si*

$$\varphi(e_i,e_j) = 0 \quad \textit{pour tout } i,j \textit{ tels que } i \neq j.$$

Une base $(e_1,\dots,e_n)$ est orthogonale pour φ si et seulement si la matrice $B = \{\varphi(e_i,e_j)\}$ de φ dans cette base est diagonale. La signature d'une forme quadratique est reliée aux signes de ses coefficients dans une telle base.

Proposition 10 *Soit E un espace de dimension finie, Q une forme quadratique sur E et $(e_1,\dots,e_n)$ une base orthogonale pour Q. Alors*

$$Q(x_1e_1 + \ldots + x_ne_n) = \sum_i Q(e_i)\, x_i^2.$$

La signature de Q est égale à (p,q), où p est égal au nombre de vecteurs e_i pour lesquels $Q(e_i) > 0$ et q au nombre de vecteurs e_i pour lesquels $Q(e_i) < 0$.

Pour montrer ce dernier point, on remarque que Q est de la forme

$$Q(x_1e_1 + \ldots + x_ne_n) = \sum_i \mathrm{signe}(Q(e_i)) \left(\sqrt{|Q(e_i)|}\, x_i\right)^2.$$

Les formes linéaires $x \mapsto \sqrt{|Q(e_i)|}\, x_i$, pour les indices i tels que $Q(e_i) \neq 0$, sont indépendantes entre elles. La signature est donnée par les signes des $Q(e_i)$ grâce à la proposition 9.

Expliquons comment construire des bases orthogonales.

Proposition 11 *Toute forme bilinéaire symétrique définie sur un espace vectoriel de dimension finie admet une base orthogonale.*

Méthode : *construction d'une base orthogonale*

On considère une forme quadratique définie sur $\mathbf{R}^n$.

– On applique l'algorithme de réduction de Gauss pour obtenir des formes linéaires $l_1, \ldots, l_{p+q}$ et on considère la matrice P dont les lignes sont données par les coordonnées de ces formes linéaires.

– Si $p+q < n$, il faut compléter la famille (l_i) de façon à obtenir une base de $\mathbf{R}^{n*}$. Il s'agit de rajouter à la matrice P des lignes afin d'en faire une matrice carrée inversible. On peut prendre ces lignes de la forme $(0\ \ldots\ 1\ \ldots\ 0)$ en positionnant judicieusement les 1. Les placer dans les colonnes correspondant aux colonnes sans pivot de la forme échelonnée de P fonctionne.

– Les vecteurs recherchés ont pour coordonnées les colonnes de l'inverse de la matrice P.

Vérifions que cette procédure donne une base orthogonale. Notons (e_i) la base obtenue et remarquons que la famille (l_j) est sa base duale. Nous avons donc la relation $l_j(x) = x_j$, où x_j est la j^{ieme} coordonnée de x dans la base $(e_1, \ldots, e_n)$. L'algorithme de Gauss exprime la forme quadratique Q en fonction des (l_i) par la formule

$$\begin{aligned} Q(x) &= l_1(x)^2 + \ldots + l_p(x)^2 - l_{p+1}(x)^2 - \ldots - l_{p+q}(x)^2 \\ &= x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2. \end{aligned}$$

La matrice de Q dans la base (e_i) est bien diagonale car aucun terme $x_i x_j$, $i \neq j$, n'est présent. Nous obtenons aussi la proposition suivante.

Théorème 4 (existence de base orthogonale) *Soit E un espace de dimension finie, Q une forme quadratique sur E de signature (p,q). Alors il existe une base orthogonale $(e_1, \ldots, e_n)$ de E dans laquelle Q a pour expression*

$$Q(x_1 e_1 + \ldots + x_n e_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2.$$

Exemples

– $Q(x_1,x_2) = x_1^2 + x_1 x_2 + x_2^2 = l_1(x_1,x_2,x_3)^2 + l_2(x_1,x_2,x_3)^2$
avec

$$l_1(x_1,x_2) = x_1 + \frac{x_2}{2}, \quad l_2(x_1,x_2) = \frac{\sqrt{3}}{2} x_2.$$

Les formes l_1 et l_2 ont pour coordonnées $(1\ \frac{1}{2})$ et $(0\ \frac{\sqrt{3}}{2})$. On forme la matrice P dont les lignes sont données par ces vecteurs lignes.

$$P = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Son inverse est donné par $P^{-1} = \begin{pmatrix} 1 & -\frac{\sqrt{3}}{3} \\ 0 & \frac{2\sqrt{3}}{3} \end{pmatrix}$.

Les colonnes de la matrice P^{-1} forment une base orthogonale pour Q.

– $Q(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_3x_1$
$\qquad\qquad = l_1(x_1,x_2,x_3)^2 - l_2(x_1,x_2,x_3)^2 - l_3(x_1,x_2,x_3)^2$

avec

$$l_1(x_1,x_2,x_3) = \frac{x_1 + x_2 + 2x_3}{2}, \quad l_2(x_1,x_2,x_3) = \frac{x_1 - x_2}{2}, \quad l_3(x_1,x_2,x_3) = x_3.$$

Les formes l_1, l_2 et l_3 ont pour coordonnées $\frac{1}{2}(1\ 1\ 2)$, $\frac{1}{2}(1\ -1\ 0)$ et $(0\ 0\ 1)$. On forme la matrice dont les lignes sont données par ces vecteurs lignes.

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

Il ne reste plus qu'à inverser : $P^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$

Les colonnes de la matrice P^{-1} forment une base orthogonale pour Q.

– $Q(x_1,x_2,x_3) = 4x_1x_3 + 4x_2x_3$
$\qquad\qquad = 4(x_1 + x_2)\,x_3$
$\qquad\qquad = l_1(x_1,x_2,x_3)^2 - l_2(x_1,x_2,x_3)^2$

avec

$$l_1(x_1,x_2,x_3) = x_1 + x_2 + x_3, \quad l_2(x_1,x_2,x_3) = x_1 + x_2 - x_3.$$

Les formes l_1 et l_2 ont pour coordonnées $(1\ 1\ 1)$ et $(1\ 1\ -1)$.
La matrice obtenue à partir de ces lignes vaut $\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{smallmatrix}\right)$.
Il faut lui ajouter une ligne afin de la rendre carrée inversible. Pour cela, on place un 1 en seconde colonne sur la troisième ligne. Il ne reste plus qu'à inverser la matrice pour obtenir la base orthogonale pour Q.

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -2 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}.$$

3.5 Interprétation matricielle

Le théorème précédent admet une interprétation matricielle dont l'énoncé ne fait pas référence à la théorie des formes quadratiques. Rappelons qu'une matrice est symétrique si elle est égale à sa transposée.

Théorème 5 (réduction des matrices symétriques) *Soit S une matrice symétrique. Alors il existe une matrice diagonale D dont les termes diagonaux appartiennent tous à $\{-1,0,1\}$ et une matrice inversible P telles que*

$$S = {}^tPDP.$$

Démontrons ce résultat. Considérons la forme quadratique sur $\mathbf{R}^n$ dont la matrice est donnée par S. Le théorème précédent donne une base (e_i) dans laquelle la forme quadratique a une matrice D diagonale, avec des coefficients diagonaux appartenant à $\{-1,0,1\}$. Notons P la matrice de passage de la base (e_i) à la base canonique, c'est la matrice dont les lignes sont données par les formes linéaires l_i. D'après la formule de changement de base pour les formes quadratiques, nous avons $S = {}^tPDP$.

Exemples
Les exemples précédents donnent

$$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} = {}^t\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = {}^t\begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

3.6 Étude de la signature

Il reste à démontrer la proposition 9 concernant la signature. La preuve repose sur le lemme suivant.

Lemme 1 *Soit E un espace vectoriel de dimension finie, Q une forme quadratique de signature (p,q). Supposons qu'il existe des sous-espaces vectoriels E_1, E_2 de E tels que*

- $E_1 + E_2 = E$,
- *Q est définie positive en restriction à E_1,*
- *Q est négative en restriction à E_2.*

Alors $E_1 \cap E_2 = \{0\}$ et $p = dim\ E_1$.

Démonstration du lemme
La forme quadratique Q est à la fois définie positive et négative sur $E_1 \cap E_2$. On a donc $Q(x) > 0$ et $Q(x) \leq 0$ pour tout vecteur x non nul dans cette intersection, ce qui montre qu'elle ne contient que le vecteur nul. Les sous-espaces E_1 et E_2 sont en somme directe, on en déduit

$$dim\ E_1 = dim\ E - dim\ E_2.$$

Soit F un sous-espace de E tel que Q soit définie positive en restriction à F. Là encore, la forme Q est définie positive et négative sur l'intersection $F \cap E_2$, qui est donc restreinte à $\{0\}$. Ceci permet de conclure :

$$dim\ F \leq dim\ E - dim\ E_2 = dim\ E_1.$$

Démontrons à partir de ce lemme qu'une forme quadratique de la forme

$$Q(x_1 e_1 + \ldots + x_n e_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$$

est nécessairement de signature (p,q). Pour cela, on remarque que Q est définie positive sur l'espace vectoriel E_1 engendré par $(e_1, \ldots, e_p)$ et négative sur l'espace E_2 engendré par $(e_{p+1}, \ldots, e_n)$. Par le lemme, on voit que p correspond bien au premier coefficient de la signature. On applique le même raisonnement à la forme quadratique $-Q$ qui a pour signature (q,p) afin de conclure.

Finalement, notons que la matrice de Q est diagonale. Les p premiers coefficients diagonaux sont égaux à un, les q suivants valent -1 et les autres sont nuls. Une telle matrice a un rang égal à $p+q$ et un noyau de dimension $n-p-q$. On en déduit le corollaire suivant.

Corollaire 5 *Soit E un espace vectoriel de dimension finie n, Q une forme quadratique de signature (p,q). Alors*

$$p + q = rang(Q), \quad p + q + dim \ker(Q) = n.$$

Méthode : *calcul de la signature d'une forme quadratique*

Il suffit de réduire la forme quadratique et de compter le nombre de signes $+$ et de signes $-$ devant les carrés des formes linéaires apparaissant dans l'expression réduite. Il est crucial que ces formes linéaires soient indépendantes entre elles.

Il est parfois possible de déterminer la signature d'une forme quadratique sans utiliser l'algorithme de réduction de Gauss, en faisant appel au déterminant. La méthode est expliquée dans les compléments de ce chapitre.

4. Coniques et quadriques affines

4.1 Coniques

Définition 19 *Une* conique $\mathcal{C}$ *est un sous-ensemble du plan composé des points dont les coordonnées $(x,y) \in \mathbf{R}^2$ dans une certaine base satisfont une équation de la forme*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

avec a, b, c, d, e et f des nombres réels tels que a, b et c ne soient pas tous les trois nuls.

Nous allons montrer qu'un changement de variables judicieux permet de simplifier l'équation de la conique.

Proposition 12 *Considérons une conique $\mathcal{C}$. Alors il existe un changement de variables affine de la forme*

$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + b_2 \end{cases}$$

qui transforme l'expression de la conique en une des formes suivantes :

$X^2 + Y^2 = 1$ — *la conique est une ellipse,*

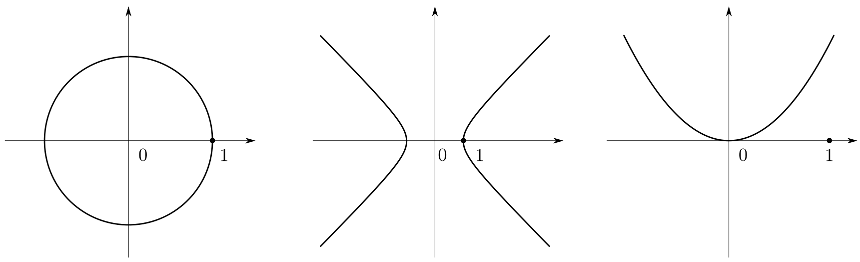
$X^2 - Y^2 = 1$ — *la conique est une hyperbole,*

$X^2 - Y = 0$ — *la conique est une parabole,*

ou une des formes :

$$X^2 = 1,\ X^2 - Y^2 = 0,\ X^2 = 0,\ X^2 + Y^2 = 0,\ X^2 + Y^2 = -1,\ X^2 = -1,$$

auquel cas la conique est composée respectivement de deux droites parallèles, deux droites sécantes, une droite, un point ou est vide.



$X^2 + Y^2 = 1$ $\quad$ $X^2 - Y^2 = 1$ $\quad$ $X^2 - Y = 0$

On peut déterminer de quelle conique il s'agit en regardant la signature de la forme quadratique $Q(x,y) = ax^2 + bxy + cy^2$ ou le *discriminant* associé $\Delta = ac - \frac{b^2}{4}$.

signature (2,0) $\rightarrow$ *ellipse ou point ou vide* ($\Delta > 0$)

signature (1,1) $\rightarrow$ *hyperbole ou deux droites sécantes* ($\Delta < 0$)

signature (1,0) $\rightarrow$ *parabole ou deux droites ou droite ou vide* ($\Delta = 0$)

Pour distinguer, par exemple, entre une ellipse, un point et l'ensemble vide, il suffit d'exhiber au moins deux points appartenant à l'ellipse.

Le discriminant d'une forme quadratique est égal au déterminant de la matrice associée à la forme quadratique. Attention à ne pas le confondre avec le discriminant d'un polynôme. Le lien avec la signature sera expliqué en complément.

Méthode : *réduction de l'équation d'une conique*
– On effectue le changement de variable qui permet de réduire la forme quadratique $Q(x,y) = ax^2 + bxy + cy^2$.
– Quitte à intervertir les deux variables, l'équation de la conique prend une des trois formes suivantes :

$$
\begin{array}{ll}
x'^2 + y'^2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (2,0) \text{ ou } (0,2) \\
x'^2 - y'^2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (1,1) \\
x'^2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (1,0) \text{ ou } (0,1)
\end{array}
$$

Dans les deux premiers cas, on fait disparaître les termes de degré 1 en mettant les polynômes $x'^2 + \alpha x'$ et $\pm y'^2 + \beta y'$ sous forme canonique. De même pour x' dans le dernier cas. Il reste alors un terme ou bien constant auquel cas on prend $Y = y'$, ou bien de degré un en y' et on l'affecte à Y.

Il est aussi possible d'appliquer directement l'algorithme de réduction de Gauss à l'expression de la conique tout entière, pour obtenir plus rapidement le changement de variables.

Exemple
– Déterminons la nature de la conique

$$\mathcal{C} = \{(x,y) \in \mathbf{R}^2 \mid x^2 + 5y^2 + 2xy + 2x + 10y + 4 = 0\}.$$

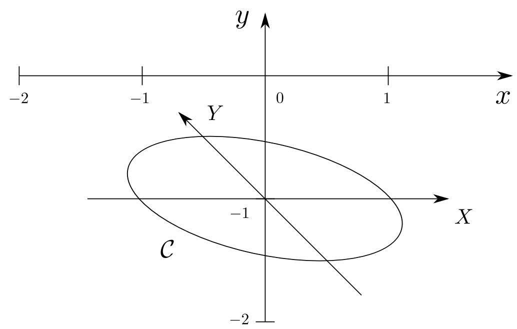
Par réduction,

$$
\begin{aligned}
&x^2 + 5y^2 + 2xy + 2x + 10y + 4 \\
&\qquad = (x+y)^2 + (2y)^2 + 2x + 10y + 4 \\
&\qquad = x'^2 + y'^2 + 2x' + 4y' + 4 \\
&\qquad = x'^2 + 2x' + y'^2 + 4y' + 4 \\
&\qquad = (x'+1)^2 + (y'+2)^2 - 1 \;=\; X^2 + Y^2 - 1
\end{aligned}
$$

en posant

$$
\left\{\begin{array}{lcr} x' & = & x+y \\ y' & = & 2y \end{array}\right., \qquad
\left\{\begin{array}{lclcl} X & = & x'+1 & = & x + y + 1 \\ Y & = & y'+2 & = & 2y + 2 \end{array}\right.
$$

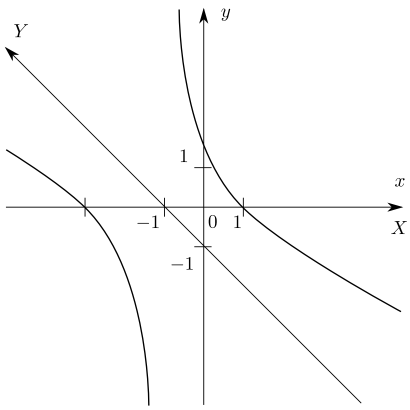
La conique est une ellipse centrée au point de coordonnées $X = 0$, $Y = 0$ ou $x = 0$, $y = -1$.



Appliquons directement l'algorithme de réduction à toute l'équation.
$$\begin{aligned}
&x^2+5y^2+2xy+2x+10y+4\\
&\qquad = x^2+x(2y+2)+5y^2+10y+4\\
&\qquad = (x+y+1)^2-(y+1)^2+5y^2+10y+4\\
&\qquad = (x+y+1)^2+4y^2+8y+3\\
&\qquad = (x+y+1)^2+(2y+2)^2-1 \;=\; X^2+Y^2-1.
\end{aligned}$$
On obtient bien sûr le même résultat.

– Pour la conique $\mathcal{C}=\{(x,y)\in\mathbf{R}^2 \mid x^2+2xy+2y+2x-3=0\}$,
$$\begin{aligned}
&x^2+2xy+2y+2x-3\\
&\qquad = (x+y)^2-y^2+2y+2x-3\\
&\qquad = x'^2-y'^2+2y+2x-3\\
&\qquad = x'^2+2x'-y'^2-3\\
&\qquad = (x'+1)^2-y'^2-4 \;=\; 4\,(X^2-Y^2-1)
\end{aligned}$$
avec $\begin{cases} x' &= x+y\\ y' &= y\end{cases}$, $\begin{cases} X &= \frac{1}{2}(x'+1) &= \frac{1}{2}(x+y+1)\\ Y &= \frac{1}{2}y' &= \frac{1}{2}y\end{cases}$

La conique est une hyperbole.



4.2 Quadriques

La même méthode fonctionne en toute dimension. Traitons le cas de la dimension 3.

Définition 20 *Une* quadrique $\mathcal{Q}$ *est un sous-ensemble de l'espace composé des points dont les coordonnées* $(x,y,z)\in\mathbf{R}^3$ *dans une certaine base satisfont une équation de la forme*

$$ax^2+by^2+cz^2+dxy+exz+fyz+gx+hy+iz+j=0$$

avec des coefficients réels tels que a*,* b*,* c*,* d*,* e *et* f *ne soient pas tous nuls.*

La méthode précédente donne le résultat suivant.

Proposition 13 *Considérons une quadrique* $\mathcal{Q} \subset \mathbf{R}^3$. *Alors il existe un changement de variables affine de la forme*

$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + a_{1,3}z + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + a_{2,3}z + b_2 \\ Z &= a_{3,1}x + a_{3,2}y + a_{3,3}z + b_3 \end{cases}$$

qui transforme l'expression de la quadrique en une des formes suivantes :

$X^2 + Y^2 + Z^2 = 1$ *ellipsoïde,*

$X^2 + Y^2 - Z^2 = 1$ *hyperboloïde à une nappe,*

$X^2 + Y^2 - Z^2 = -1$ *hyperboloïde à deux nappes,*

$X^2 + Y^2 - Z^2 = 0$ *cône,*

$X^2 + Y^2 - Z = 0$ *paraboloïde elliptique,*

$X^2 - Y^2 - Z = 0$ *paraboloïde hyperbolique,*

ou une des formes suivantes, qui correspondent au cas où la quadrique est le produit d'une conique par une droite :

$X^2 + Y^2 = 1$ *cylindre elliptique,*

$X^2 - Y^2 = 1$ *cylindre hyperbolique,*

$X^2 - Y = 0$ *cylindre parabolique,*

ou enfin une des formes suivantes :

$$X^2 = 1,\ X^2 - Y^2 = 0,\ X^2 + Y^2 + Z^2 = 0,\ X^2 + Y^2 = 0,\ X^2 = 0,$$

$$X^2 + Y^2 + Z^2 = -1,\ X^2 + Y^2 = -1,\ X^2 = -1,$$

auquel cas la quadrique correspond à deux plans parallèles, deux plans sécants, un point, une droite, un plan ou est vide.

Là encore, on peut s'aider de la signature de la forme quadratique

$$Q(x,y,z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

pour déterminer la nature de la quadrique.

signature (3,0) $\rightarrow$ *ellipsoïde, point ou vide ;*

signature (2,1) $\rightarrow$ *hyperboloïde à une ou deux nappes, cône ;*

signature (2,0) $\rightarrow$ *paraboloïde ou cylindre elliptique, droite, vide ;*

signature (1,1) $\rightarrow$ *paraboloïde ou cylindre hyperbolique, deux plans sécants ;*

signature (1,0) $\rightarrow$ *cylindre parabolique, deux plans parallèles, plan, vide.*

Exemple
On considère la quadrique $\mathcal{Q}$ de $\mathbf{R}^3$ donnée par

$$\mathcal{Q} = \{(x,y,z) \in \mathbf{R}^3 \mid x^2 + 3\,y^2 + z^2 + 4\,x\,y + 4\,y\,z - 10z = 0\}.$$

Déterminons la nature de cette quadrique. La réduction donne

$$\begin{aligned}
&x^2 + 3\,y^2 + z^2 + 4\,x\,y + 4\,y\,z - 10z \\
&\qquad = (x+2y)^2 - (y-2z)^2 + 5z^2 - 10z \\
&\qquad = (x+2y)^2 - (y-2z)^2 + 5(z-1)^2 - 5 \\
&\qquad = (x+2y)^2 - (y-2z)^2 + 5(z-1)^2 - 5 \\
&\qquad = 5\Big(\frac{1}{5}(x+2y)^2 - \frac{1}{5}(y-2z)^2 + (z-1)^2 - 1\Big) \\
&\qquad = 5\Big((\frac{x+2y}{\sqrt{5}})^2 - (\frac{y-2z}{\sqrt{5}})^2 + (z-1)^2 - 1\Big).
\end{aligned}$$

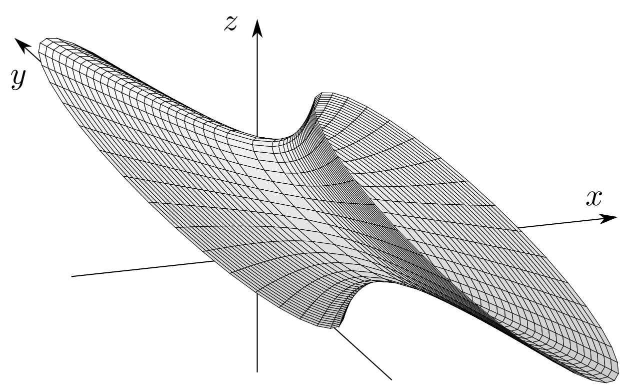
On pose donc

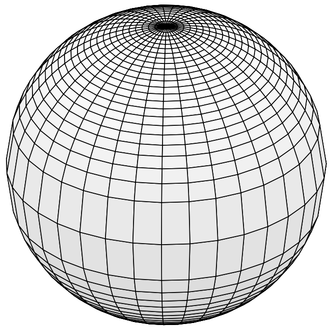
$$\begin{cases} X &= \frac{\sqrt{5}}{5}(x+2y) \\ Y &= z-1 \\ Z &= \frac{\sqrt{5}}{5}(y-2z) \end{cases}$$

et l'équation devient

$$X^2 + Y^2 - Z^2 = 1.$$

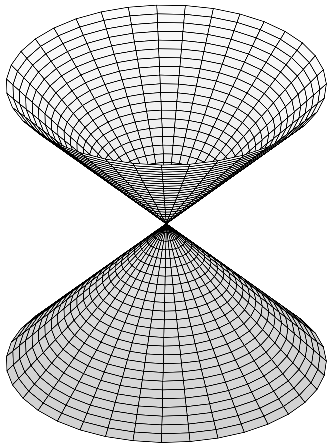
La quadrique $\mathcal{Q}$ est un hyperboloïde à une nappe.





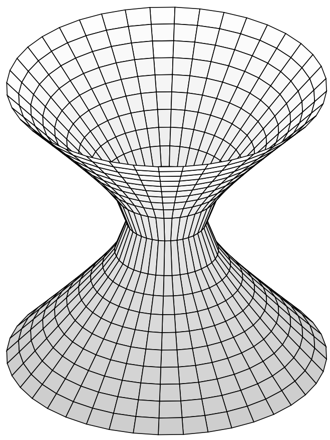
Sphère

$X^2 + Y^2 + Z^2 = 1$



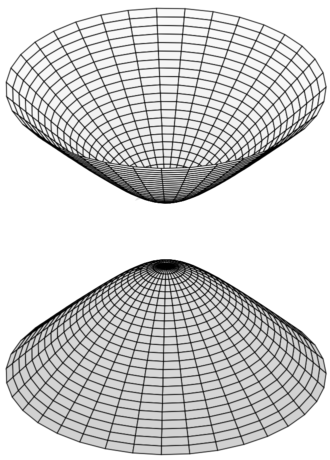
Cône

$X^2 + Y^2 - Z^2 = 0$



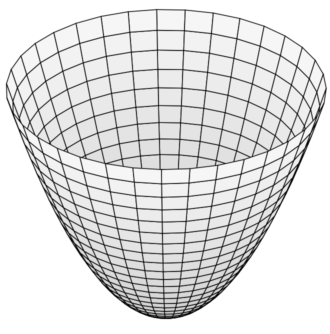
Hyperboloïde à une nappe

$X^2 + Y^2 - Z^2 = 1$



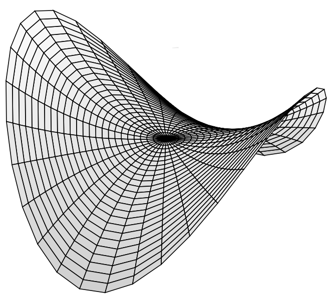
Hyperboloïde à deux nappes

$X^2 + Y^2 - Z^2 = -1$



Paraboloïde elliptique

$X^2 + Y^2 - Z = 0$



Paraboloïde hyperbolique

$X^2 - Y^2 - Z = 0$

5. Compléments

5.1 Forme quadratique et déterminant

Rappelons qu'une forme quadratique est dite non dégénérée si son noyau est nul. De manière équivalente, le déterminant de sa matrice calculée dans une base quelconque est non nul.

En général, le déterminant de la matrice associée à une forme quadratique dépend de la base choisie pour la calculer. Rappelons que les matrices B et B' dans deux bases différentes (e_i) et (e'_i) sont reliées par la formule

$$B' = {}^tPBP.$$

On en déduit les égalités

$$\det(B') = \det({}^tP)\det(B)\det(P) = \det(P)^2\ \det(B)$$

et les deux déterminants sont différents en général. On note tout de même qu'ils ont même signe. Le signe du déterminant de la matrice associée à une forme quadratique ne dépend pas de la base choisie pour exprimer cette matrice.

Examinons comment ce signe est relié à la signature (p,q) de la forme quadratique Q en se plaçant dans une base orthogonale pour Q. Dans une telle base, la matrice de Q possède p coefficients strictement positifs et q coefficients négatifs. Le déterminant est égal au produit de ces coefficients.

Proposition 14 *Une forme quadratique de signature (p,q) est non dégénérée si et seulement si le déterminant de sa matrice est non nul. Dans ce cas, le signe de ce déterminant est strictement négatif si et seulement si q est impair.*

On en déduit des informations importantes sur la signature en petite dimension. Par exemple, en dimension 2, si le déterminant de la matrice d'une forme quadratique dans une certaine base est strictement négatif, la signature vaut nécessairement (1,1) et s'il est strictement positif, elle vaut (2,0) ou (0,2).

On distingue entre ces deux derniers cas en regardant les coefficients diagonaux de la matrice de Q. Ces coefficients correspondent aux valeurs que prend Q sur les vecteurs de la base dans laquelle sa matrice est calculée. Si Q a pour signature (2,0), ils sont tous strictement positifs.

Proposition 15 *Soit E un espace vectoriel de dimension 2, Q une forme quadratique non dégénérée définie sur E et $B = \{b_{i,j}\}$ sa matrice dans une base de E. Alors la signature de Q vaut*

- (1,1) *si* $\det(B) < 0$,
- (2,0) *si* $\det(B) > 0$ *et* $b_{1,1} > 0$,
- (0,2) *si* $\det(B) > 0$ *et* $b_{1,1} < 0$.

5.2 Classification des formes quadratiques

Les théorèmes précédents peuvent se formuler sous forme d'un résultat de classification. Introduisons une relation d'équivalence sur l'ensemble des formes quadratiques définies sur un espace donné.

Proposition 16 *Deux formes quadratiques Q et Q' définies sur un espace vectoriel E sont dites équivalentes s'il existe une application linéaire inversible $f : E \to E$ telle que $Q'(x) = Q(f(x))$ pour tout $x \in E$.*

Deux formes quadratiques équivalentes ont même signature car l'application f met en bijection les espaces sur lesquelles les deux formes sont définies positives ou définies négatives. On va montrer la réciproque. Le théorème suivant est dû à James Joseph Sylvester (1814-1897).

Théorème 6 (loi d'inertie de Sylvester) *Deux formes quadratiques définies sur un espace vectoriel de dimension finie sont équivalentes si et seulement si elles ont même signature.*

Démonstration
Soit Q et Q' deux formes quadratiques ayant même signature. On a vu qu'il existe deux bases (e_i) et (e'_i) telles que

$$Q(x_1e_1 + \ldots + x_ne_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2,$$
$$Q'(x_1e'_1 + \ldots + x_ne'_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2.$$

Notons f l'application qui envoie la base (e_i) sur la base (e'_i) et posons $x = x_1e_1 + \ldots + x_ne_n$. On a alors $f(x) = x_1e'_1 + \ldots + x_ne'_n$, ce qui implique

$$Q'(f(x)) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2 = Q(x).$$

C'est l'égalité désirée.

Comme application, donnons la classification des formes quadratiques définies sur un espace de dimension deux. Il suffit d'énumérer les valeurs possibles pour la signature (p,q) et de donner un exemple de forme quadratique pour chacune de ces valeurs.

Proposition 17 *Soit E un espace vectoriel de dimension deux et Q une forme quadratique sur E. Alors il existe une base (e_1,e_2) de E dans laquelle Q prend l'une des formes suivantes :*

- $Q(x_1e_1 + x_2e_2) = 0$ *signature* $(0,0)$
- $Q(x_1e_1 + x_2e_2) = x_1^2$ *signature* $(1,0)$
- $Q(x_1e_1 + x_2e_2) = -x_1^2$ *signature* $(0,1)$
- $Q(x_1e_1 + x_2e_2) = x_1^2 + x_2^2$ *signature* $(2,0)$
- $Q(x_1e_1 + x_2e_2) = -x_1^2 - x_2^2$ *signature* $(0,2)$
- $Q(x_1e_1 + x_2e_2) = x_1^2 - x_2^2$ *signature* $(1,1)$

5.3 Équations cartésiennes et dualité

Nous avons vu au premier chapitre comment passer d'une représentation paramétrique d'un sous-espace vectoriel à sa représentation cartésienne. Faisons le lien avec la notion d'espace dual introduite en début de chapitre.

Définition 21 *Soit E un espace vectoriel de dimension finie et $F \subset E$ un sous-espace vectoriel. L'*annulateur *de F est le sous-espace vectoriel de l'espace dual E^* défini par*

$$F^0 = \{l \in E^* \mid \forall\ x \in E,\ l(x) = 0\}.$$

Cet espace F^0 est constitué de toutes les formes linéaires s'annulant sur F. Considérons une base $(e_1, \ldots, e_k)$ de F que nous interprétons comme un système d'équations paramétriques pour F. Un vecteur $x \in E$ est dans F si et seulement si on peut trouver λ_1, ..., λ_k tels que $x = \sum_{i \leq k} \lambda_i e_i$.

Proposition 18 *Soit E un espace vectoriel de dimension finie n, F un sous-espace de E de dimension k, $(e_1, \ldots, e_k)$ une base de F complétée en une base $(e_1, \ldots, e_k, \ldots, e_n)$ de E. Alors un vecteur $x \in E$ est dans F si et seulement si $e_j^*(x) = 0$ pour tout $j > k$. De plus, la famille $(e_{k+1}^*, \ldots, e_n^*)$ est une base de F^0.*

Les formes linéaires e_{k+1}^*, ..., e_n^* forment donc un système d'équations cartésiennes pour F.

Démonstration
Un vecteur $x \in E$ a pour décomposition dans la base $(e_1, \ldots, e_n)$

$$x = \sum_{i \leq n} e_i^*(x)\ e_i.$$

Il appartient à $F = \mathrm{vect}(e_1, \ldots, e_k)$ si et seulement si ses $n - k$ dernières coordonnées sont nulles, c'est-à-dire si et seulement si $e_j^*(x) = 0$ pour $j > k$. De plus, pour $j > k$, les formes e_j^* s'annulent sur chacun des e_i, $i \leq k$. Elles s'annulent donc sur F et appartiennent à F^0. Il reste à voir qu'elles l'engendrent. Soit $l \in F^0$, par le corollaire 4 nous pouvons écrire

$$l = \sum_{i \leq n} l(e_i)\, e_i^*.$$

Pour $i \leq k$, le vecteur e_i est dans F d'où $l(e_i) = 0$ et $l = \sum_{i>k} l(e_i) e_i^*$.

La démonstration précédente montre également qu'un vecteur $x \in E$ est dans F si et seulement si $l(x) = 0$ pour tout $l \in F^0$. On en déduit que toute base de F^0 donne un système d'équations cartésiennes pour F.

Corollaire 6 *Soit $(l_1, \ldots, l_m)$ une base de F^0, alors un vecteur x appartient à F si et seulement si $l_1(x) = 0$, ..., $l_m(x) = 0$.*

5.4 Construction géométrique de bases orthogonales

Nous avons construit des bases orthogonales pour les formes quadratiques grâce à l'algorithme de réduction de Gauss. C'est la proposition 11.

Toute forme bilinéaire symétrique définie sur un espace vectoriel de dimension finie admet une base orthogonale.

Voici une autre preuve de cette proposition, de nature géométrique.

Démonstration
Raisonnons par récurrence sur la dimension n de E. Notons Q la forme quadratique et φ sa forme polaire. Si Q n'est pas identiquement nulle, prenons un vecteur $e_n \in E$ tel que $Q(e_n) \neq 0$. Soit $l \in E^*$ la forme linéaire définie par $l(x) = \varphi(x, e_n)$. Le sous-espace $H = \ker(\varphi)$ est de dimension $n-1$ car $l(e_n)$ est non nul. Par hypothèse de récurrence, la restriction de Q à l'hyperplan H admet une base orthogonale $(e_1, \ldots, e_{n-1})$. Les relations

$$e_n \notin H \quad \text{et} \quad l(e_i) = 0 = \varphi(e_i, e_n) \text{ pour } i < n$$

montrent que la famille $(e_1, \ldots, e_n)$ est une base orthogonale de E. La preuve est terminée.

On en déduit que toute forme quadratique se décompose en carrés de formes linéaires. Notons e_i^* les vecteurs de la base duale de $(e_1, \ldots, e_n)$. Par orthogonalité,

$$Q(x_1 e_1 + \ldots + x_n e_n) = \sum_i Q(e_i)\, x_i^2 = \sum_i \operatorname{signe}(Q(e_i)) \Big(\sqrt{Q(e_i)}\, e_i^*(x)\Big)^2$$

ce qui donne la décomposition recherchée.

La démonstration précédente est plus courte que celle reposant sur la réduction de Gauss, mais il est malaisé de calculer à la main les formes linéaires par cette méthode. Les raisonnements géométriques précédents se prêtent par contre bien à une implémentation sur ordinateur et on peut en déduire un algorithme explicite pour calculer une base orthogonale, en programmant de manière récursive. La récursion nécessite la construction d'une base de l'hyperplan H pour exprimer la forme Q en restriction à H, nous avons vu comment obtenir une telle base à la fin du premier chapitre.

Chapitre 3

Espaces euclidiens

Nous étudions dans ce chapitre la notion d'espace euclidien, en nous appuyant sur les résultats généraux concernant les formes quadratiques vus au chapitre précédent. Une fois introduite la notion de produit scalaire, nous montrons comment il intervient en géométrie plane avant de généraliser les concepts d'aire, longueur et angle à des espaces euclidiens de dimension quelconque. On étudie ensuite la notion d'orthogonalité et les transformations de l'espace euclidien associées, projections orthogonales et isométries.

1. Produit scalaire

1.1 Définition d'un produit scalaire

Définition 22 *Un* produit scalaire *défini sur un espace vectoriel réel est la donnée d'une forme bilinéaire symétrique sur cet espace dont la forme quadratique associée est définie positive.*

Un espace euclidien *est un espace vectoriel réel muni d'un produit scalaire.*

Il est d'usage d'employer une notation particulière pour le produit scalaire de deux vecteurs $x,y \in E$: $\langle x,y\rangle$, $(x \mid y)$ où $\langle x \mid y\rangle$. On pose également

$$\langle x,x\rangle = \|x\|^2.$$

La quantité $\|x\| = \sqrt{\langle x,x\rangle}$ est la *norme euclidienne* du vecteur x.

Exemples

– L'espace $\mathbf{R}^n$ muni du produit scalaire usuel est bien sûr un espace euclidien.

– L'espace $C^0([0,1],\mathbf{R})$ muni de la forme bilinéaire symétrique

$$(f,g) \mapsto \int_0^1 f(t)g(t)\,dt$$

est un espace euclidien. Il n'est pas de dimension finie.

Lorsque l'espace est de dimension finie, on peut trouver une base orthogonale en vertu du théorème 11. Comme le produit scalaire est défini positif, on peut diviser chacun des vecteurs de cette base par sa norme euclidienne de manière à ce qu'ils soient tous de norme un.

Proposition 19 *Soit* $(E,\langle .,. \rangle)$ *un espace euclidien de dimension finie* n*. Alors il existe une base* $(e_1, \ldots, e_n)$ *de* E *qui vérifie*

$$\langle e_i,e_j \rangle = 0 \text{ pour tout } i,j \text{ distincts}, \quad \|e_i\| = 1 \text{ pour tout } i.$$

Une telle base est dite orthonormée*. Dans cette base, tout vecteur* $x \in E$ *a pour coefficients* $x_i = \langle x,e_i \rangle$*,*

$$x = \sum_{i=1}^{n} x_i\, e_i = \sum_{i=1}^{n} \langle x,e_i \rangle\, e_i = \langle x,e_1 \rangle\, e_1 + \ldots + \langle x,e_n \rangle\, e_n,$$

le produit scalaire et la forme quadratique sont de la forme

$$\langle x,y \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n,$$

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 = x_1^2 + \ldots + x_n^2.$$

Grâce aux formules précédentes, on peut exprimer les coefficients de la matrice associée à une application linéaire $f : E \to E$ dans une base orthonormée (e_i) à l'aide produit scalaire. Ces coefficients sont les coordonnées des vecteurs $f(e_j)$ dans la base (e_i). La matrice de f est donc égale à $\{\langle e_i,f(e_j) \rangle\}_{i,j}$.

1.2 Norme euclidienne

Vérifions que la fonction $x \mapsto \|x\|$ est bien une norme, ce qui signifie qu'elle satisfait les propriétés suivantes.

Proposition 20 *Soit* $(E,\langle .,. \rangle)$ *un espace euclidien. Alors*

- *pour tout* $\lambda \in \mathbf{R}$*,* $x \in E$*,* $\quad \|\lambda x\| = |\lambda| \|x\|$*,*
- *pour tout* $x,y \in E$*,* $\quad \|x + y\| \leq \|x\| + \|y\|$*,* $\quad$ *(inégalité triangulaire)*
- *pour tout* $x \in E$*,* $\quad \|x\| = 0$ *si et seulement si* $x = 0$*.*

La démonstration de l'inégalité triangulaire repose sur l'inégalité suivante, attribuée à Augustin Louis Cauchy (1789-1857) et Hermann Amandus Schwarz (1843-1921).

Théorème 7 (inégalité de Cauchy-Schwarz) *Soit* $(E,\langle .,. \rangle)$ *un espace euclidien et* $x,y \in E$*. Alors*

$$|\langle x,y \rangle| \leq \|x\| \|y\|$$

avec égalité si et seulement si x *et* y *sont proportionnels entre eux.*

Démonstration de l'inégalité de Cauchy-Schwarz
Si x et y sont proportionnels entre eux, par exemple $x = \lambda y$, les deux termes de l'inégalité valent $\lambda\|x\|^2$. Sinon, x et y engendrent un plan vectoriel vect(x,y), en restriction duquel la matrice du produit scalaire est donnée par

$$B = \begin{pmatrix} \langle x,x\rangle & \langle x,y\rangle \\ \langle y,x\rangle & \langle y,y\rangle \end{pmatrix}.$$

Choisissons une base orthonormée de vect(x,y). On a vu que la matrice B est de la forme ${}^tPP$, où P est la matrice de passage de cette base orthonormée à la base (x,y). On a donc

$$\|x\|^2\|y\|^2 - \langle x,y\rangle^2 = \det(B) = \det(P)^2 > 0.$$

Exemples
– Appliquons l'inégalité aux deux vecteurs (x_1,x_2) et (y_1,y_2) de $\mathbf{R}^2$ muni du produit scalaire usuel. On obtient

$$|x_1y_1 + x_2y_2| \leq \sqrt{x_1^2 + x_2^2}\ \sqrt{y_1^2 + y_2^2}.$$

– Appliquons l'inégalité à deux fonctions continues $f,g : [0,1] \to \mathbf{R}$, en utilisant le produit scalaire $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$. On obtient

$$\Big|\int_0^1 f(t)g(t)\,dt\Big| \leq \sqrt{\int_0^1 f(t)^2\,dt}\ \sqrt{\int_0^1 g(t)^2\,dt}.$$

Démonstration de l'inégalité triangulaire

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

Revenons sur l'inégalité de Cauchy-Schwarz. On voit que même si cette inégalité est vraie dans tout espace euclidien, c'est avant tout un résultat de géométrie plane. On va en donner une interprétation en terme d'aire.

1.3 Aire, longueur et angle

Définition 23 *Soit $(E,\langle .,.\rangle)$un espace euclidien et $x,y \in E$. Si x et y sont liés, on pose $aire(x,y) = 0$. Sinon, soit (e_1,e_2) une base orthonormée de* vect(x,y) *et (x_1,x_2), (y_1,y_2) les coordonnées de x et y dans cette base. L'*aire *du parallélogramme de côtés x et y est définie par*

$$aire(x,y) = |x_1y_2 - x_2y_1|,$$

elle ne dépend pas de la base orthonormée (e_1,e_2) utilisée pour la calculer.

Cette aire est égale à la valeur absolue du déterminant de la matrice P considérée dans la preuve du théorème précédent. De fait, P est la matrice de passage de la base (e_1,e_2) à la base (x,y), elle s'écrit donc

$$P = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

La formule vue plus haut, $\det(B) = \det(P)^2$, devient

Proposition 21 *Soit* $(E,\langle .,. \rangle)$*un espace euclidien et* $x,y \in E$. *Alors*

$$\langle x,y \rangle^2 + aire(x,y)^2 = \|x\|^2 \|y\|^2.$$

Cette formule montre que l'aire est indépendante de la base orthonormée choisie pour la calculer. Par contre, le signe du déterminant $\det(P) = x_1y_2 - x_2y_1$ dépend de cette base. Cela nous amènera à introduire la notion d'espace vectoriel euclidien orienté.

Exemple
Considérons deux vecteurs (x_1,x_2) et (y_1,y_2) de $\mathbf{R}^2$ muni du produit scalaire usuel. La proposition entraîne l'égalité

$$(x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

Cette formule s'appelle *identité de Lagrange*, elle est associée au nom du mathématicien Joseph-Louis de Lagrange (1736-1813). On peut la démontrer directement en développant les carrés.

Méthode : *calcul de l'aire d'un parallélogramme de* $\mathbf{R}^n$

Pour calculer l'aire du parallélogramme dont les côtés sont donnés par deux vecteurs x, y de $\mathbf{R}^n$, on pourrait construire une base orthogonale du plan $\mathrm{vect}(x,y)$ comme expliqué au précédent chapitre, exprimer x et y dans cette base et calculer un déterminant. Il est plus rapide d'utiliser la formule précédente.

Exemple
– Calculons l'aire du parallélogramme de $\mathbf{R}^3$ dont les côtés sont donnés par $x = (1,1,0)$ et $y = (0,1,1)$.

$\|x\|^2 = 1^2+1^2+0^2 = 2$, $\|y\|^2 = 0^2+1^2+1^2 = 2$, $\langle x,y \rangle = 1\times0+1\times1+0\times1 = 1$,

d'où $aire(x,y) = \sqrt{2 \times 2 - 1^2} = \sqrt{3}$.

L'inégalité de Cauchy-Schwarz montre que pour tous vecteurs $x,y \in E$, le réel $\langle x,y \rangle/(\|x\|\,\|y\|)$ est compris entre -1 et 1. Rappelons que pour tout réel a compris entre $[-1,1]$, il existe un unique nombre réel $\theta \in [0,\pi]$ tel que $a = \cos(\theta)$. Ceci permet de définir l'angle non orienté de deux vecteurs de la façon suivante.

Définition 24 (angles non orientés) *Soit* $(E,\langle .,. \rangle)$ *un espace euclidien et* $x,y \in E$ *non nuls. La mesure de l'angle non orienté défini par la paire* $\{x,y\}$ *est l'unique* $\theta \in [0,\pi]$ *satisfaisant*

$$\langle x,y \rangle = \|x\| \, \|y\| \, \cos(\theta).$$

Méthode : *calculs d'angles et de produits scalaires*

Pour calculer l'angle non orienté de deux vecteurs dont les coordonnées dans une base orthonormée sont données, on calcule leur norme, leur produit scalaire et on fait appel à la fonction inverse du cosinus, qu'on note *acos* dans la suite. Réciproquement, on peut calculer le produit scalaire de deux vecteurs si on connaît leurs normes et leur angle.

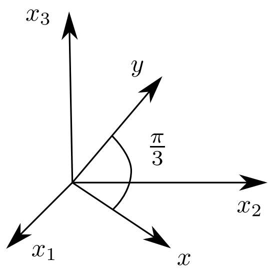
Exemples

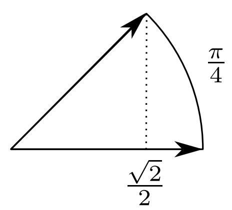
– Reprenons les deux vecteurs $x = (1,1,0)$ et $y = (0,1,1) \in \mathbf{R}^3$. Nous avons montré que $\|x\|^2 = 2$, $\|y\|^2 = 2$, $\langle x,y \rangle = 1$. La mesure de l'angle non orienté entre x et y est donc donnée par

$$\mathrm{acos}\Big(\frac{1}{\sqrt{2}\sqrt{2}}\Big) = \mathrm{acos}\Big(\frac{1}{2}\Big) = \frac{\pi}{3}.$$

– Considérons deux vecteurs x, y de norme un et faisant un angle égal à $\frac{\pi}{4}$.

$$\langle x,y \rangle = \cos\Big(\frac{\pi}{4}\Big) = \frac{\sqrt{2}}{2}.$$





On termine par quelques formules concernant les parallélogrammes et les triangles d'un espace euclidien.

Proposition 22 *Soit* $(E,\langle .,. \rangle)$ *un espace euclidien et* $x,y \in E$*. Alors*

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2 \qquad \textit{(formule du parallélogramme)}$$

$$\|x\|^2 + \|y\|^2 = \frac{1}{2}\|x-y\|^2 + 2\left\|\frac{x+y}{2}\right\|^2 \qquad \textit{(formule de la médiane)}$$

Pour interpréter géométriquement la première de ces formules, on considère un parallélogramme de sommets A, B, C et D situés dans un même plan et on prend pour x le vecteur allant de A à B, pour y le vecteur allant de A à D. On obtient

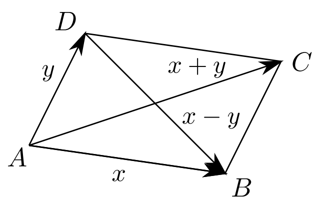
$$2AB^2 + 2BC^2 = AC^2 + BD^2.$$

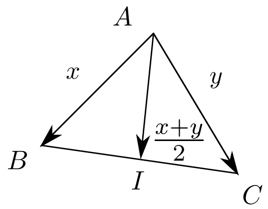
La somme des carrés des longueurs des quatre côtés d'un parallélogramme est égale à la somme des carrés des longueurs des diagonales.

Pour la seconde formule, on considère un triangle de sommets A, B et C et on note I le milieu de BC.

$$AB^2 + AC^2 = \frac{1}{2}BC^2 + 2AI^2.$$

Cette formule exprime le carré de la longueur de la médiane du triangle en fonction des carrés des côtés du triangle. Elle remonte à Thalès de Milet (-625 -547).





Démonstration de la formule du parallélogramme
Il suffit de développer les carrés dans le premier terme.

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle + \|x\|^2 + \|y\|^2 - 2\langle x,y\rangle \\ &= 2\,\|x\|^2 + 2\,\|y\|^2.\end{aligned}$$

Donnons enfin un analogue multidimensionnel du théorème de Pythagore.

Théorème 8 (de Pythagore) *Soit $(E,\langle .,.\rangle)$un espace euclidien et v_1, ..., v_n des vecteurs de E orthogonaux deux à deux: $\langle v_i,v_j\rangle = 0$ pour tout i, j distincts. Alors*

$$\|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2.$$

Démonstration

$$\|v_1 + \ldots + v_n\|^2 = \left\langle \sum_i v_i, \sum_j v_j \right\rangle = \sum_{i,j} \langle v_i,v_j\rangle = \sum_i \|v_i\|^2.$$

1.4 Espace euclidien orienté

Orientation d'un espace euclidien

Définition 25 *On se place dans un espace euclidien $(E,\langle .,.\rangle)$de dimension finie. Nous dirons que deux bases (e_i) et (e'_i) ont* même orientation *si la matrice de passage de l'une à l'autre base est de déterminant strictement positif. S'il est au contraire négatif, nous dirons que les bases ont des orientations opposées.*

Orienter un espace euclidien de dimension finie consiste à choisir une base qu'on dit orientée positivement, ou encore de manière directe. Toutes les bases ayant même orientation que la base choisie sont dites orientées positivement. Les autres sont dites orientées négativement, de manière indirecte, ou encore dans le sens rétrograde. Deux bases orientées de manière indirecte ont même orientation, en vertu des formules de changement de base.

Une application inversible définie d'un espace euclidien orienté dans lui-même *préserve l'orientation* si une base et son image ont même orientation. On verra plus loin comment caractériser cette propriété en terme de déterminant.

Exemples

– La base canonique définit l'orientation canonique de l'espace $\mathbf{R}^n$.

– Un plan de $\mathbf{R}^3$ hérite d'une structure euclidienne obtenue par restriction du produit scalaire usuel de $\mathbf{R}^3$ mais il ne possède pas d'orientation donnée a priori. Se donner une orientation d'un tel plan revient à choisir un vecteur e_3 de norme un orthogonal au plan. Une base (e_1,e_2) du plan est considérée comme directe si (e_1,e_2,e_3) est une base directe de $\mathbf{R}^3$ muni de son orientation canonique. Il existe deux choix possibles pour le vecteur e_3, chacun correspondant à une orientation différente.

Déterminant d'une famille de n vecteurs

La notion d'espace euclidien orienté de dimension n permet de définir le déterminant d'une famille de n vecteurs $(v_1, \dots, v_n)$. Il s'agit du déterminant de la matrice dont les colonnes sont données par les coordonnées des v_i dans une base orthonormée directe. On démontrera dans la suite qu'il ne dépend pas de la base orthonormée directe choisie.

Angles orientés dans le plan

Une autre application de la notion d'orientation est celle d'angle orienté. On se place dans un espace E euclidien orienté de dimension deux. Nous avons vu précédement que pour tout $x,y \in E$ non nuls,

$$\left(\frac{\langle x,y\rangle}{\|x\|\|y\|}\right)^2 + \left(\frac{\det(x,y)}{\|x\|\|y\|}\right)^2 = 1.$$

Rappelons que pour tous réels a, b satisfaisant $a^2 + b^2$, on peut trouver un réel θ, unique à un multiple de 2π près, tel que $a = \cos(\theta)$ et $b = \sin(\theta)$.

Définition 26 (angles orientés) *Soit $(E,\langle .,.\rangle)$un espace euclidien orienté de dimension deux et x, y deux vecteurs de E non nuls. La mesure de l'angle orienté défini par le couple (x,y) est l'unique $\theta \in \mathbf{R}/2\pi\mathbf{Z}$ satisfaisant*

$$\langle x,y\rangle = \|x\|\,\|y\|\,\cos(\theta),$$

$$\det(x,y) = \|x\|\,\|y\|\,\sin(\theta),$$

le déterminant étant calculé dans une base orthonormée orientée de E.

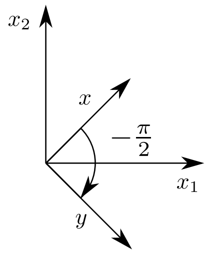
Exemple
Plaçons-nous dans $\mathbf{R}^2$ muni de son orientation canonique et considérons les vecteurs $x = (1,1)$ et $y = (1, -1)$. Nous avons

$$\|x\|^2 = 2,\ \|y\|^2 = 2,\ \langle x,y\rangle = 0,\ \det(x,y) = \det\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2.$$

La mesure θ de l'angle orienté formé par ces deux vecteurs vérifie

$$\cos(\theta) = 0/2 = 0, \quad \sin(\theta) = -2/2 = -1,$$

ce qui implique $\theta = -\frac{\pi}{2} \bmod 2\pi$.



2. Orthogonalité

On a introduit précédemment le concept de base orthonormée. On va étudier plus en détail la notion d'orthogonalité dans un espace euclidien.

Définition 27 *Soit $(E,\langle .,.\rangle)$un espace euclidien. Deux sous-espaces vectoriels F_1 et F_2 de E sont orthogonaux entre eux si tout vecteur de F_1 est orthogonal à tout vecteur de F_2 :* $\forall\, x \in F_1,\ \forall\, y \in F_2,\ \langle x,y\rangle = 0.$

2.1 Projection orthogonale

Commençons par introduire la notion d'espace orthogonal à un sous-espace donné avant de définir la projection orthogonale sur un sous-espace.

Définition 28 *Soit $(E,\langle .,.\rangle)$un espace euclidien et F un sous-espace vectoriel de E. L'*orthogonal *de F est le sous-espace vectoriel de E constitué des vecteurs qui sont orthogonaux à tous les vecteurs de F :*

$$F^\perp = \{x \in E \mid \forall\, y \in F,\ \langle x,y\rangle = 0\}$$

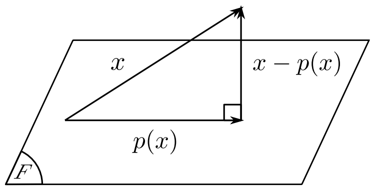
Un vecteur à la fois dans F et $F^\perp$ est orthogonal à lui-même, il est donc nul. On en déduit que $F \cap F^\perp = \{0\}$.

Proposition 23 *Soit $(E,\langle .,.\rangle)$un espace euclidien et F un sous-espace vectoriel de E de dimension finie. Alors pour tout $x \in E$, il existe un unique vecteur $p(x)$ qui satisfait les deux conditions suivantes :*

- *$p(x) \in F$,*
- *$x - p(x) \in F^\perp$.*

Le vecteur $p(x)$ est la projection orthogonale *de x sur F. Soit (e_i) une base orthonormée de F, alors*

$$p(x) = \sum_i \langle x,e_i\rangle\, e_i = \langle x,e_1\rangle\, e_1 + \ldots + \langle x,e_n\rangle\, e_n.$$



Démonstration
La somme $\sum_i \langle x,e_i\rangle\, e_i$ satisfait les deux propriétés voulues : elle est dans F et pour tout j,

$$\langle x - \sum_i \langle x,e_i\rangle\, e_i, e_j\rangle = \langle x,e_j\rangle - \langle x,e_j\rangle = 0.$$

Soit y, y' des vecteurs tels que y, $y' \in F$ et $x-y$, $x-y' \in F^\perp$. La différence $y - y' = (x-y') - (x-y)$ est à la fois dans F et dans $F^\perp$, elle est donc nulle.

Méthode : *calcul du projecteur orthogonal sur un sous-espace*

Il suffit de calculer une base orthonormée (e_i) de ce sous-espace et d'utiliser la formule $p(x) = \sum \langle x,e_i\rangle e_i$.

Remarquons que l'application $x \mapsto p(x)$ est linéaire. Posons

$$x_1 = p(x) \in F, \quad x_2 = x - p(x) \in F^\perp.$$

De manière remarquable, le vecteur x_2 est dans $F^\perp$ et vérifie $x - x_2 \in F$. Cela montre l'existence de la projection orthogonale de x sur $F^\perp$, donnée par $x - p(x)$. On a montré le résultat suivant.

Corollaire 7 *Tout $x \in E$ s'écrit de manière unique sous la forme*

$$x = x_1 + x_2 \quad \text{avec } x_1 \in F \text{ et } x_2 \in F^\perp.$$

On dit que E est la somme directe de F et $F^\perp$ et on note

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

La projection orthogonale sur F est l'unique application linéaire $p : E \to E$ satisfaisant

$$\forall\, x \in F, \quad p(x) = x, \quad \forall\, x \in F^\perp, \quad p(x) = 0.$$

Rappelons que la dimension de deux sous-espaces en somme directe est égale à la somme de leurs dimensions. De plus, on a toujours l'inclusion $F \subset (F^\perp)^\perp$. Si E est de dimension finie, le corollaire précédent montre que ces deux espaces ont même dimension. Ils sont donc égaux.

Corollaire 8 *Soit $(E,\langle .,. \rangle)$un espace euclidien de dimension finie et F un sous-espace vectoriel de E. Alors*

– *$\dim F + \dim F^\perp = \dim E$,*
– *$F = (F^\perp)^\perp$.*

Exemple
Considérons un hyperplan $H = \{x \in \mathbf{R}^n \mid a_1x_1 + \ldots + a_nx_n = 0\}$. Posons $v = (a_1, \ldots, a_n)$. Alors H est l'orthogonal de $\mathrm{vect}(v)$:

$$\mathrm{vect}(v)^\perp = H, \quad H^\perp = \mathrm{vect}(v).$$

Une base orthonormée de $\mathrm{vect}(v)$ est donnée par le singleton $(v/\|v\|)$. La projection sur $\mathrm{vect}(v)$ vaut donc

$$p(x) = \frac{\langle x,v\rangle}{\|v\|^2}\, v = \frac{\sum a_kx_k}{\sum a_k^2}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

La matrice de p a pour coefficients $\{\frac{a_ia_j}{\sum a_k^2}\}$, elle est donnée par

$$\frac{1}{\sum_{k=1}^n a_k^2}\begin{pmatrix} a_1^2 & a_1a_2 & \ldots & a_1a_n \\ \vdots & \vdots & & \vdots \\ a_na_1 & a_na_2 & \ldots & a_n^2 \end{pmatrix}.$$

La projection orthogonale sur H a pour expression

$$p_H(x) = x - \frac{\langle x,v\rangle}{\|v\|^2}\, v.$$

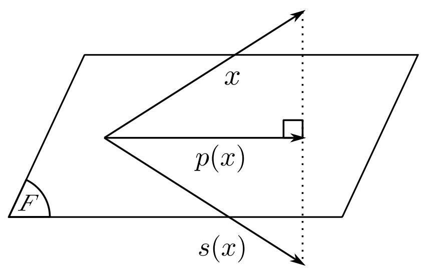
Terminons par une notion reliée à celle de projection orthogonale.

Définition 29 *Soit $(E,\langle .,.\rangle)$un espace euclidien, F un sous-espace de E de dimension finie. La* symétrie orthogonale *relativement à F est l'unique application linéaire $s : E \to E$ telle que*

$$\forall\, x \in F, \quad s(x) = x, \quad \forall\, x \in F^\perp, \quad s(x) = -x.$$

Lorsque F est un hyperplan, on parle de réflexion orthogonale*. Notons p la projection orthogonale sur F. Alors*

$$s(x) = 2p(x) - x.$$



Exemples
– La symétrie relativement au sous-espace $\{0\}$ est appelée *symétrie centrale*. Elle est donnée par $s(x) = -x$ pour tout x. La symétrie par rapport à E est égale à l'identité.

– La réflexion orthogonale relativement à un hyperplan $H = vect(v)^\perp$ est donnée par l'expression

$$s(x) = x - 2\,\frac{\langle x,v\rangle}{\|v\|^2}\,v.$$

– La symétrie orthogonale par rapport à une droite $\mathrm{vect}(v)$ est donnée par

$$s(x) = 2\,\frac{\langle x,v\rangle}{\|v\|^2}\,v - x.$$

En dimension 3, une symétrie orthogonale relativement à une droite est appelée un *demi-tour*.

2.2 Procédé d'orthonormalisation de Gram-Schmidt

Soit E un espace vectoriel muni d'une forme bilinéaire symétrique φ. Nous avons vu dans le chapitre portant sur les formes quadratiques (page 44) comment trouver une base orthogonale pour la forme φ.

– On commence par réduire la forme quadratique Q associée à φ. On obtient des formes linéaires $l_1, \ldots, l_{p+q}$ linéairement indépendantes telles que

$$Q = \sum \pm l_i^2.$$

– On complète la famille (l_i) en une base de E^* et on construit une base (e_i) de telle sorte que la base (l_i) soit duale à celle-ci.

Cette méthode générale est assez longue à mettre en œuvre. Dans le cas d'un produit scalaire, il existe un procédé plus rapide attribué à Jørgen Pedersen Gram (1850-1916) et Erhard Schmidt (1876-1959) mais qu'on trouve déjà dans les travaux de Pierre-Simon de Laplace (1749-1827).

Méthode : *procédé d'orthonormalisation de Gram-Schmidt*

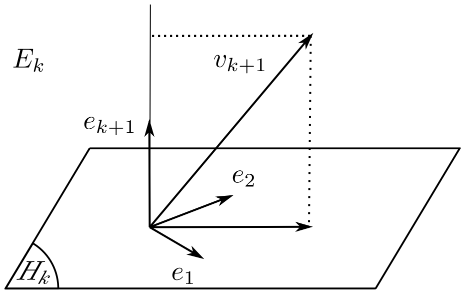
On part d'une base $(v_1, \ldots, v_n)$ de E et on cherche à construire par récurrence une base orthonormée $(e_1, \ldots, e_n)$ de façon à ce que les vecteurs $(v_1, \ldots, v_k)$ et $(e_1, \ldots, e_k)$ engendrent le même espace vectoriel pour tout k allant de 1 à n.

Dans ce cas, v_1 et e_1 sont nécessairement proportionnels et e_1 doit être de norme un, ce qui laisse une seule possibilité si on suppose $\langle e_1,v_1\rangle > 0$:

$$e_1 = \frac{v_1}{\|v_1\|}.$$

Supposons construits $(e_1, \ldots, e_k)$ et construisons e_{k+1}. Ce doit être un vecteur de norme un appartenant à $E_k = \text{vect}(e_1, \ldots, e_k, v_{k+1})$ orthogonal à l'espace $H_k = \text{vect}(e_1, \ldots, e_k)$. Il s'agit donc de projeter un vecteur de l'espace E_k sur l'orthogonal $H_k^\perp$ à l'hyperplan $H_k \subset E_k$ (vu page 67) et de le normaliser.

$$e_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle \, e_i}{\| v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle \, e_i \|}.$$



Théorème 9 (orthonormalisation de Gram-Schmidt) *Soit* $(E, \langle .,. \rangle)$ *un espace euclidien de dimension finie et* $(v_1, \ldots, v_n)$ *une base de* E*. Alors il existe une unique base orthonormée* $(e_1, \ldots, e_n)$ *telle que, pour tout entier* i *entre 1 et* n*,*

– $\text{vect}(e_1, \ldots, e_i) = \text{vect}(v_1, \ldots, v_i)$,

– $\langle e_i, v_i \rangle > 0$.

Exemple
Comparons les deux méthodes sur un exemple. On se place sur $\mathbf{R}^2$ et on prend $v_1 = (3,4)$ et $v_2 = (4,3)$. Nous avons

$$\|v_1\| = 5, \quad \|v_2\| = 5, \quad \langle v_1, v_2 \rangle = 24.$$

On obtient par la première méthode

$$\begin{aligned} Q(x_1 v_1 + x_2 v_2) &= 25x_1^2 + 25x_2^2 + 48 x_1 x_2 \\ &= 25\Big(x_1 + \tfrac{24}{25}\, x_2\Big)^2 + \tfrac{49}{25}\, x_2^2 \\ &= l_1(x_1,x_2)^2 + l_2(x_1,x_2)^2. \end{aligned}$$

avec $l_1(x_1,x_2) = 5x_1 + \frac{24}{5}\, x_2$ et $l_2(x_1,x_2) = \frac{7}{5} x_2$.

$$\begin{pmatrix} 5 & \frac{24}{5} \\ 0 & \frac{7}{5} \end{pmatrix}^{-1} = \frac{1}{7} \begin{pmatrix} \frac{7}{5} & -\frac{24}{5} \\ 0 & 5 \end{pmatrix} = \frac{1}{35} \begin{pmatrix} 7 & -24 \\ 0 & 25 \end{pmatrix}.$$

Les coordonnées obtenues sont celles d'une base orthogonale (e_1,e_2) dans la base (v_1,v_2).

$$e_1 = \frac{1}{5}\, v_1 = \frac{1}{5}\begin{pmatrix}3\\4\end{pmatrix},$$

$$e_2 = \frac{1}{35}(-24v_1+25v_2) = \frac{1}{35}\begin{pmatrix}-24\times 3+25\times 4\\ -24\times 4+25\times 3\end{pmatrix} = \frac{1}{35}\begin{pmatrix}28\\-21\end{pmatrix} = \frac{1}{5}\begin{pmatrix}4\\-3\end{pmatrix}.$$

Le procédé de Gram-Schmidt donne

$$e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{5}\begin{pmatrix}3\\4\end{pmatrix},$$

$$e_2 = \frac{v_2 - \langle v_2,e_1\rangle\, e_1}{\|v_2 - \langle v_2,e_1\rangle\, e_1\|} = \frac{1}{\|\ldots\|}\begin{pmatrix}4-24/5\times 3/5\\ 3-24/5\times 4/5\end{pmatrix} = \frac{1}{35}\begin{pmatrix}28\\-21\end{pmatrix}.$$

Ce procédé nécessite moins de calculs. Sans surprise, les deux méthodes conduisent au même résultat. Dans le cas défini positif, l'algorithme général produit un système triangulaire, qui doit coïncider aux signes près avec le résultat obtenu par le procédé de Gram-Schmidt.

3. Isométries

3.1 Définition d'une isométrie

Définition 30 *Soit $(E,\langle .,.\rangle)$un espace euclidien. Une application linéaire $f : E \to E$ est une* isométrie *de E si elle préserve le produit scalaire :*

$$\textit{Pour tout } x,y \in E, \quad \langle f(x),f(y)\rangle = \langle x,y\rangle.$$

Grâce aux identités de polarisation, on vérifie qu'une application linéaire est une isométrie si et seulement si elle préserve la norme :

$$\|f(x)\| = \|x\| \text{ pour tout } x \in E.$$

Exemples

– L'identité est une isométrie.

– Les symétries orthogonales sont des isométries. De fait une symétrie orthogonale s est de la forme $s(x) = p(x) - (x - p(x))$ avec x et $x - p(x)$ deux vecteurs orthogonaux de somme égale à x, si bien que

$$\|s(x)\|^2 = \|p(x)\|^2 + \|-(x - p(x))\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 = \|x\|^2.$$

Remarquons qu'une isométrie a un noyau réduit au vecteur nul. En effet, si $f(x) = 0$ alors $\|x\| = \|f(x)\| = 0$. En dimension finie, cela implique qu'une isométrie est inversible.

De plus, si $(e_1,\ldots,e_n)$ est une base orthonormée, il en va de même de son image $(f(e_1),\ldots,f(e_n))$ car les produits scalaires des vecteurs e_i, e_j sont préservés par f. Cette propriété caractérise les isométries.

Proposition 24 *Soit $(E,\langle .,. \rangle)$un espace euclidien de dimension finie muni d'une base orthonormée $(e_1, \dots, e_n)$ et $f : E \to E$ une application linéaire. Alors f est une isométrie si et seulement si l'image $(f(e_1), \dots, f(e_n))$ de cette base orthonormée est une base orthonormée.*

Démonstration
Soit $x = \sum x_i e_i$ et $y = \sum y_j e_j \in E$. Si $(f(e_1), \dots, f(e_n))$ est orthogonale,

$$\langle f(x),f(y)\rangle = \sum_{i,j} x_i y_j \langle f(e_i),f(e_j)\rangle = \sum_i x_i y_i = \langle x,y\rangle.$$

3.2 Groupe orthogonal

La matrice d'une isométrie exprimée dans une base orthonormée a une forme particulière. Voici comment la décrire.

Définition 31 *Une matrice P de taille $n \times n$ est dite* orthogonale *si elle vérifie l'égalité*

$$^tPP = id.$$

L'ensemble des matrices orthogonales est appelé groupe orthogonal *et noté*

$$O_n(\mathbf{R}) = \{P \in M_n(\mathbf{R}) \mid {}^tPP = id\}.$$

Le groupe spécial orthogonal *est composé des matrices orthogonales de déterminant* 1.

$$SO_n(\mathbf{R}) = \{P \in M_n(\mathbf{R}) \mid {}^tPP = id,\ \det(P) = 1\}.$$

Remarquons qu'une matrice orthogonale a nécessairement un déterminant égal à ± 1,

$$\det(P)^2 = \det({}^tPP) = \det(id) = 1,$$

et que son inverse est égal à sa transposée. Les matrices orthogonales de déterminant 1 sont celles qui préservent l'orientation canonique de $\mathbf{R}^n$.

Méthode : *caractérisation des matrices orthogonales*

Pour montrer qu'une matrice est orthogonale, il suffit de vérifier que ses colonnes sont de norme un et orthogonales entre elles. En effet, le coefficient i,j de la matrice ${}^tPP$ est obtenu en faisant le produit scalaire de la colonne i de P par sa colonne j. Cette matrice est égale à l'identité si et seulement si ses colonnes forment une famille orthonormée de $\mathbf{R}^n$.

Exemple
La matrice

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

est orthogonale de déterminant un. C'est un élément de $SO_2(\mathbf{R})$. La transformation de $\mathbf{R}^2$ associée est appelée *rotation d'angle θ*.

Les matrices orthogonales forment un groupe : le produit de deux matrices orthogonales est une matrice orthogonale et toute matrice orthogonale admet une matrice inverse qui est orthogonale.

$$\forall A, B \in O_n(\mathbf{R}),\ AB \in O_n(\mathbf{R}),$$

$$\forall P \in O_n(\mathbf{R}),\ P^{-1} = {}^tP \in O_n(\mathbf{R}).$$

On le vérifie par un calcul direct :

$${}^t(AB)AB = {}^tB^tAAB = {}^tBB = id,$$

$${}^t(P^{-1}) = ({}^tP)^{-1} = P^{-1}.$$

Faisons à présent le lien entre matrice orthogonale et isométrie.

Proposition 25 *Considérons un espace euclidien $(E,\langle .,. \rangle)$de dimension finie, une application linéaire $f : E \to E$ et une base orthonormée $(e_1, \ldots, e_n)$ de E. Alors f est une isométrie si et seulement si sa matrice dans la base (e_i) est orthogonale.*

Démonstration
Les colonnes de la matrice de f sont données par les coordonnées des vecteurs de la base $(f(e_i))$ dans la base (e_i). Ces vecteurs colonnes sont orthogonaux relativement au produit scalaire usuel sur $\mathbf{R}^n$ et de norme un si et seulement si la base $(f(e_i))$ est orthonormée. On a vu que cela caractérise les isométries.

3.3 Isométries du plan euclidien

Pour illustrer les notions qu'on vient de définir, nous allons classer les isométries en dimension deux. On commence par travailler en coordonnées avec des matrices orthogonales.

Proposition 26 *Toute matrice de $SO_2(\mathbf{R})$ est la matrice d'une rotation. Toute matrice de $O_2(\mathbf{R})$ de déterminant -1 est la matrice d'une réflexion orthogonale.*

Rappelons qu'une réflexion orthogonale est une symétrie orthogonale par rapport à un hyperplan. En dimension 2, les hyperplans sont des droites.

Démonstration
Une matrice de $O_2(\mathbf{R})$ est de la forme $P = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ avec ses deux colonnes de norme un et orthogonales entre elles, ce qui implique

$$\det \begin{pmatrix} a & -d \\ b & c \end{pmatrix} = ac + bd = 0.$$

Les vecteurs $\begin{pmatrix} a \\ b \end{pmatrix}$ et $\begin{pmatrix} -d \\ c \end{pmatrix}$ sont donc proportionnels et de norme un, le coefficient de proportionnalité vaut ± 1 et la matrice peut prendre une des deux formes suivantes.

- $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$: on obtient une rotation en posant $a = \cos(\theta)$, $b = \sin(\theta)$.
- $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$: c'est une réflexion orthogonale. Calculons son axe de symétrie.

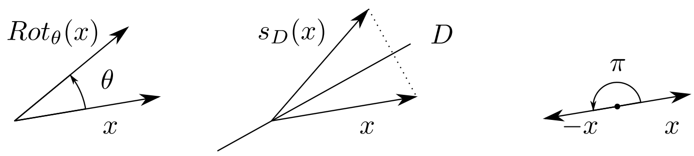
La droite de symétrie doit être fixe, on l'obtient en résolvant l'équation $Pv = v$, ce qui donne $v = \begin{pmatrix} b \\ 1-a \end{pmatrix}$.

L'espace orthogonal à $\mathrm{vect}(v)$ est une droite engendrée par le vecteur $w = \begin{pmatrix} a-1 \\ b \end{pmatrix}$. L'image de ce vecteur par notre matrice vaut

$$Pw = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \begin{pmatrix} a-1 \\ b \end{pmatrix} = \begin{pmatrix} a(a-1)+b^2 \\ (a-1)b - ab \end{pmatrix} = \begin{pmatrix} 1-a \\ -b \end{pmatrix} = -w$$

en prenant en compte la relation $a^2 + b^2 = 1$. Il s'agit donc bien d'une symétrie orthogonale relativement à $\mathrm{vect}(v)$. La proposition est démontrée.

Attention, en dimension deux, une symétrie centrale est une rotation d'angle π, ce n'est pas une réflexion.



Il faut aussi faire attention à ne pas se laisser tromper par la terminologie que nous venons d'introduire. La matrice d'une symétrie orthogonale est orthogonale. Par contre, la matrice d'une projection orthogonale n'est pas orthogonale en général : une matrice orthogonale est inversible alors qu'une projection différente de l'identité n'est jamais inversible.

Méthode : *caractérisation des isométries de* $\mathbf{R}^2$

Un élément $P \in O_2(\mathbf{R})$ est une rotation si son déterminant est égal à 1 et une réflexion sinon. On détermine l'angle d'une rotation en mettant la première colonne sous la forme $(\cos(\theta), \sin(\theta))$. On trouve la droite fixée par une réflexion en résolvant le système linéaire $Px = x$.

Exemples

– La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est la matrice de la réflexion orthogonale relativement à la droite engendrée par le vecteur $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

– La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est la matrice de la rotation d'angle $\frac{\pi}{2}$.

Corollaire 9 *Les isométries d'un espace vectoriel euclidien de dimension deux sont des rotations ou des réflexions orthogonales.*

4. Compléments

4.1 Isométries affines

Nous avons défini une isométrie comme une application linéaire d'un espace euclidien $(E,\langle .,. \rangle)$qui préserve la norme euclidienne. Il existe une notion plus générale d'isométrie reliée à la structure d'espace métrique d'un espace euclidien.

Définissons la *distance euclidienne* entre deux points $x,y \in E$ par la formule $d(x,y) = \|x - y\|$. Une application $f : E \to E$ est une *isométrie* si elle préserve la distance :

$$\forall\, x, y \in E, \quad d(f(x), f(y)) = d(x, y).$$

Ici, on ne suppose pas f linéaire. La *translation* de vecteur v est l'application définie de E dans E par $t_v(x) = x + v$. C'est un exemple d'isométrie qui n'est pas une application linéaire. On peut composer les translations avec les isométries linéaires pour obtenir de nouvelles isométries.

Définition 32 *Une* isométrie affine *est la composée d'une translation et d'isométrie linéaire.*

Il n'existe pas d'autres isométries dans un espace euclidien.

Proposition 27 *Dans un espace euclidien, toute isométrie est affine.*

Démonstration
Soit $f : E \to E$ une isométrie qui fixe l'origine : $f(0) = 0$. Par définition d'une isométrie, cela implique $\|f(x)\| = \|x\|$. Montrons que f préserve le produit scalaire.

$$\begin{array}{rcl}
\langle f(x),f(y)\rangle & = & -\langle f(x), -f(y)\rangle \\
 & = & -\frac{1}{2}\Big(\|f(x) - f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2\Big) \\
 & = & -\frac{1}{2}\Big(\|x - y\|^2 - \|x\|^2 - \|y\|^2\Big) \\
 & = & \langle x,y\rangle.
\end{array}$$

Montrons que f est linéaire.

$$\begin{array}{rl}
\multicolumn{2}{l}{\|f(x+y) - f(x) - f(y)\|^2} \\
= & \|f(x+y)\|^2 + \|f(x)\|^2 + \|f(y)\|^2 \\
 & -2\langle f(x+y),f(x)\rangle - 2\langle f(x+y),f(y)\rangle - 2\langle f(x),f(y)\rangle \\
= & \|x+y\|^2 + \|x\|^2 + \|y\|^2 - 2\langle x+y,x\rangle - 2\langle x+y,y\rangle - 2\langle x,y\rangle \\
= & \|x + y - x - y\|^2 \\
= & 0
\end{array}$$

Un calcul analogue donne l'égalité $f(\lambda x) = \lambda f(x)$.

Considérons maintenant une isométrie f qui ne fixe pas l'origine et posons $\tilde{f}(x) = f(x) - f(0)$. Cette nouvelle application préserve l'origine, elle est donc linéaire et $f = t_{f(0)} \circ \tilde{f}$. La proposition est démontrée.

4.2 Déterminant de Gram

Soit $(E,\langle .,. \rangle)$un espace euclidien et $v_1, \ldots, v_n$ des vecteurs de E. On peut restreindre le produit scalaire au sous-espace engendré par les v_i. Si ces vecteurs forment une famille libre, la matrice du produit scalaire prend la forme $B = \{\langle v_i,v_j \rangle\}_{i,j}$. Le déterminant de cette matrice est appelé déterminant de Gram.

$$Gram(v_1, \ldots, v_n) = \det \begin{pmatrix} \langle v_1,v_1 \rangle & \ldots & \langle v_1,v_n \rangle \\ \vdots & & \vdots \\ \langle v_n,v_1 \rangle & \ldots & \langle v_n,v_n \rangle \end{pmatrix}.$$

Si les vecteurs $v_1, \ldots, v_n$ sont linéairement dépendants, une relation de dépendance sur ces vecteurs induit une relation de dépendance sur les colonnes de la matrice B et le déterminant précédent est nul. On en déduit un critère d'indépendance en terme de produit scalaire.

Proposition 28 *Soit $(E,\langle .,. \rangle)$un espace euclidien et $v_1, \ldots, v_n$ des vecteurs de E. Alors ces vecteurs forment une famille libre si et seulement si*

$$Gram(v_1, \ldots, v_n) \neq 0.$$

Le déterminant de Gram est relié à la notion de volume. Par analogie avec le cas de la dimension deux, on définit le volume n-dimensionnel du parallélépipède de côtés $v_1, \ldots, v_n$ à partir du déterminant de la matrice dont les colonnes sont données par les coordonnées des v_j dans une base orthonormée (e_i) de l'espace engendré par les v_j.

Définition 33 *Soit $(E,\langle .,. \rangle)$un espace euclidien, $(v_1, \ldots, v_n)$ une famille libre de E et $(e_1, \ldots, e_n)$ une base orthonormée de l'espace engendré par les v_j. Le* volume n-dimensionnel *du parallélépipède de côtés $v_1, \ldots, v_n$ est défini par*

$$\mathrm{vol}_n(v_1, \ldots, v_n) = |\det(v_1, \ldots, v_n)| = |\det(\{\langle e_i,v_j \rangle\}_{i,j})|.$$

Exprimons ce volume comme un déterminant de Gram. La formule de changement de base pour les formes quadratiques définies positives s'exprime sous la forme $B = {}^tPP$, où P est la matrice de passage de la base (e_i) à la base (v_j). Elle a pour colonnes les coordonnées des v_j dans la base (e_i) : $P = \{\langle e_i,v_j \rangle\}_{i,j}$. En prenant le déterminant, on obtient $\det(B) = \det(P)^2$, ce qui donne la formule suivante.

Proposition 29

$$vol_n(v_1, \ldots, v_n) = |\det(\{\langle e_i,v_j \rangle\}_{i,j})| = \sqrt{Gram(v_1, \ldots, v_n)}.$$

Le dernier terme ne dépend pas de la base (e_i), on voit que le volume ne dépend pas de la base orthonormée choisie pour le calculer.

4.3 Distance à un sous-espace

Donnons une autre application du déterminant de Gram au calcul de la distance d'un vecteur à un sous-espace F d'un espace euclidien $(E,\langle .,. \rangle)$. On définit la distance entre deux points x, $y \in E$ ainsi que la distance entre un point $x \in E$ et un sous-espace vectoriel $F \subset E$ de dimension finie par

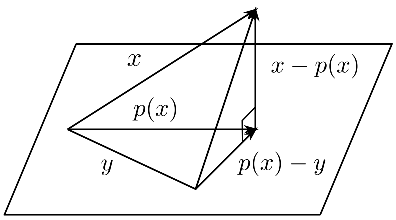
$$d(x,y) = \|x - y\|,$$
$$d(x,F) = \min\{d(x,y) \mid y \in F\}.$$

Proposition 30 *Soit $(E,\langle .,. \rangle)$un espace euclidien et F un sous-espace vectoriel de E de dimension finie. Notons par p la projection orthogonale sur F. Alors*

$$d(x, F) = \|x - p(x)\|.$$

Démonstration
Soit $y \in F$, les vecteurs $x - p(x) \in F^\perp$ et $p(x) - y \in F$ sont orthogonaux d'où

$$\begin{aligned}\|x - y\|^2 &= \|x - p(x) + p(x) - y\|^2 \\ &= \|x - p(x)\|^2 + \|p(x) - y\|^2.\end{aligned}$$



La quantité $\|x - p(x)\|$ minore donc $\|x - y\|$ pour tout $y \in F$ avec égalité si et seulement si $y = p(x)$.

Soit maintenant $(v_1, \dots, v_n)$ une base de F et $v_{n+1} \in E$. Nous pouvons exprimer la distance de v_{n+1} à F à l'aide des déterminants de Gram.

Proposition 31 *Soit $(E,\langle .,. \rangle)$un espace euclidien, $(v_1, \dots, v_n)$ une famille libre de E et $v_{n+1} \in E$. Alors le carré de la distance de v_{n+1} à l'espace F engendré par les v_i est donné par*

$$d(v_{n+1},F)^2 = \frac{Gram(v_1, \dots, v_{n+1})}{Gram(v_1, \dots, v_n)}.$$

Démonstration
Si v_{n+1} est dans F, il n'y a rien à démontrer. Sinon, on se place dans l'espace euclidien engendré par $(v_1, \dots, v_{n+1})$. Soit $(e_1, \dots, e_n,e_{n+1})$ une base orthogonale de cet espace choisie de telle sorte que $(e_1, \dots, e_n)$ soit une base de F. Le vecteur $v_{n+1} - p(v_{n+1})$ est la projection de v_{n+1} sur e_{n+1} et vaut $\langle e_{n+1},v_{n+1} \rangle e_{n+1}$.

$$Gram(v_1, \dots, v_{n+1}) = \det \begin{pmatrix} \langle v_1,e_1 \rangle & \dots & \langle v_n,e_1 \rangle & \langle v_{n+1},e_1 \rangle \\ \vdots & & & \vdots \\ \langle v_1,e_n \rangle & \dots & \langle v_n,e_n \rangle & \langle v_{n+1},e_n \rangle \\ \langle v_1,e_{n+1} \rangle & \dots & \langle v_n,e_{n+1} \rangle & \langle v_{n+1},e_{n+1} \rangle \end{pmatrix}^2$$

Tous les coefficients de la dernière ligne sont nuls hormis le dernier. On conclut en développant le déterminant relativement à cette ligne.

Chapitre 4

Réduction des applications linéaires

Ce chapitre est consacré à l'étude des applications linéaires. On commence par quelques rappels sur les déterminants avant d'aborder le procédé général de diagonalisation. Ce procédé consiste à trouver une base de l'espace constituée de vecteurs propres de l'application, afin d'en simplifier l'étude. Il n'est pas toujours possible de diagonaliser une application linéaire. Dans le cas euclidien, on introduira des classes d'applications pour lesquelles on peut effectuer la réduction en se basant sur les outils géométriques développés au chapitre précédent. On termine par quelques applications de la réduction à la géométrie.

1. Déterminant

1.1 Définition du déterminant

Le déterminant d'une matrice carrée peut se définir par récurrence sur la dimension de la matrice, par développement relativement à la première colonne.

– Le déterminant d'une matrice 1×1 est égal à son unique coefficient.

$$\det(a_{1,1}) = a_{1,1}.$$

– Le déterminant d'une matrice $n \times n$ est alors obtenu grâce à la formule

$$\det\Big(\{a_{i,j}\}_{\substack{i=1..m\\ j=1..n}}\Big) = \sum_{k=1}^{n} (-1)^{k+1}\, a_{k,1}\, \det(\{a_{l,m}\}_{l\neq k, m\neq 1}).$$

Le déterminant d'une matrice est parfois noté en remplaçant les parenthèses qui entourent les coefficients de la matrice par des barres verticales. Avec cette notation, la formule précédente s'écrit

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ a_{3,1} & a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2,2} & \dots & a_{2,n} \\ a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{n,n} \end{vmatrix} - a_{2,1} \begin{vmatrix} a_{1,2} & \dots & a_{1,n} \\ a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{n,n} \end{vmatrix} +$$

$$\ldots + (-1)^{n+1} a_{n,1} \begin{vmatrix} a_{1,1} & \ldots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \ldots & a_{n-1,n-1} \end{vmatrix}.$$

Exemple

$$\begin{vmatrix} \mathbf{-1} & 2 & 2 \\ \mathbf{3} & -1 & 3 \\ \mathbf{5} & 5 & -1 \end{vmatrix} = \mathbf{-1} \times \begin{vmatrix} -1 & 3 \\ 5 & -1 \end{vmatrix} - \mathbf{3} \times \begin{vmatrix} 2 & 2 \\ 5 & -1 \end{vmatrix} + \mathbf{5} \times \begin{vmatrix} 2 & 2 \\ -1 & 3 \end{vmatrix}$$
$$= 14 + 36 + 40$$
$$= 90.$$

Il est aussi possible de définir le déterminant en développant en lignes plutôt qu'en colonnes.

1.2 Calcul du déterminant

Il est possible de calculer le déterminant de certaines matrices directement à partir de la définition précédente.

Déterminant d'une matrice 2×2

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Exemple

$$\det \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} = 5 \times 5 - 2 \times 2 = 21.$$

Déterminant d'une matrice 3×3

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + dhc + gbf - ceg - fha - ibd.$$

Exemple

$$\det \begin{pmatrix} -1 & 2 & 2 \\ 3 & -1 & 3 \\ 5 & 5 & -1 \end{pmatrix} = -1 + 30 + 30 + 10 + 15 + 6 = 90.$$

Déterminant d'une matrice triangulaire

Une matrice est *triangulaire supérieure* si $a_{i,j} = 0$ pour tout i,j tels que $i > j$. Le déterminant d'une telle matrice est égal au produit de ses termes diagonaux.

$$\det \begin{pmatrix} a_{1,1} & a_{2,1} & \ldots & a_{n-1,1} & a_{n,1} \\ 0 & a_{2,2} & \ldots & a_{n-1,2} & a_{n,2} \\ 0 & 0 & \ldots & a_{n-1,3} & a_{n,3} \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \vdots & 0 & a_{n,n} \end{pmatrix} = a_{1,1}\, a_{2,2} \,\ldots\, a_{n,n}.$$

Cette proposition se démontre par récurrence sur la taille de la matrice. On en déduit que le déterminant de la matrice identité de taille $n \times n$ est égal à 1.

Le déterminant peut aussi se calculer à l'aide de l'algorithme de Gauss.

Méthode : *calcul du déterminant par l'algorithme de Gauss*

Ajouter à une ligne un multiple d'une autre ligne ne modifie pas la valeur du déterminant. Permuter deux lignes change son signe. En utilisant ces deux opérations, on peut mettre la matrice sous forme triangulaire en employant la méthode du pivot de Gauss. Le déterminant est alors égal au produit des termes diagonaux. Attention à ne pas oublier de changer le signe du déterminant chaque fois qu'on permute deux lignes.

Exemple

$$\begin{vmatrix} -1 & 2 & 2 \\ 3 & -1 & 3 \\ 5 & 5 & -1 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 2 \\ 0 & 5 & 9 \\ 0 & 15 & 9 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 2 \\ 0 & 5 & 9 \\ 0 & 0 & -18 \end{vmatrix} = (-1) \times 5 \times (-18) = 90.$$

1.3 Propriétés des déterminants

Soit A une matrice de taille $n \times n$. Notons $A_1, \ldots, A_n$ ses colonnes et considérons le déterminant de A comme fonction des colonnes de A :

$$(A_1, \ldots, A_n) \mapsto \det(A_1, \ldots, A_n) = \det(A).$$

Les deux propriétés clef du déterminant sont les suivantes.

– Le déterminant est une fonction alternée des colonnes de la matrice : si deux colonnes sont égales, alors le déterminant est nul.

– Le déterminant dépend de manière linéaire de chacune des colonnes de la matrice : pour tout j, l'application $A_j \mapsto \det(A_1, \ldots, A_j, \ldots, A_n)$ est linéaire.

On démontre ces propriétés à partir de la définition du déterminant par une récurrence sur la taille de la matrice. On vérifie ensuite que toute fonction de n colonnes qui vérifie les deux propriétés précédentes est nécessairement proportionnelle au déterminant. À partir de cette caractérisation, on démontre les formules suivantes.

Proposition 32 *Le déterminant d'un produit de matrices carrées est égal au produit des déterminants.*

$$\det(AB) = \det(A)\det(B).$$

Le déterminant d'une matrice inversible est égal à l'inverse de son déterminant.

$$\det(A^{-1}) = \det(A)^{-1}.$$

La multilinéarité implique également la formule

$$\det(\lambda A) = \lambda^n \det(A).$$

Rappelons que la *transposée* de la matrice carrée $A = \{a_{i,j}\}$ est la matrice de coefficients $\{a_{j,i}\}$. Elle est notée ${}^tA$. Les colonnes de A deviennent les lignes de ${}^tA$ et vice-versa.

Proposition 33 *Le déterminant d'une matrice est égal à celui de sa transposée.*

$$\det({}^tA) = \det(A).$$

Le déterminant permet aussi de déterminer si une matrice est inversible. C'est une propriété importante que nous utiliserons dans la suite pour calculer les valeurs propres d'une matrice.

Proposition 34 *Une matrice carrée est inversible si et seulement si son déterminant est non nul.*

Enfin, on peut faire appel au déterminant pour résoudre des systèmes d'équations linéaires. La solution d'un système linéaire s'exprime comme un quotient de deux déterminants grâce aux formules de Cramer.

Proposition 35 (formules de Cramer) *Soit A une matrice de taille $n \times n$ et x, y deux vecteurs colonnes de taille n. Notons $x_1, \ldots, x_n$ les coordonnées de x. Supposons que $Ax = y$ et det$(A) \neq 0$. Alors, pour tout j,*

$$x_j = \frac{\det(A_1, \ldots, A_{j-1}, y, A_{j+1}, \ldots, A_n)}{\det(A_1, \ldots, A_n)}.$$

Le déterminant au dénominateur est le déterminant de la matrice A. Le déterminant au numérateur est celui de la matrice obtenue à partir de A en remplaçant sa colonne j par y.

Démonstration des formules de Cramer
L'égalité $Ax = y$ se traduit en colonnes par la relation $\sum_k x_k A_k = y$. On a alors

$$\begin{aligned}
\det(A_1,\ldots,A_{j-1},y,A_{j+1},\ldots,A_n) &= \det(A_1,\ldots,A_{j-1},\sum_k x_k A_k, A_{j+1},\ldots,A_n) \\
&= \sum_k x_k \det(A_1,\ldots,A_{j-1},A_k,A_{j+1},\ldots,A_n) \\
&= x_j \det(A_1,\ldots,A_{j-1},A_j,A_{j+1},\ldots,A_n)
\end{aligned}$$

par linéarité et en utilisant le caractère alterné du déterminant.

Les formules de Cramer permettent d'exprimer les coordonnées d'un vecteur de $\mathbf{R}^n$ dans une base donnée sous forme de déterminants. Considérons une base $(v_1,\dots,v_n)$ de $\mathbf{R}^n$ et appliquons le résultat précédent à la matrice A dont les colonnes sont données par les v_j.

Corollaire 10 *Soit $(v_1,\dots,v_n)$ une base de $\mathbf{R}^n$ et y un vecteur de $\mathbf{R}^n$. Alors*

$$y = \sum_{j=1}^{n} \frac{\det(v_1,\dots,v_{j-1},y,v_{j+1},\dots,v_n)}{\det(v_1,\dots,v_n)}\, v_j.$$

Méthode : *résolution d'un système par les formules de Cramer*

Pour résoudre une équation linéaire de la forme $AX = Y$, on commence par calculer le déterminant de la matrice A qui doit être non nul pour qu'il y ait une unique solution. Pour calculer la coordonnée j de la solution X, on remplace la colonne j de A par Y, on calcule son déterminant et on divise par le déterminant de A.

Exemple
Résolvons le système suivant avec les formules de Cramer.

$$\begin{cases} 5x_1 + 2x_2 = 10 \\ 2x_1 + 5x_2 = 8 \end{cases}$$

$$x_1 = \frac{\det\begin{pmatrix} 10 & 2 \\ 8 & 5 \end{pmatrix}}{\det\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}} = \frac{10\times 5 - 8\times 2}{5\times 5 - 2\times 2} = \frac{34}{21},$$

$$x_2 = \frac{\det\begin{pmatrix} 5 & 10 \\ 2 & 8 \end{pmatrix}}{\det\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}} = \frac{5\times 8 - 2\times 10}{5\times 5 - 2\times 2} = \frac{20}{21}.$$

1.4 Déterminant d'une application linéaire

Définissons le déterminant d'une application linéaire en faisant appel à sa matrice dans une certaine base.

Définition 34 *Soit E un espace vectoriel de dimension finie et $f : E \to E$ une application linéaire. Le déterminant de f est égal au déterminant de sa matrice dans une base quelconque de E. Il ne dépend pas de la base choisie.*

Vérifions qu'il ne dépend pas de la base choisie. Soit (e_i), (e'_i) deux bases de E, A et A' les matrices de f dans ces bases et P la matrice de passage de la base (e_i) à la base (e'_i). La formule de changement de base donne

$$A' = P^{-1}AP$$

d'où $\det(A') = \det(P)^{-1} \det(A) \det(P) = \det(A)$. Deux matrices conjuguées ont même déterminant.

Exemples
– Le déterminant de l'application identité $id : x \mapsto x$ est égal à 1.

– Une homothétie $f(x) = \lambda x$ définie sur un espace de dimension n a un déterminant égal à λ^n.

– Une isométrie définie sur un espace de dimension finie a un déterminant égal à ± 1 car sa matrice est orthogonale.

On dira qu'une application linéaire $f : E \to E$ *préserve l'orientation*, ou encore qu'elle est directe, si son déterminant est strictement positif. Si l'espace E est orienté, l'image d'une base directe par une telle application est directe. Une application dont le déterminant est strictement négatif renverse l'orientation : l'image d'une base directe est indirecte et réciproquement. C'est une conséquence de la formule du produit pour le déterminant.

Dans ce cours, nous utiliserons le déterminant pour
– déterminer si une matrice est inversible,
– déterminer si une famille de vecteurs de $\mathbf{R}^n$ est une base,
– résoudre des systèmes linéaires,
– calculer des valeurs propres.

En analyse, le déterminant est aussi employé pour calculer des volumes, faire des changements de variables dans les intégrales multiples, etc.

2. Diagonalisation

2.1 Somme de sous-espaces vectoriels

Soit $E_1, \ldots, E_n$ des sous-espaces vectoriels d'un espace vectoriel E. Rappelons que la somme de ces sous-espaces est définie par

$$E_1 + \ldots + E_n = \{v_1 + \ldots + v_n \mid v_1 \in E_1, \ldots, v_n \in E_n\}.$$

Définition 35 *Soit E un espace vectoriel et $E_1, \ldots, E_n$ des sous-espaces vectoriels de E. On dit que ces sous-espaces sont en* somme directe *si tout élément $x \in E_1 + \ldots + E_n$ s'écrit de manière unique sous la forme*

$$x = x_1 + \ldots + x_n \quad \textit{avec } x_1 \in E_1, \ldots, x_n \in E_n.$$

La somme de ces sous-espaces est alors notée $E_1 \oplus \ldots \oplus E_n$.

On dit qu'un espace vectoriel E est somme directe des sous-espaces E_1, $E_2, \ldots, E_n$ si $E = E_1 \oplus E_2 \oplus \ldots \oplus E_n$.

Remarquons que des vecteurs $v_1, \ldots, v_n$ tous non nuls sont linéairement indépendants entre eux si et seulement si les droites associées $\mathrm{vect}(v_1), \ldots, \mathrm{vect}(v_n)$ sont en somme directe.

Méthode : *caractérisation des sommes directes*

Pour montrer que des sous-espaces $E_1, \ldots, E_n$ sont en somme directe, il suffit de vérifier que pour tout $x_1 \in E_1, \ldots, x_n \in E_n$, la condition $x_1 + \ldots + x_n = 0$ implique que tous les x_i sont nuls. Pour montrer que deux sous-espaces E_1 et E_2 sont en somme directe, il suffit de montrer que leur intersection ne contient que le vecteur nul : $E_1 \cap E_2 = \{0\}$.

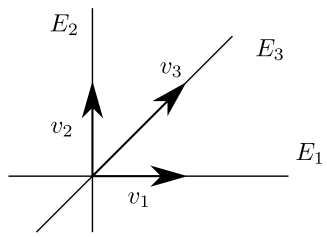
De fait, si un vecteur x admettait deux décompositions différentes, la différence donnerait une somme nulle, comprenant au moins un vecteur non nul. Dans le cas de deux sous-espaces, si $x_1 + x_2$ est nul sans que x_1 et x_2 ne soient nuls, le vecteur $x = x_1 = -x_2$ appartient à l'intersection sans être nul.

On peut vérifier que les sous-espaces $E_1 \ldots, E_n$ sont en somme directe si et seulement si pour tout $k > 1$, E_k est en somme directe avec $E_1 + \ldots + E_{k-1}$. Par contre, il ne suffit pas qu'ils soient en somme directe deux à deux.

Exemples

– Considérons les sous-espaces de $\mathbf{R}^2$ donnés par $E_i = \mathrm{vect}(v_i)$ avec

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$



On a $E_1 \cap E_2 = E_2 \cap E_3 = E_1 \cap E_3 = \{0\}$ si bien que deux quelconques de ces sous-espaces sont en somme directe. Par contre, les trois sous-espaces ne sont pas en somme directe car le vecteur v_2 admet deux décompositions possibles en somme d'éléments de E_1, E_2 et E_3 :

$$\begin{array}{ccccccc} v_2 & = & 0 & + & v_2 & + & 0 \\ & = & v_1 & + & 0 & + & v_3. \end{array}$$

– Soit $(E, \langle .,. \rangle)$un espace euclidien et E_1 un sous-espace de dimension finie. On a vu au chapitre précédent que $E = E_1 \oplus E_1^\perp$ (corollaire 7).

– Les deux sous-espaces vectoriels de $C^0([0,1],\mathbf{R})$ donnés par

$$E_1 = \left\{ f \in E \;\middle|\; \int_0^1 f(t)\,dt = 0 \right\}, \quad E_2 = \{\text{fonctions constantes}\}$$

sont en somme directe : $E = E_1 \oplus E_2$. Cela découle de l'exemple précédent en prenant $E_2 = \mathrm{vect}(\mathbf{1}_{[0,1]})$ et $E_1 = E_2^\perp$, en notant $\mathbf{1}_{[0,1]}$ la fonction constante égale à un sur $[0,1]$.

La dimension d'une somme directe se calcule aisément.

Proposition 36 *Soient E un espace vectoriel, $E_1, \ldots, E_n$ des sous-espaces vectoriels de dimension finie en somme directe et $\mathcal{B}_1, \ldots, \mathcal{B}_n$ des bases de chacun des E_i. L'ensemble des vecteurs appartenant à toutes ces bases forme une base de la somme*

$$E_1 \oplus \ldots \oplus E_n$$

et

$$dim\ E_1 \oplus \ldots \oplus E_n = dim\ E_1 + \ldots + dim\ E_n.$$

Démonstration
Pour chaque i, choisissons une base $(e_{i,j})_j$ de chacun des E_i et montrons que la famille $(e_{i,j})_{i,j}$ qui contient tous les éléments de ces bases est une base de la somme. C'est bien une famille génératrice car tout élément de la somme est somme de vecteurs qui s'expriment comme combinaisons linéaires des éléments de ces bases. Elle est libre : s'il existe des coefficients $c_{i,j}$ tels que $\sum_{i,j} c_{i,j} e_{i,j} = 0$, la somme sur i des vecteurs $\sum_j c_{i,j} e_{i,j} \in E_i$ est nulle. Chacun de ces vecteurs est donc nul car la somme est directe et on conclut que pour tout i, les $(c_{i,j})_j$ sont nuls par indépendance des $(e_{i,j})_j$.

Lorsque deux espaces ne sont pas en somme directe, on peut tout de même donner une formule pour la dimension de la somme.

Proposition 37 *Soit E un espace vectoriel et E_1, E_2 deux sous-espaces vectoriels de E de dimension finie. Alors*

$$dim\ (E_1 + E_2) = dim\ E_1 + dim\ E_2 - dim\ (E_1 \cap E_2).$$

Démonstration
Choisissons un sous-espace E' de E_1 qui est un supplémentaire de $E_1 \cap E_2$. Montrons que E' et E_2 sont en somme directe. Soit $x \in E' \cap E_2$, ce vecteur est dans E' donc dans E_1 et dans $E' \cap E_1 \cap E_2 = \{0\}$. On a bien $x = 0$.

$$E_1 = E' \oplus (E_1 \cap E_2), \qquad E_1 + E_2 = E' \oplus E_2,$$

d'où $dim\ E_1 + E_2 = dim\ E' + dim\ E_2 = dim\ E_1 - dim\ E_1 \cap E_2 + dim\ E_2$.

Terminons par un résultat propre aux espaces euclidiens qui donne un lien entre orthogonalité et somme directe.

Proposition 38 *Soit $(E,\langle .,. \rangle)$un espace euclidien et $E_1, \ldots, E_n$ des sous-espaces vectoriels de E orthogonaux entre eux :*

$$\textit{pour tout } i,j \textit{ distincts, pour tout } x_i \in E_i,\ x_j \in E_j, \quad \langle x_i, x_j \rangle = 0.$$

Alors ces sous-espaces sont en somme directe.

Démonstration
Soit $x_i \in E_i$ tels que $x_1 + \ldots + x_n = 0$. Faisons le produit scalaire avec x_j pour tout j :

$$0 = \langle x_1 + \ldots + x_n, x_j \rangle = \langle x_1, x_j \rangle + \ldots + \langle x_n, x_j \rangle = \|x_j\|^2.$$

On conclut que $x_j = 0$ pour tout j.

2.2 Valeurs propres et vecteurs propres

Définition 36 *Soit E un espace vectoriel et $f : E \to E$ une application linéaire. Un nombre réel λ est appelé* valeur propre *réelle de f s'il existe un vecteur $v \in E$ non nul tel que*

$$f(v) = \lambda v.$$

Les vecteurs non nuls satisfaisant cette égalité sont appelés vecteurs propres *de f associés à la valeur propre λ.*

On définit de la même façon la notion de valeur propre complexe, associée à un vecteur propre $v \in \mathbf{C}^n$ non nul qui satisfait $f(v) = \lambda v$. Remarquons que les vecteurs propres associés à une valeur propre λ sont exactement les éléments du noyau de $f - \lambda id$, où $id : E \to E$ est l'application identité : $id(x) = x$.

Définition 37 *Le sous-espace vectoriel $ker(f - \lambda id)$ est appelé sous-espace propre associé à la valeur propre λ. Il est composé des vecteurs propres associés à λ et du vecteur nul.*

Exemples
– L'application identité a une unique valeur propre égale à 1 et tous les vecteurs sont propres.

– Une symétrie centrale a une unique valeur propre égale à -1. Là encore, tous les vecteurs de l'espace sur lequel est définie la symétrie sont des vecteurs propres.

Nous allons faire appel à la notion de polynôme caractéristique pour calculer ces valeurs propres.

Proposition 39 *Soit $f : E \to E$ une application linéaire définie sur un espace vectoriel E de dimension finie n. Le* polynôme caractéristique *P_c de f est le polynôme défini sur $\mathbf{R}$ ou $\mathbf{C}$ par*

$$P_c(X) = \det(X\,id - f).$$

Un nombre réel (ou complexe) λ est valeur propre de f si et seulement si il est racine réelle (ou complexe) du polynôme caractéristique.

On démontre par récurrence, en développant par rapport à la première colonne, que le polynôme caractéristique a un degré égal à n. Le polynôme caractéristique est parfois défini par la formule

$$P_c(X) = \det(f - X\,id) = (-1)^n \det(X\,id - f).$$

C'est le même polynôme à un signe près.

Démonstration
Il existe un vecteur v non nul tel que $f(v) = \lambda v$ si et seulement si on a l'égalité $(f - \lambda id)v = 0$, c'est-à-dire si et seulement si $f - \lambda id$ est une application qui n'est pas inversible. C'est équivalent à l'annulation de son déterminant. La proposition est démontrée.

Méthode : *Calcul des valeurs propres d'une application linéaire*

On se place dans une base de E, notons A la matrice de f dans cette base. Il faut calculer le polynôme caractéristique $P_c(X) = \det(Xid - A)$, par exemple en développant ce déterminant en colonne, puis trouver ses racines.

Exemples
– L'identité, définie de $\mathbf{R}^n$ dans $\mathbf{R}^n$, a pour polynôme caractéristique

$$P_c(X) = \det(Xid - id) = \det((X-1)id) = (X-1)^n.$$

– La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ a pour polynôme caractéristique

$$P_c(X) = \det \begin{pmatrix} X & -1 \\ -1 & X \end{pmatrix} = X^2 - 1.$$

Elle possède deux valeurs propres réelles -1 et 1.
– La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ a pour polynôme caractéristique

$$P_c(X) = \det \begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix} = X^2 + 1.$$

La matrice n'a pas de valeurs propres réelles. Elle possède cependant deux valeurs propres complexes i et $-i$.
– La matrice $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ a pour polynôme caractéristique

$$P_c(X) = \det \begin{pmatrix} X-1 & -1 \\ -1 & X \end{pmatrix} = X^2 - X - 1.$$

Le discriminant de ce polynôme vaut 5. La matrice possède deux valeurs propres réelles données par

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

De manière générale, le polynôme caractéristique de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est donné par

$$\det\left(X\,id - \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \det\begin{pmatrix} X-a & -b \\ -c & X-d \end{pmatrix} = X^2 - (a+d)X + ad - bc.$$

Le discriminant de ce polynôme vaut $\Delta = (a+d)^2 - 4(ad-bc) = (a-d)^2 + 4bc$. S'il est strictement positif, la matrice a deux valeurs propres réelles. S'il est strictement négatif, la matrice a deux valeurs propres complexes qui ne sont pas réelles.

Terminons par un corollaire de la proposition précédente, qui fait le lien entre déterminant et valeurs propres complexes d'une application linéaire.

Corollaire 11 *Soit f une application linéaire définie sur un espace vectoriel réel ou complexe de dimension finie n et $\lambda_1, \ldots, \lambda_n$ les racines* complexes *de son polynôme caractéristique. Alors*

$$\det(f) = \prod_{i=1}^{n} \lambda_i.$$

On dit que le déterminant est égal au produit des valeurs propres, comptées avec multiplicités, la *multiplicité* d'une valeur propre faisant référence au nombre de fois où elle apparaît dans l'expression factorisée du polynôme caractéristique. Il faut bien noter que dans le cas réel, il ne suffit pas de faire le produit des valeurs propres réelles pour obtenir le déterminant en général.

Pour démontrer ce corollaire, il suffit de factoriser le polynôme P_c sur $\mathbf{C}$,

$$P_c(X) = \det(Xid - f) = \prod_{i=1}^{n} (X - \lambda_i)$$

et de l'évaluer en $X = 0$. L'existence d'une telle factorisation est présentée en complément (proposition 45, page 108) et découle du théorème de d'Alembert-Gauss.

Exemples
– L'homothétie de $\mathbf{R}^n$ donnée par $x \mapsto \lambda x$ a pour polynôme caractéristique

$$P_c(X) = \det(Xid - \lambda id) = (X - \lambda)^n = \prod_{i=1}^{n} (X - \lambda).$$

Cette application a pour valeur propre λ avec multiplicité n et son déterminant vaut λ^n.
– La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est de déterminant 1, on a vu précédemment qu'elle n'a pas de valeurs propres réelles. Elle possède cependant deux valeurs propres complexes i et $-i$ et on a bien l'égalité $1 = i \times (-i)$.

$$P_c(X) = X^2 + 1 = (X - i)(X + i).$$

2.3 Diagonalisation

On cherche une base dans laquelle l'application a une expression particulièrement simple.

Définition 38 *Une application linéaire $f : E \to E$ définie sur un espace vectoriel réel E est dite diagonalisable sur $\mathbf{R}$ s'il existe une base de E composée de vecteurs propres de f.*

On a une définition analogue pour les applications définies sur un espace vectoriel complexe.

On parle d'application diagonalisable car la matrice d'une telle application dans la base de vecteurs propres est diagonale. Pour déterminer si une matrice est diagonalisable, il faut calculer ses valeurs propres et les vecteurs propres associés. Si l'ensemble de ces vecteurs propres forment une base, la matrice est diagonalisable.

Méthode : *Calcul des vecteurs propres d'une application linéaire*

Il s'agit de déterminer une base de $\ker(f - \lambda id)$. On a vu au premier chapitre comment procéder. Soit A la matrice de f dans une base de E. On résout le système $(A - \lambda id)v = 0$ par la méthode du pivot de Gauss. Les coordonnées obtenues sont celles des vecteurs propres dans la base considérée.

Exemples

– La réflexion orthogonale de $\mathbf{R}^2$ relativement à la droite dirigée par le vecteur $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ a pour matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et pour valeurs propres -1 et 1.

Calculons les vecteurs propres $x = (x_1, x_2)$ associés à la valeur propre 1. L'équation $Ax = x$ donne le système

$$\begin{cases} -x_1 + x_2 & = & 0 \\ x_1 - x_2 & = & 0 \end{cases}$$

soit $\ker(A - id) = \text{vect}(\begin{pmatrix} 1 \\ 1 \end{pmatrix})$.

Calculons les vecteurs propres $x = (x_1, x_2)$ associés à la valeur propre -1. L'équation $Ax = -x$ donne le système

$$\begin{cases} x_1 + x_2 & = & 0 \\ x_1 + x_2 & = & 0 \end{cases}$$

soit $\ker(A + id) = \text{vect}(\begin{pmatrix} 1 \\ -1 \end{pmatrix})$.

Une base diagonalisant A est donc donnée par $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$.

– La rotation d'angle $\frac{\pi}{2}$ a pour matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, elle n'a pas de valeurs propres réelles ni de vecteurs propres réels. Considérons A comme agissant sur $\mathbf{C}^2$ plutôt que $\mathbf{R}^2$. Ses valeurs propres complexes sont i et $-i$, calculons les vecteurs propres complexes associés.

Calculons les vecteurs propres $z = (z_1,z_2)$ associés à la valeur propre i. L'équation $Az = iz$ donne le système

$$\begin{cases} -iz_1 - z_2 & = & 0 \\ z_1 - iz_2 & = & 0 \end{cases}$$

soit $\ker(A - i\,id) = \mathrm{vect}(\begin{pmatrix} i \\ 1 \end{pmatrix})$.

Calculons les vecteurs propres $z = (z_1,z_2)$ associés à la valeur propre $-i$. L'équation $Az = -z$ donne le système

$$\begin{cases} iz_1 - z_2 & = & 0 \\ z_1 + iz_2 & = & 0 \end{cases}$$

soit $\ker(A + i\,id) = \mathrm{vect}(\begin{pmatrix} -i \\ 1 \end{pmatrix})$.

Une base diagonalisant A sur $\mathbf{C}$ est donnée par $\left(\begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$.

– La matrice $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ a pour valeurs propres $\lambda_1 = \frac{1+\sqrt{5}}{2}$, $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Calculons les vecteurs propres $x = (x_1,x_2)$ associés à la valeur propre λ_1. L'équation $Az = \lambda_1 z$ donne le système

$$\begin{cases} (1 - \lambda_1)\,x_1 + x_2 & = & 0 \\ x_1 - \lambda_1 x_2 & = & 0 \end{cases}$$

soit $\ker(A - \lambda_1\,id) = \mathrm{vect}(\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix})$. De même, $\ker(A - \lambda_2\,id) = \mathrm{vect}(\begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix})$.

Une base diagonalisant A sur $\mathbf{C}$ est donnée par $\left(\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}, \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}\right)$.

2.4 Interprétation matricielle

Les résultats précédents peuvent s'exprimer en termes de matrices. Soit A une matrice de taille $n \times n$. On peut parler de valeurs, vecteurs et sous-espaces propres de A. Ce sont ceux de l'application linéaire définie de $\mathbf{R}^n$ dans $\mathbf{R}^n$ associée à A. De même, une matrice est diagonalisable si cette application est diagonalisable. Cette propriété peut se reformuler comme suit.

Proposition 40 *Une matrice A de taille $n\times n$ est diagonalisable si et seulement si il existe une matrice diagonale D et une matrice inversible P de tailles $n \times n$ telles que*

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

Les coefficients diagonaux de D sont les valeurs propres de A et les colonnes de la matrice P sont des vecteurs propres associés à chacune de ces valeurs.

Démonstration

Supposons A diagonalisable. Soit (e'_i) une base de vecteurs propres de A, P

la matrice de passage de la base canonique à la base (e'_i) et D la matrice de A dans cette base (e'_i). Cette matrice D est diagonale et par la formule de changement de base, $A = PDP^{-1}$.

Réciproquement, si A est de la forme PDP^{-1}, notons $(e_i)_i$ la base canonique de $\mathbf{R}^n$, e'_j la j^{ieme} colonne de P et λ_j le j^{ieme} coefficient diagonal de D. Des relations

$$De_j = \lambda_j e_j, \quad e'_j = Pe_j, \quad D = P^{-1}AP,$$

on en déduit que $Ae'_j = \lambda_j e'_j$. La famille (e'_j) forme une base de $\mathbf{R}^n$ car la matrice P est inversible.

Exemples
Les exemples précédents donnent les relations

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

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

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1}$$

Mettre une matrice sous forme diagonale est très utile du point de vue calculatoire. Cela permet d'inverser ou de calculer les puissances de la matrice facilement. Si deux matrices A et B satisfont la relation de conjugaison $A = PBP^{-1}$ alors

$$A^n = PB^nP^{-1} \text{ et } A^{-1} = PB^{-1}P^{-1}.$$

Montrons la première égalité : $A^n = PBP^{-1}PBP^{-1} \dots PBP^{-1} = PB^nP^{-1}$.

Donnons une application concernant la suite de Fibonacci (1175 - 1250). Cette suite est définie par la relation de récurrence suivante.

$$u_0 = 0,\, u_1 = 1 \text{ et } u_{n+1} = u_n + u_{n-1} \text{ pour } n \geq 1.$$

Les premiers termes de cette suite sont donnés par

$$0,\, 1,\, 1,\, 2,\, 3,\, 5,\, 8,\, 13,\, 21,\, 34,\, 55,\, 89,\, 144,\, 233,\, 377,\, 610,\, 987,\, 1597 \dots$$

On peut écrire la relation de récurrence définissant $(u_n)_{n\geq 0}$ sous la forme

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$$

ce qui implique

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}.$$

D'après l'exemple précédent,

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}^{-1}$$

On a posé $\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$ pour alléger les expressions. On obtient

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Tout calcul fait, on obtient

$$u_n = \frac{1}{\lambda_1 - \lambda_2}(\lambda_1^n - \lambda_2^n) = \frac{\sqrt{5}}{5}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Le même principe marche pour toute suite définie par une récurrence linéaire.

2.5 Cas des valeurs propres distinctes

Nous allons voir plusieurs situations dans lesquelles il est possible d'affirmer qu'une matrice est diagonalisable sans procéder au calcul des vecteurs propres. La première situation repose sur le théorème suivant.

Théorème 10 (décomposition en sous-espaces propres) *Soit E un espace vectoriel, $f : E \to E$ une application linéaire et $\lambda_1, \ldots, \lambda_k$ des valeurs propres distinctes de f. Alors les sous-espaces propres associés à ces valeurs propres sont en somme directe :*

$$ker(f - \lambda_1 id) \oplus \ldots \oplus ker(f - \lambda_k id).$$

Démonstration
On procède par récurrence sur k. Il n'y a rien à montrer pour $k = 1$. Soit $v_i \in \ker(f - \lambda_i id)$ tels que $\sum v_i = 0$. Appliquons $f - \lambda_k id$ à cette égalité pour faire disparaître le vecteur v_k. Nous avons $(f - \lambda_k id)\, v_i = (\lambda_i - \lambda_k)\, v_i$ ce qui entraîne

$$(\lambda_1 - \lambda_k)v_1 + \ldots + (\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Par l'hypothèse de récurrence, chacun des termes de cette somme est nul et comme les valeurs propres sont distinctes, les v_i sont nuls pour $i < k$. Comme la somme de tous les v_i est nulle, le vecteur v_k est aussi nul. Le résultat est démontré.

On en déduit que f est diagonalisable si et seulement si l'une des deux conditions suivantes est réalisée,

$$E = \bigoplus_k \ker(f - \lambda_k id), \qquad dim\ E = \sum_k dim\ \ker(f - \lambda_k id).$$

Nous pouvons maintenant énoncer un premier critère de diagonalisation.

Théorème 11 (diagonalisation, valeurs propres simples) *Soit E un espace vectoriel réel de dimension finie n et soit $f : E \to E$ une application linéaire. Si le polynôme caractéristique de f possède n racines réelles distinctes, alors f est diagonalisable sur $\mathbf{R}$.*

Démonstration
Soit $\lambda_1, \ldots, \lambda_n$ les racines du polynôme caractéristique. Les sous-espaces propres $\ker(f - \lambda_k id)$ sont non vides, en somme directe. Pour chaque k, choisissons $v_k \in \ker(f - \lambda_k id)$ non nul. La famille (v_k) est libre car ces sous-espaces sont en somme directe. Elle contient autant de vecteurs que la dimension de E. C'est donc une base de vecteurs propres. La proposition est démontrée.

Le même résultat est vrai avec des espaces vectoriels sur $\mathbf{C}$.

Le théorème précédent est intéressant car il y a des conditions explicites pour déterminer si un polynôme a toutes ses racines complexes distinctes à partir des coefficients du polynôme.

Par exemple, un polynôme de degré deux a deux racines distinctes si et seulement si son discriminant est non nul. Pour un polynôme de la forme $P(X) = X^2 + aX + b$, le discriminant vaut

$$\Delta = a^2 - 4b.$$

Il existe un analogue du discriminant en tout degré. Pour un polynôme de degré 3 de la forme $P(X) = X^3 + aX^2 + bX + c$, il vaut

$$\Delta = a^2b^2 + 18abc - 27c^2 - 4b^3 - 4a^3c.$$

Il y a bien sûr des applications diagonalisables dont les racines ne sont pas toutes distinctes, comme par exemple l'identité ou les homothéties. On en verra d'autres dans la suite.

3. Transformations autoadjointes

3.1 Adjoint d'une application linéaire

On introduit la notion de transformation adjointe à une application linéaire définie sur un espace euclidien.

Définition 39 *Soit $(E,\langle .,. \rangle)$un espace euclidien et $f : E \to E$ une application linéaire. Une application linéaire $f^* : E \to E$ est dite adjointe à f si*

$$\forall\, x, y \in E, \quad \langle f(x), y\rangle = \langle x, f^*(y)\rangle.$$

Si E est de dimension finie, alors l'application adjointe f^* existe et est unique. La matrice de l'adjointe f^* dans une base orthonormée (e_i) de E est égale à la transposée de celle de f.

Vérifions qu'en dimension finie, une application f^* est adjointe à f si et seulement si sa matrice est égale à la transposée de celle de f dans une base orthonormée donnée (e_i). Rappelons que la matrice de f a pour coefficients $a_{i,j} = \langle e_i, f(e_j)\rangle$ dans une telle base. Les coefficients de la matrice de f^* sont alors égaux à

$$\langle e_i, f^*(e_j)\rangle = \langle f(e_i), e_j\rangle = \langle e_j, f(e_i)\rangle = a_{j,i}.$$

Réciproquement, si f^* est l'application associée à la matrice transposée de celle de f, on a pour tous $x,y \in E$ de coordonnées $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans la base (e_i),

$$\langle f(x), y\rangle = \sum_i \Big(\sum_j a_{i,j} x_j\Big) y_i = \sum_j \Big(\sum_i a_{i,j} y_i\Big) x_j = \langle f^*(y), x\rangle.$$

Exemples

– Une matrice P est orthogonale si ${}^tP = P^{-1}$. L'adjointe d'une application orthogonale est donc égale à son inverse.

– Les projections orthogonales sont égales à leurs adjointes. Cela est une conséquence de la proposition suivante.

Proposition 41 *Une application linéaire $p : E \to E$ définie sur un espace euclidien de dimension finie est une projection orthogonale si et seulement si elle satisfait les deux conditions suivantes.*

– $p \circ p = p$,

– $p^* = p$.

Démonstration

Soit p une projection orthogonale sur $F = \mathrm{im}(p)$ et $x,y \in E$. On sait que $p(x) \in F$ et $y - p(y) \in F^\perp$. Cela entraîne l'égalité $\langle p(x), y - p(y)\rangle = 0$ et en inversant le rôle de x et y,

$$\langle p(x), y\rangle = \langle p(x), p(y)\rangle = \langle x, p(y)\rangle.$$

Réciproquement, posons $F = \mathrm{im}(p)$ et soit $x \in E$. Le vecteur $p(x)$ est dans F, il faut vérifier que le vecteur $x - p(x)$ est orthogonal à F. Tout vecteur $y \in F$ est de la forme $y = p(z)$ pour un certain $z \in E$, ce qui implique

$$p(y) = p(p(z)) = p(z) = y.$$

Montrons que $x - p(x)$ est orthogonal à y :

$$\langle x - p(x), y\rangle = \langle x, y\rangle - \langle p(x), y\rangle = \langle x, y\rangle - \langle x, p(y)\rangle = \langle x, y\rangle - \langle x, y\rangle = 0.$$

Le théorème est démontré.

On montre de la même façon qu'une symétrie orthogonale est caractérisée par les deux égalités $s \circ s = id$ et $s^* = s$.

3.2 Diagonalisation des applications autoadjointes

On va étudier le problème de la diagonalisation pour les applications d'un espace euclidien qui sont égales à leurs adjointes.

Définition 40 *Soit* $(E,\langle .,. \rangle)$*un espace euclidien de dimension finie et soit* $f : E \to E$ *une application linéaire. L'application* f *est dite* autoadjointe *si*

$$\textit{pour tout } x \in E, \quad \langle f(x),y \rangle = \langle x,f(y) \rangle.$$

En d'autres termes, f *est autoadjointe si elle est égale à son adjointe.*

Rappelons que la matrice de l'adjoint d'une transformation est égale à la transposée de sa matrice et qu'une matrice est *symétrique* si elle est égale à sa transposée. Une transformation est donc autoadjointe si et seulement si sa matrice est symétrique.

Théorème 12 (diagonalisation des applications autoadjointes) *Soit* $(E,\langle .,. \rangle)$*un espace euclidien de dimension finie et* $f : E \to E$ *une application autoadjointe. Alors il existe une base orthonormée constituée de vecteurs propres de* f*.*

On dit que f est diagonalisable en base orthonormée. La démonstration du théorème repose sur deux lemmes.

Lemme 2 *Les valeurs propres complexes d'une application linéaire autoadjointe d'un espace euclidien de dimension finie sont toutes réelles.*

Démonstration
On se place dans une base orthonormée de E et on note A la matrice de f. Faisons agir cette matrice sur $\mathbf{C}^n$. Supposons que A admet une valeur propre $\lambda \in \mathbf{C}$. On peut alors trouver un vecteur propre $v = (v_1, \dots, v_n) \in \mathbf{C}^n$ tel que $Av = \lambda v$. Montrons que λ est en fait réelle.

Notons $\bar{v}$ le vecteur dont les coordonnées sont les conjuguées de celles de v. La matrice A est à coefficients réels et on obtient par conjugaison

$$Av = \lambda v, \quad A\bar{v} = \bar{\lambda}\bar{v}.$$

En coordonnées, le produit scalaire est défini par

$$\langle x,y \rangle = \sum_i x_i y_i$$

pour tous vecteurs x,y à coefficients réels et nous l'étendons aux vecteurs à coefficients complexes en utilisant la même égalité. Comme A est autoadjointe, nous avons

$$\langle Av,w \rangle = \langle v,Aw \rangle$$

pour tous vecteurs v, w à coefficients réels. Cette égalité s'étend aux vecteurs à coefficients complexes par bilinéarité. Au final,

$$\lambda\langle v,\bar{v}\rangle = \langle Av,\bar{v}\rangle = \langle v,A\bar{v}\rangle = \bar{\lambda}\langle v,\bar{v}\rangle.$$

On conclut que $\lambda = \bar{\lambda}$ car $\langle v,\bar{v}\rangle = \sum_i |v_i|^2$ est non nul.

Lemme 3 *Soit $(E,\langle .,.\rangle)$un espace euclidien, F un sous-espace vectoriel de E et $f : E \to E$ une application linéaire. Si $f(F) \subset F$ Alors $f^*(F^\perp) \subset F^\perp$.*

Démonstration
Soit $x \in F^\perp$, il faut montrer que pour tout vecteur $y \in F$, on a $\langle f^*(x),y\rangle = 0$. Mais

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

et ce dernier terme est nul car $f(y)$ est dans $f(F) \subset F$ et x est dans $F^\perp$.

Démonstration du théorème de diagonalisation
On procède par récurrence sur la dimension de E. Toute matrice est diagonalisable en dimension un. En dimension plus grande, d'après le théorème de d'Alembert-Gauss, rappelé en complément (théorème 17, page 108), le polynôme caractéristique de A admet une racine dans $\mathbf{C}$, la matrice A admet donc une valeur propre $\lambda_1 \in \mathbf{C}$ et par le premier lemme, nous savons que λ_1 est réelle. Notons $F = \ker(f - \lambda_1 id)$ son sous-espace propre. Par le lemme précédent, on peut restreindre $f = f^*$ à $F^\perp$ et par l'hypothèse de récurrence, il existe une base orthonormée (e_i) de $F^\perp$ constituée de vecteurs propres de f.

Nous avons vu précédemment que $F \oplus F^\perp = E$ (corollaire 7). Il suffit de rajouter à cette base (e_i) les éléments d'une base orthonormée de F pour obtenir une base orthonormée de E constituée de vecteurs propres de f. Le théorème est démontré.

3.3 Interprétation matricielle

Le théorème de diagonalisation des applications autoadjointes admet une interprétation en termes de matrices. Sous cette forme, il est dû à Karl Weierstrass (1815-1897).

Théorème 13 (diagonalisation des matrices symétriques) *Soit S une matrice symétrique. Alors il existe une matrice orthogonale P et une matrice diagonale D telles que*

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

Les termes diagonaux de D sont les valeurs propres de S et P est la matrice de passage de la base canonique à la base des vecteurs propres de S.

Notons que comme P est une matrice orthogonale, nous avons $P^{-1} = {}^tP$. Les matrices S et D sont non seulement conjuguées mais aussi congruentes.

Méthode : *diagonalisation des matrices symétriques*

On commence par calculer les valeurs propres de la matrice. Pour chaque valeur propre, on trouve une base du sous-espace propre associé en utilisant la méthode du pivot de Gauss. Il faut ensuite orthonormaliser cette base, par exemple en utilisant le procédé de Gram-Schmidt. Notons que si l'espace propre est de dimension un, il suffit de normaliser un vecteur de cet espace. L'ensemble des vecteurs appartenant à toutes ces bases forme une base orthonormée de $\mathbf{R}^n$. La matrice P recherchée a pour colonnes ces vecteurs et la matrice D a pour coefficients diagonaux les valeurs propres associées.

La proposition suivante montre que les vecteurs obtenus par la méthode précédente forment bien une base orthonormée.

Lemme 4 *Les sous-espaces propres d'une application autoadjointes sont orthogonaux entre eux.*

Démonstration
Soit λ_1 et λ_2 deux valeurs propres distinctes de f et v_1, v_2 deux vecteurs propres associés. Nous avons

$$\lambda_1\langle v_1,v_2\rangle = \langle Av_1,v_2\rangle = \langle v_1,Av_2\rangle = \lambda_2\langle v_1,v_2\rangle,$$

ce qui implique $\langle v_1,v_2\rangle = 0$.

Exemple
On considère la matrice symétrique $S = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}$.

Son polynôme caractéristique vaut $P_c(X) = \det \begin{pmatrix} X-1 & -2 & 0 \\ -2 & X-3 & -2 \\ 0 & -2 & X-1 \end{pmatrix}$.

$$\begin{aligned} P_c(X) &= (X-1)((X-3)(X-1)-4) + 2\times -2(X-1) \\ &= (X-1)\Big(X^2-4X-5\Big) \\ &= (X-1)(X+1)(X-5). \end{aligned}$$

Les valeurs propres sont égales à 1, -1 et 5.

Les vecteurs propres $x = (x_1,x_2,x_3)$ associés à la valeur propre 1 vérifient

$$\left\{\begin{array}{ccccccc} & & 2x_2 & & & = & 0 \\ 2x_1 & + & 3x_2 & + & 2x_3 & = & 0 \\ & & 2x_2 & & & = & 0 \end{array}\right.$$

ce qui donne $\ker(S - id) = \mathrm{vect}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right)$.

Les vecteurs propres $x = (x_1,x_2,x_3)$ associés à la valeur propre -1 vérifient

$$\begin{cases} 2x_1 + 2x_2 = 0 \\ 2x_1 + 4x_2 + 2x_3 = 0 \\ 2x_2 + 2x_3 = 0 \end{cases}$$

ce qui donne $\ker(S + id) = \mathrm{vect}\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right)$.

Les vecteurs propres $x = (x_1,x_2,x_3)$ associés à la valeur propre 5 vérifient

$$\begin{cases} -4x_1 + 2x_2 = 0 \\ 2x_1 - 2x_2 + 2x_3 = 0 \\ 2x_2 - 4x_3 = 0 \end{cases}$$

ce qui donne $\ker(S - 5id) = \mathrm{vect}\left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right)$.

Les espaces propres sont de dimension un, il n'est donc pas nécessaire d'utiliser le procédé d'orthonormalisation de Gram-Schmidt sur chacun de ces espaces, il suffit de normaliser pour obtenir les colonnes de la matrice P. La matrice D a pour coefficients diagonaux les valeurs propres de S. On obtient :

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

et la matrice symétrique S se diagonalise comme suit :

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}^{-1}$$

Remarquons que la matrice P est facile à inverser car son inverse est égal à sa transposée.

Le théorème 5 affirme que toute matrice symétrique S est de la forme ${}^tPDP$ avec P inversible et D une matrice diagonale dont les coefficients diagonaux appartiennent à $\{-1,0,1\}$. Expliquons comment on peut retrouver ce résultat à partir du théorème 14. La matrice D apparaissant dans ce théorème a pour coefficients diagonaux $\lambda_1, \ldots, \lambda_n$. Considérons la matrice diagonale D' dont les coefficients diagonaux sont donnés par les signes $+1$ ou -1 des λ_i pour les λ_i non nuls et zéro sinon. Notons D'' la matrice diagonale dont les coefficients diagonaux sont les racines $\sqrt{|\lambda_i|}$ pour les λ_i non nuls et 1 pour les λ_i nuls. On a alors

$$D = D''D'D'',$$

$$S = {}^tPDP = {}^tPD''D'D''P = {}^t(D''P)D'(D''P),$$

ce qui donne la décomposition du théorème 5.

3.4 Réduction simultanée

Donnons une application de la diagonalisation des matrices symétriques à l'étude des formes quadratiques.

Théorème 14 (réduction simultanée) *Soit $(E,\langle .,. \rangle)$un espace euclidien de dimension finie et Q une forme quadratique définie sur E. Alors il existe une base (e_i) de E qui est à la fois orthonormée pour le produit scalaire de E et orthogonale pour la forme quadratique Q.*

Dans une telle base, Q prend la forme

$$Q(x_1 e_1 + \ldots + x_n e_n) = \sum_{k=1}^{n} \lambda_k x_k^2,$$

où les λ_k sont les valeurs propres de la matrice de Q dans la base (e_i). La signature de Q est donnée par le signe de ces valeurs propres.

Démonstration
Soit B la matrice de la forme quadratique dans une base orthonormée (e_i) de E. Cette matrice est symétrique. D'après le théorème précédent, il existe une matrice orthogonale P et une matrice diagonale D telles que

$$D = P^{-1} B P = {}^t P B P.$$

Considérons la base (e'_i) de E dont les coordonnées des vecteurs sont données par les colonnes de P. La matrice P est la matrice de passage de la base (e_i) à la base (e'_i). La matrice P est orthogonale si bien que la base (e'_i) est orthonormée. La matrice de la forme quadratique Q dans cette base est égale à D, en vertu de la formule du changement de base pour les formes quadratiques. Cette base est donc orthogonale pour Q.

Méthode : *réduction simultanée des formes quadratiques*

Pour obtenir une base qui est simultanément orthonormée relativement à un produit scalaire et orthogonale pour une forme quadratique définie sur le même espace, on commence par écrire la matrice de la forme quadratique dans une base orthonormée relativement au produit scalaire. On diagonalise alors cette matrice en base orthonormée. La base obtenue est celle recherchée.

Exemple
Plaçons-nous sur $\mathbf{R}^3$ muni de son produit scalaire usuel et considérons la forme quadratique

$$Q(x_1,x_2,x_3) = x_1^2 + 3\,x_2^2 + x_3^2 + 4\,x_1\,x_2 + 4\,x_2\,x_3.$$

Sa matrice vaut $B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}$.

C'est la matrice de l'exemple précédent. Ses valeurs propres sont égales à 1, -1 et 5. Deux valeurs propres sont strictement positives et une valeur propre est strictement négative. La signature de Q vaut donc (2,1). La base recherchée est composée des vecteurs propres que nous avons calculés précédemment :

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix},\ \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix},\ \frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\1\end{pmatrix}.$$

Elle est orthonormée pour le produit scalaire usuel de $\mathbf{R}^3$ et orthogonale pour la forme quadratique Q. Il ne reste plus qu'à effectuer le changement de variables

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

pour simplifier l'expression de Q :

$$\begin{aligned}Q(x_1,x_2,x_3) &= x_1^2+3\,x_2^2+x_3^2+4\,x_1\,x_2+4\,x_2\,x_3\\ &= y_1^2-y_2^2+5y_3^2.\end{aligned}$$

4. Espaces préhilbertiens

Il existe une généralisation naturelle de la notion d'espace euclidien pour les espaces vectoriels complexes. Elle consiste à placer une conjugaison sur les coefficients du second vecteur du produit scalaire, de façon à ce que la forme quadratique associée soit à valeurs réelles. La théorie procède de manière analogue au cas réel.

4.1 Forme hermitienne

Définition 41 *Soit E un espace vectoriel défini sur $\mathbf{C}$. Une forme hermitienne sur E est une application $\varphi : E\times E \to \mathbf{C}$ qui satisfait pour tout z,v,w dans E et $\lambda \in \mathbf{C}$,*

- $\varphi(\lambda v + w,z) = \lambda\,\varphi(v,z) + \varphi(w,z)$,
- $\varphi(z,\lambda v + w) = \bar{\lambda}\,\varphi(z,v) + \varphi(z,w)$,
- $\varphi(v,w) = \overline{\varphi(w,v)}$.

La forme quadratique associée $Q(z) = \varphi(z,z)$ est à valeurs réelles en raison de la conjugaison complexe qui apparaît dans la définition.

Exemples
– Sur $\mathbf{C}^n$, le produit hermitien usuel est défini par

$$\varphi(z,w) = \sum_{k=1}^{n} z_k \overline{w_k} = z_1\overline{w_1} + \ldots + z_n\overline{w_n}.$$

pour tout $z = (z_1, \ldots, z_n) \in \mathbf{C}^n$, $w = (w_1, \ldots, w_n) \in \mathbf{C}^n$. La forme quadratique associée vaut

$$Q(z) = \sum_{k=1}^{n} |z_k|^2 = |z_1|^2 + \ldots + |z_n|^2.$$

– Sur $C^0([0,1],\mathbf{C})$, le produit hermitien usuel est défini par

$$\varphi(f,g) = \int_0^1 f(t)\overline{g(t)}\, dt.$$

La forme quadratique associée vaut

$$Q(f) = \int_0^1 |f(t)|^2\, dt.$$

Une base (e_i) de E sur $\mathbf{C}$ étant donnée, on pose $B_{i,j} = \varphi(e_i,e_j)$ afin de représenter une forme hermitienne φ par une matrice B satisfaisant ${}^tB = \overline{B}$. Notons $z = (z_i)$ et $w = (w_i)$ les coordonnées de deux vecteurs de E dans la base (e_i), nous avons alors la formule

$$\varphi(z,w) = {}^tz\, B\, \overline{w} = \sum_{i,j} z_i\, b_{i,j}\, \overline{w_j} = b_{1,1}\, z_1\overline{w_1} + b_{1,2}\, z_1\overline{w_2} + \ldots + b_{n,n}\, z_n\overline{w_n}.$$

4.2 Réduction des formes hermitiennes

La théorie des formes hermitiennes procède de manière analogue au cas réel. En particulier, on a une notion de signature et un algorithme de Gauss qui repose sur les règles suivantes, avec z, w, a, a_1, $a_2 \in \mathbf{C}$.

Règle 1 : $|z|^2 + 2Re(z\overline{a}) = |z + a|^2 - |a|^2$.

Règle 2 : $z\overline{w} + z\overline{a_1} + a_2\overline{w} = (z + a_2)\overline{(w + a_1)} - a_2\overline{a_1}$.

Règle 3 : $4Re(z\overline{w}) = |z + w|^2 - |z - w|^2$.

La partie réelle de z a été notée $Re(z)$, elle vaut $Re(z) = \dfrac{z + \overline{z}}{2}$.

Proposition 42 *Soit E un espace vectoriel complexe et soit Q une forme quadratique hermitienne de signature (p,q). Alors il existe $p + q$ formes $\mathbf{C}$-linéaires indépendantes l_1, ..., l_{p+q} définies sur E telles que*

$$\forall z \in E, \quad Q(z) = \sum_{i=1}^{p} |l_i(z)|^2 - \sum_{j=p+1}^{p+q} |l_j(z)|^2.$$

Notons le module qui apparaît dans la proposition précédente. Les formes l_i sont à valeurs complexes et linéaires sur $\mathbf{C}$.

Exemples
Plaçons-nous sur $E = \mathbf{C}^2$ et posons $z = \binom{z_1}{z_2}$, $w = \binom{w_1}{w_2} \in \mathbf{C}^2$.

– $\varphi(z,w) = z_1\overline{w_1} + 3z_2\overline{w_2} + 2z_1\overline{w_2} + 2z_2\overline{w_1}$.

Sa matrice vaut $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. L'algorithme de Gauss donne

$$\begin{aligned} Q(z) &= |z_1|^2 + 3|z_2|^2 + 2z_1\overline{z_2} + 2z_2\overline{z_1} \\ &= |z_1|^2 + 3|z_2|^2 + 4Re(z_1\overline{z_2}) \\ &= |z_1 + 2z_2|^2 - |z_2|^2. \end{aligned}$$

Nous obtenons la décomposition $Q(z) = |l_1(z_1,z_2)|^2 - |l_2(z_1,z_2)|^2$ avec

$$l_1(z_1,z_2) = z_1 + 2z_2, \quad l_2(z_1,z_2) = z_2.$$

– $\varphi(z,w) = z_1\overline{w_1} + 3z_2\overline{w_2} + 2iz_1\overline{w_2} - 2iz_2\overline{w_1}$.

Sa matrice vaut $\begin{pmatrix} 1 & 2i \\ -2i & 3 \end{pmatrix}$. L'algorithme de Gauss donne

$$\begin{aligned} Q(z) &= |z_1|^2 + 3|z_2|^2 + 2iz_1\overline{z_2} - 2iz_2\overline{z_1} \\ &= |z_1|^2 + 3|z_2|^2 + 4Re(z_1(\overline{-iz_2})) \\ &= |z_1 - 2iz_2|^2 - |z_2|^2. \end{aligned}$$

Cette fois-ci, $Q(z) = |l_1(z_1,z_2)|^2 - |l_2(z_1,z_2)|^2$ avec

$$l_1(z_1,z_2) = z_1 - 2iz_2, \quad l_2(z_1,z_2) = z_2.$$

Définition 42 *Un espace* préhilbertien *est un espace vectoriel complexe muni d'une forme hermitienne dont la forme quadratique associée est définie positive. Une telle forme est appelée* produit scalaire hermitien.

Utilisons la notation $\langle .,. \rangle$ pour un produit scalaire hermitien. L'inégalité de Cauchy-Schwarz devient

$$\forall\, z, w \in E, \quad |\langle z,w \rangle|^2 \leq Q(z)Q(w)$$

et on peut définir la notion de projection orthogonale comme dans le cas réel.

4.3 Diagonalisation des matrices normales

En matière de diagonalisation, on introduit les notions suivantes.

Définition 43 *Soit E un espace préhilbertien et $f : E \to E$ une application $\mathbf{C}$-linéaire. On dit que f admet un adjoint s'il existe une application $\mathbf{C}$-linéaire $f^* : E \to E$ qui satisfait pour tout $z,w \in E$,*

$$\langle f(z),w\rangle = \langle z,f^*(w)\rangle.$$

L'application f est dite

- *hermitienne si $f = f^*$,*
- *unitaire si $f^* \circ f = id$,*
- *normale si $f \circ f^* = f^* \circ f$.*

En dimension finie, toute application f a un adjoint. Si A est la matrice de f dans une base orthonormée, son adjoint a pour matrice ${}^t\bar{A}$. On pose

$$A^* = {}^t\bar{A}.$$

La notion d'application hermitienne généralise la notion d'application autoadjointe dans le cas réel. La notion d'application unitaire généralise celle d'application orthogonale. La nouveauté est la notion d'application normale. Notons que les applications hermitiennes et unitaires sont normales. Le théorème de diagonalisation s'énonce comme suit.

Théorème 15 (théorème spectral) *Soit E un espace préhilbertien et soit $f : E \to E$ une application normale. Alors f est diagonalisable en base orthonormée.*

Lorsqu'on travaille en coordonnées sur $\mathbf{C}^n$, on dit qu'une matrice A de taille $n \times n$ à coefficients complexes est hermitienne, unitaire ou normale si $A = A^*$, $A^*A = id$ ou $A^*A = AA^*$ respectivement. L'ensemble des matrices unitaires de taille $n \times n$ est noté

$$U_n(\mathbf{C}) = \{M \in M_n(\mathbf{C}) \mid {}^t\bar{M}M = id\}.$$

Ce théorème implique que les matrices unitaires sont diagonalisables sur $\mathbf{C}$ alors que les matrices orthogonales ne sont pas diagonalisables sur $\mathbf{R}$.

Théorème 16 (diagonalisation des matrices normales) *Soit A une matrice à coefficients complexes vérifiant*

$$A^*A = AA^*.$$

Alors il existe une matrice diagonale D et une matrice unitaire U telles que

$$A = UDU^*.$$

Les coefficients diagonaux de D sont les valeurs propres de A. Ils sont réels si et seulement si A est hermitienne. Ils sont de module un si et seulement si elle est unitaire.

En particulier, toute matrice à coefficients réels qui est orthogonale est unitaire donc diagonalisable sur **C**.

La démonstration du théorème spectral dans le cas hermitien est similaire au cas réel. Pour les matrices normales, on les décompose sous la forme $A' + iA''$ avec A' et A'' des matrices hermitiennes qui commutent, en posant $A' = \frac{1}{2}(A + A^*)$ et $A'' = \frac{1}{2i}(A - A^*)$. Comme ces deux matrices commutent, les sous-espaces propres de A' sont invariants par A''. Il devient possible de diagonaliser A' puis ensuite de diagonaliser A'' sur chacun des sous-espaces propres de A'' de façon à trouver une base qui diagonalise simultanément les deux matrices.

Pour ce qui est des valeurs propres de f, soit $x \in E$ et $\lambda \in \mathbf{C}$ tels que $f(x) = \lambda x$. Ceci entraîne les égalités

$$\bar{\lambda}\langle x,x\rangle = \langle x,f(x)\rangle = \langle f^*(x),x\rangle$$

expression qui est égale à $\lambda\langle x,x\rangle$ dans le cas hermitien et $\lambda^{-1}\langle x,x\rangle$ dans le cas unitaire. On a donc $\lambda = \bar{\lambda}$ dans le premier cas et $\lambda = \bar{\lambda}^{-1}$ dans le second cas, c'est-à-dire $|\lambda|^2 = 1$.

Réciproquement, si D est diagonalisable à coefficients diagonaux réels, on aura $D^* = D$ et $(UDU^*)^* = UDU^*$ pour toute matrice U unitaire. Si D est diagonalisable à coefficients diagonaux de module un, on aura $D^* = D^{-1}$ et $(UDU^*)^{-1} = (UDU^*)^*$ pour toute matrice U unitaire, ce qui termine la démonstration du théorème spectral.

Exemple
Diagonalisation des rotations du plan hermitien.

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

5. Compléments

Donnons deux applications de la réduction simultanée.

5.1 Comparaison des normes euclidiennes

La première permet de comparer des normes euclidiennes. Soit $(E,\langle .,. \rangle)$un espace euclidien muni d'une forme quadratique Q définie positive. Nous avons deux structures euclidiennes sur E, une associée au produit scalaire $\langle .,. \rangle$ et l'autre à la forme bilinéaire venant de Q. Nous voulons les comparer.

Proposition 43 *Soit $(E,\langle .,. \rangle)$un espace euclidien et Q une forme quadratique définie sur E. Notons $\lambda_1, \ldots, \lambda_k$ les valeurs propres de la matrice de Q dans une base orthonormée pour $\langle .,. \rangle$. Alors pour tout $x \in E$,*

$$(\min_k \lambda_k) \ \|x\|^2 \leq Q(x) \leq (\max_k \lambda_k) \ \|x\|^2.$$

Démonstration
Il suffit de se placer dans une base (e_i) de diagonalisation simultanée.

$$Q(x_1e_1 + \ldots + x_ne_n) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 \leq (\max_k \lambda_k)(x_1^2 + \ldots + x_n^2).$$

On procède de manière analogue pour l'autre inégalité.

On voit que les normes associées à deux formes quadratiques définies positives sont équivalentes. Remarquons qu'on a égalité dans la majoration de $Q(x)$ pour les vecteurs appartenant au sous-espace propre associé à la plus grande valeur propre.

Exemple
La forme quadratique définie sur $\mathbf{R}^2$ par

$$Q(x_1,x_2) = x_1^2 + x_1x_2 + x_2^2$$

a pour matrice $\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$ dans la base canonique et pour polynôme caractéristique

$$P_c(X) = (X-1)^2 - (1/2)^2 = (X - 1/2)(X - 3/2).$$

Les racines de ce polynôme valent $\frac{1}{2}$ et $\frac{3}{2}$. On en déduit l'encadrement

$$\frac{1}{2}(x_1^2 + x_2^2) \ \leq \ x_1^2 + x_1x_2 + x_2^2 \ \leq \ \frac{3}{2}(x_1^2 + x_2^2).$$

5.2 Caractéristiques géométriques des coniques

Rappelons qu'une conique est un sous-ensemble de $\mathbf{R}^2$ défini par une équation de la forme

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

On se restreint au cas où la conique est une ellipse ou une hyperbole pour alléger la présentation : $b^2 - 4ac \neq 0$.

On peut reprendre les méthodes du chapitre portant sur les coniques en substituant à la réduction de Gauss la diagonalisation simultanée. On obtient le résultat suivant.

Proposition 44 *Considérons une conique $\mathcal{C}$ de $\mathbf{R}^2$ associée à une forme quadratique non dégénérée : $b^2 - 4ac \neq 0$. Alors il existe un changement de variables de la forme*

$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + b_2 \end{cases}$$

avec $\{a_{i,j}\}$ matrice orthogonale, qui transforme l'expression de la conique en

$$\lambda_1 X^2 + \lambda_2 Y^2 = \mu.$$

Les coefficients λ_1 et λ_2 sont les valeurs propres de la matrice associée à la forme quadratique $ax^2 + bxy + cy^2$.

Comme le changement de variable est associé à une matrice orthogonale, il préserve la distance euclidienne et il est possible de calculer les caractéristiques métriques de la conique à partir de son expression réduite. Pour une ellipse par exemple, les demi-longueurs du grand axe et du petit axe sont données par $a = \sqrt{\mu/\lambda_1}$ et $b = \sqrt{\mu/\lambda_2}$.

Les coefficients μ, λ_1 et λ_2 n'étant déterminés qu'à un facteur multiplicatif près, il est naturel d'introduire un invariant métrique reposant sur le quotient de deux de ces nombres.

Définition 44 *Quitte à intervertir les variables X et Y dans la forme réduite d'une conique et à changer son signe, on peut supposer $\lambda_2 \geq \lambda_1$ et $\lambda_2 \geq 0$.*

L'excentricité $e \in [0,\infty[$ d'une conique est alors définie par $e = \sqrt{1 - \frac{\lambda_1}{\lambda_2}}$.

L'excentricité d'une ellipse est strictement inférieure à 1, celle d'une hyperbole strictement supérieure à 1. Par convention, on pose $e = 1$ pour les paraboles. Une ellipse proche d'un cercle a une excentricité proche de 0.

En exercice, on peut vérifier par un calcul direct que la courbe définie en coordonnées polaires (r, θ) par l'équation

$$r(\theta) = \frac{p}{1 + e\cos(\theta)}$$

est une conique d'excentricité e.

5.3 Théorème de d'Alembert-Gauss

Le théorème suivant est dû à Jean Le Rond d'Alembert (1717-1783) et Carl Friedrich Gauss (1777-1855). C'est un résultat que nous avons utilisé dans la preuve du théorème de diagonalisation des applications autoadjointes.

Théorème 17 (d'Alembert-Gauss) *Tout polynôme non constant défini sur* $\mathbf{C}$ *à coefficients complexes admet une racine complexe.*

Démonstration
On procède par l'absurde. Soit $P(X)$ un polynôme non constant défini sur $\mathbf{C}$ sans racine. Comme $|P(z)|$ tend vers l'infini quand $|z|$ tend vers l'infini et que $|P(z)|$ ne s'annule pas, il possède un minimum non nul atteint en un point $z_0 \in \mathbf{C}$. On peut le supposer égal à 1 et atteint en 0, en remplaçant $P(z)$ par $P(z+z_0)/P(z_0)$.

Nous pouvons écrire le polynôme $P(z)$ sous la forme

$$P(z) = 1 + cz^{n_0} + cz^{n_0}Q(z)$$

pour un certain entier n_0, un coefficient $c \neq 0$ et $Q(z)$ un polynôme qui s'annule en 0. Soit $\alpha \in \mathbf{C}$ tel que $\alpha^{n_0} = -1/c$. Il suffit de se déplacer dans la direction α pour faire décroître le module de P et obtenir une contradiction. Soit $t \in]0,1[$ tel que $|Q(\alpha t)| \leq 1/2$. Alors

$$P(\alpha t) = 1 - t^{n_0} - t^{n_0}Q(\alpha t),$$

$$|P(\alpha t)| \leq 1 - t^{n_0} + t^{n_0}\,|Q(\alpha t)| \leq 1 - t^{n_0}/2 < 1.$$

Le résultat est démontré.

Le théorème de d'Alembert-Gauss implique que tout polynôme complexe est un produit de polynômes de degré un.

Proposition 45 *Soit* $P(X)$ *un polynôme à coefficients complexes de degré* $n > 0$. *Alors il existe des nombres complexes* $\lambda_1, \ldots, \lambda_n$, *pas nécessairement distincts, et un nombre complexe* C *tels que*

$$P(X) = C\,(X-\lambda_1)(X-\lambda_2)\ldots(X-\lambda_n).$$

La démonstration se fait par récurrence sur le degré. Le polynôme $P(X)$ possède une racine λ_1 par le théorème précédent. Par division euclidienne, il existe un polynôme $Q(X)$ de degré un de moins que le degré de $P(X)$ tel que $P(X) = (X-\lambda_1)Q(X)$. On peut alors appliquer l'hypothèse de récurrence à $Q(X)$ s'il est de degré non nul. Dans le cas contraire, il est constant et on conclut.

La preuve du théorème de d'Alembert-Gauss ne donne pas de moyen explicite pour calculer les racines d'un polynôme. C'est une preuve non constructive qui affirme l'existence d'un objet mathématique, un zéro de $P(X)$, sans donner de méthode pour le construire. En pratique, on a recours à des algorithmes comme la méthode de Newton pour obtenir des valeurs approchées des racines réelles ou complexe d'un polynôme.

5.4 Racine carrée des matrices symétriques

Le théorème de diagonalisation des matrices réelles symétriques permet de montrer que ces matrices possèdent une racine carrée, et même une unique racine carrée positive dans le cas positif.

Proposition 46 *Soit B une matrice symétrique positive. Alors il existe une unique matrice symétrique positive C satisfaisant $C^2 = B$.*

On dit que C est une racine carrée de B. La matrice C étant symétrique,

$$B = C^2 = {}^tCC$$

ce qui donne une nouvelle preuve du théorème 5 dans le cas défini positif.

Pour montrer l'existence de C, il suffit de diagonaliser $B = PDP^{-1}$, avec P une matrice orthogonale. La matrice D est diagonale et ses coefficients diagonaux sont positifs par hypothèse. Il est donc possible de prendre la racine carrée de chacun de ses coefficients diagonaux pour former une matrice D' satisfaisant $D'^2 = D$. La matrice recherchée vaut $C = PD'P^{-1}$. Cette méthode marche avec toute matrice diagonalisable dont les valeurs propres sont positives.

Pour montrer l'unicité, remarquons que C commute avec $C^2 = B$. Les matrices B et C sont donc diagonalisables dans la même base. Il y a alors un seul choix possible pour la matrice diagonale associée à C, elle s'obtient en prenant les racines carrées positives des éléments diagonaux de la matrice diagonale associée à B.

Il est possible d'exprimer explicitement la racine carrée d'une matrice symétrique positive en fonction de ses coefficients mais la formule devient compliquée lorsque la taille de la matrice s'accroît. Voici ce que l'on obtient pour une matrice de taille 2×2.

$$\sqrt{\begin{pmatrix} a & b \\ b & c \end{pmatrix}} = \frac{1}{\sqrt{a + c + 2\sqrt{ac - b^2}}} \begin{pmatrix} a + \sqrt{ac - b^2} & b \\ b & c + \sqrt{ac - b^2} \end{pmatrix}.$$

La matrice considérée est positive si et seulement si $a \geq 0$, $c \geq 0$ et $ac \geq b^2$.

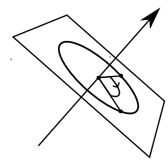
Remarquons enfin que la matrice identité de taille $n \times n$ possède une seule racine carrée symétrique positive, elle-même, mais qu'elle possède beaucoup d'autres racines carrées symétriques, en l'occurrence toutes les matrices associées aux symétries orthogonales relativement à n'importe quel sous-espace vectoriel de $\mathbf{R}^n$.

5.5 Isométries des espaces euclidiens de dimension 3

Étudions les isométries d'un espace euclidien de dimension 3 comme $\mathbf{R}^3$.

Proposition 47 *Les isométries linéaires d'un espace vectoriel euclidien de dimension trois sont les rotations autour d'un axe, ou les composées de telles rotations avec la réflexion orthogonale relativement au plan orthogonal à l'axe.*

Une *rotation autour d'un axe* est une application linéaire qui fixe les points d'une droite vectorielle, laisse invariant le plan orthogonal à cette droite et agit comme une rotation dans ce plan.



Remarquons que la symétrie centrale égale à $-id$ est composée de la rotation d'angle π autour d'une droite arbitraire et de la réflexion par relativement au plan orthogonal à la droite.

Démonstration
Soit f une isométrie, son polynôme caractéristique est unitaire de degré trois, il tend vers $+\infty$ en $+\infty$ et vers $-\infty$ en $-\infty$. Par le théorème des valeurs intermédiaires, ce polynôme possède une valeur propre réelle λ.

Nous avons vu que les valeurs propres d'une matrice unitaire sont de module 1 lors de la preuve du théorème spectral. La matrice de f dans une base orthonormée est orthogonale donc unitaire, la valeur propre λ est égale à 1 ou -1. Notons D la droite engendrée par un vecteur propre associé. Elle est invariante par f, son orthogonal $D^\perp$ est invariant par $f^* = f^{-1}$ par le lemme 3. L'application f se restreint au plan $D^\perp$ où elle agit comme une isométrie.

D'après le théorème de classification des isométries planes obtenu au chapitre précédent (corollaire 9), si cette isométrie $f_{|D^\perp}$ est directe, c'est une rotation du plan $D^\perp$ et nous obtenons les isométries mentionnées dans la proposition, en fonction du signe de λ.

Si cette isométrie renverse l'orientation, c'est une symétrie orthogonale du plan et f est aussi une symétrie orthogonale. Il peut s'agir de l'identité, de la symétrie centrale, d'une réflexion orthogonale ou d'une rotation d'angle π autour d'un axe. Dans tous les cas, elle est bien de la forme recherchée.

Il existe une formule due au mathématicien français Olinde Rodrigues (1795-1851) pour la rotation $R_{u,\theta} : \mathbf{R}^3 \to \mathbf{R}^3$ d'angle $\theta \in \mathbf{R}/2\pi\mathbf{Z}$ et d'axe dirigé par un vecteur u de norme 1 :

$$R_{u,\theta}(x) = \cos(\theta)\, x + \sin(\theta)\, u \times x + (1 - \cos(\theta))\, \langle x,u \rangle\, u.$$

Elle se démontre en décomposant x en un vecteur proportionnel à u et un vecteur qui lui est orthogonal. On a utilisé les notations usuelles

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \rangle = x_1y_1 + x_2y_2 + x_3y_3, \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Chapitre 5

Annexes

1. Erreurs courantes

Voici quelques remarques motivées par des erreurs trouvées trop souvent dans les copies d'examen.

1) Le noyau d'une matrice n'est jamais vide.

Il contient toujours le vecteur nul. Ce vecteur est habituellement noté 0. Dans $\mathbf{R}^n$, il a pour coordonnées $(0, 0, \ldots, 0)$. Attention à ne pas confondre les notations suivantes :

– l'ensemble vide : $\emptyset$,
– l'ensemble contenant le vecteur nul : $\{0\}$.

2) Le noyau d'une matrice est un espace vectoriel.

Ce n'est jamais un vecteur. Lorsqu'une matrice est inversible, son noyau contient un seul élément, le vecteur nul.

3) Un sous-espace vectoriel de $\mathbf{R}^n$ est de dimension inférieure ou égale à n.

Il ne suffit pas de compter le nombre de vecteurs dans une famille génératrice pour en déduire sa dimension. En particulier, l'espace vectoriel engendré par les vecteurs suivants n'est pas de dimension trois :

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

4) Un système de k équations cartésiennes dans $\mathbf{R}^n$ ne définit pas nécessairement un sous-espace vectoriel de dimension $n - k$.

La dimension d'un sous-espace vectoriel défini par un système linéaire homogène s'obtient en calculant la dimension du noyau de la matrice A associée au système. Elle est égale à $n - rang(A)$.

5) Le déterminant est défini pour les matrices carrées uniquement.

Il est inutile de chercher à calculer le déterminant d'une matrice non carrée. Le déterminant des matrices suivantes n'est pas défini.

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{pmatrix}, \quad \begin{pmatrix} 1 & 4 & 0 & 0 & 5 \\ 0 & 0 & 1 & 1 & 4 \\ 1 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

6) Si le noyau d'une forme quadratique Q est nul, il peut malgré tout exister des vecteurs x non nuls tels que $Q(x) = 0$.

Il ne faut pas confondre le cône des vecteurs isotropes $\{x \mid Q(x) = 0\}$ et le noyau d'une forme quadratique $\{y \mid \forall\, x,\ \varphi(x,y) = 0\}$. Le noyau est inclu dans le cône mais la réciproque n'est pas toujours vraie.

7) Une conique est un sous-ensemble du plan.

La conique d'équation $X^2 + Y^2 = 1$ est un cercle.
La quadrique d'équation $X^2 + Y^2 = 1$ est un cylindre.

8) Des vecteurs peuvent être linéairement dépendants sans être proportionnels deux à deux.

C'est toujours le contre-exemple $\binom{0}{1}, \binom{1}{0}, \binom{1}{-1}$. Remarquons par contre que des vecteurs non nuls orthogonaux deux à deux sont orthogonaux entre eux et linéairement indépendants.

9) Il existe des matrices diagonalisables qui ne sont pas symétriques.

Une matrice symétrique est diagonalisable mais la réciproque n'est pas vraie. Les matrices suivantes sont diagonalisables sur $\mathbf{R}$ sans être symétriques.

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

10) L'inégalité de Cauchy-Schwarz ne prend pas de t.

Il s'agit de Hermann Amandus Schwarz (1843-1921) et non de Laurent Schwartz (1915-2002).

11) Attention aux nombreux sens du mot orthogonal.

Il ne suffit pas que les colonnes d'une matrice forment une famille orthogonale pour que la matrice soit orthogonale. Rappelons aussi que la matrice d'une projection orthogonale n'est pas orthogonale en général.

12) On obtient une base à la fois orthonormée pour un produit scalaire et orthonormale pour une forme quadratique grâce à l'algorithme de réduction simultanée.

Construire une base orthogonale pour la forme quadratique par l'algorithme de Gauss puis la transformer en base orthonormée par le procédé d'orthonormalisation de Gram-Schmidt ne donne pas un résultat correct.

13) Les calculs doivent être menés jusqu'au bout.

Un résultat de la forme $\frac{1{,}45}{12} + 1$ ne rapporte aucun point à l'examen. Il est conseillé d'utiliser pour les réponses numériques un format en accord avec celui de l'énoncé. Les fractions doivent être mises sous forme irréductible, $\frac{2}{3}$ plutôt que $\frac{12}{18}$. Les racines doivent figurer au numérateur, $\frac{\sqrt{5}}{5}$ plutôt que $\frac{1}{\sqrt{5}}$.

2. Formulaire

Dans ce qui suit,
E est un espace vectoriel muni de deux bases $(e_i)_{i=1..n}$, $(e'_i)_{i=1..n}$,
P est la matrice de passage entre ces deux bases,
$(x_1, \ldots, x_n)$ sont les coordonnées d'un vecteur x dans la base (e_i),
$(y_1, \ldots, y_n)$ sont les coordonnées d'un vecteur y dans la base (e_i),
$(x'_1, \ldots, x'_n)$, $(y'_1, \ldots, y'_n)$, sont les analogues dans la base (e'_i),
φ est une forme bilinéaire sur E,
B est la matrice de b dans la base (e_i),
B' est la matrice de b dans la base (e'_i),
Q est la forme quadratique associée à φ,
X, Y sont des nombres réels,
v_1, ..., v_n sont des vecteurs de E,
θ est l'angle, orienté ou non, entre x et y,
det est le déterminant,
A est une matrice,
${}^tA$ est sa transposée,
f est une application linéaire,
f^* est son adjointe.

Formes quadratiques

Forme bilinéaire

$$\varphi(x,y) = {}^tx\, B\, y = \sum_{i,j} x_i\, b_{i,j}\, y_j = b_{1,1}\, x_1 y_1 + b_{1,2}\, x_1 y_2 + \ldots + b_{n,n}\, x_n y_n$$

Formule du rang

$$dim\ \ker(\varphi) + rang(\varphi) = dim\ E$$

Changement de base pour les vecteurs

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

Changement de base pour les applications linéaires

$$A' = P^{-1}\, A\, P$$

Changement de base pour les formes bilinéaires

$$B' = {}^tP\, B\, P$$

Forme quadratique

$$Q\Big(\sum_{k=1}^{n} x_k e_k\Big) = {}^t x B x = \sum_{i,j} b_{i,j} x_i x_j = \sum_i b_{i,i}\, x_i^2 + 2 \sum_{\substack{i,j \\ i<j}} b_{i,j} x_i x_j$$

Formules de polarisation

$$Q(x+y) = Q(x) + 2\varphi(x,y) + Q(y)$$

$$Q(x-y) = Q(x) - 2\varphi(x,y) + Q(y)$$

$$\varphi(x,y) = \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big)$$

$$\varphi(x,y) = \frac{1}{2}\Big(Q(x) + Q(y) - Q(x-y)\Big)$$

$$\varphi(x,y) = \frac{1}{4}\Big(Q(x+y) - Q(x-y)\Big)$$

Identités remarquables

$$X^2 + Xa = \Big(X + \frac{a}{2}\Big)^2 - \frac{a^2}{4}$$

$$XY + Xa_1 + Ya_2 = (X + a_2)(Y + a_1) - a_1 a_2$$

$$XY = \frac{1}{4}(X+Y)^2 - \frac{1}{4}(X-Y)^2$$

Espaces euclidiens

Décomposition en base orthonormée (e_i)

$$x = \sum_{i=1}^{n} \langle x,e_i\rangle\, e_i = \langle x,e_1\rangle\, e_1 + \ldots + \langle x,e_n\rangle\, e_n$$

Produit scalaire

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n,$$

Carré de la norme

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 = x_1^2 + \ldots + x_n^2$$

Inégalité de Cauchy-Schwarz

$$|\langle x,y\rangle| \leq \|x\|\|y\|$$

Inégalité triangulaire

$$\|x+y\| \leq \|x\| + \|y\|$$

Aire des parallélogrammes

$$aire(x,y) = |x_1y_2 - x_2y_1| \quad \text{dans une base orthonormée}$$

$$\langle x,y\rangle^2 + aire(x,y)^2 = \|x\|^2\|y\|^2$$

Identité de Lagrange

$$(x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2)$$

Formule du parallélogramme

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

$$2AB^2 + 2BC^2 = AC^2 + BD^2$$

Formule de la médiane

$$\|x\|^2 + \|y\|^2 = \frac{1}{2}\|x-y\|^2 + 2\left\|\frac{x+y}{2}\right\|^2$$

$$AB^2 + AC^2 = \frac{1}{2}BC^2 + 2AI^2$$

Théorème de Pythagore

$$\|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2 \quad \text{si } \langle v_i,v_j\rangle = 0 \text{ pour tout } i \neq j$$

Relations entre aire, longueur et angle

$$\langle x,y\rangle = \|x\|\,\|y\|\,\cos(\theta)$$

$$\det(x,y) = \|x\|\,\|y\|\,\sin(\theta) \qquad (\theta \text{ orienté})$$

Espace orthogonal en dimension finie

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

$$dim\ F + dim\ F^{\perp} = dim\ E$$

Projection sur vect*$(e_1,\ \ldots,\ e_n)$*

$$p(x) = \sum_i \langle x,e_i\rangle\, e_i \quad \text{si } (e_i) \text{ est orthonormée}$$

Projection sur $\mathrm{vect}(e_1,\ldots,e_n)^{\perp}$

$$p(x) = x - \sum_i \langle x,e_i\rangle\, e_i \quad \text{si } (e_i) \text{ est orthonormée}$$

Projection sur la droite vect*(v)*

$$p(x) = \frac{\langle x,v\rangle}{\|v\|^2}\, v$$

Projection sur l'hyperplan $\mathrm{vect}(v)^\perp$

$$p(x) = x - \frac{\langle x,v\rangle}{\|v\|^2}\, v$$

Symétrie orthogonale

$$s(x) = 2p(x) - x$$

Procédé d'orthonormalisation de Gram-Schmidt

$$e_1 = \frac{v_1}{\|v_1\|}$$

$$e_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1},e_i\rangle\, e_i}{\|v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1},e_i\rangle\, e_i\|}$$

Déterminants

Développement relativement à la première colonne

$$\det\Big(\{a_{i,j}\}_{\substack{i=1..m\\ j=1..n}}\Big) = \sum_{k=1}^{n} (-1)^{k+1}\, a_{k,1}\, \det(\{a_{l,m}\}_{l\neq k, m\neq 1})$$

Déterminant d'une matrice 2×2

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Déterminant d'une matrice 3×3

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + dhc + gbf - ceg - fha - ibd$$

Déterminant d'un produit

$$\det(AB) = \det(A)\det(B)$$

Déterminant de l'inverse

$$\det(A^{-1}) = \det(A)^{-1}$$

Déterminant d'un multiple

$$\det(\lambda A) = \lambda^n \det(A)$$

Déterminant de la transposée

$$\det({}^t A) = \det(A)$$

Formules de Cramer

$$\text{si } Ax = y, \text{ alors} \quad x_j = \frac{\det(A_1, \ldots, A_{j-1}, y, A_{j+1}, \ldots, A_n)}{\det(A_1, \ldots, A_n)}$$

Diagonalisation

Dimension d'une somme directe

$$dim\ E_1 \oplus \ldots \oplus E_n = dim\ E_1 + \ldots + dim\ E_n$$

Dimension d'une somme

$$dim\ (E_1 + E_2) = dim\ E_1 + dim\ E_2 - dim\ (E_1 \cap E_2)$$

Polynôme caractéristique

$$P_c(X) = \det(Xid - f)$$

Caractérisation des projections orthogonales

$$p \circ p = p, \quad p^* = p$$

Caractérisation des symétries orthogonales

$$s \circ s = id, \quad s^* = s$$

Caractérisation de l'adjoint

$$\forall\, x, y \in E, \quad \langle f(x), y\rangle = \langle x, f^*(y)\rangle$$

Excentricité d'une conique

$$e = \sqrt{1 - \frac{\lambda_1}{\lambda_2}} \quad \text{si } \lambda_2 \geq \lambda_1 \text{ et } \lambda_2 > 0$$

Espaces préhilbertiens

Forme hermitienne

$$\varphi(z,w) = {}^t z\, B\, \bar{w} = \sum_{i,j} z_i\, b_{i,j}\, \bar{w}_j = b_{1,1}\, z_1\bar{w}_1 + b_{1,2}\, z_1\bar{w}_2 + \ldots + b_{n,n}\, z_n\bar{w}_n$$

Produit scalaire hermitien

$$\langle z,w\rangle = \sum_{i=1}^{n} z_i\bar{w}_i = z_1\bar{w}_1 + \ldots + z_n\bar{w}_n$$

Carré de la norme hermitienne

$$\|z\|^2 = \sum_{i=1}^{n} |z_i|^2 = |z_1|^2 + \ldots + |z_n|^2$$

Identités remarquables

$$|z|^2 + 2Re(z\overline{a}) = |z + a|^2 - |a|^2$$

$$z\overline{w} + z\overline{a_1} + a_2\overline{w} = (z + a_2)\overline{(w + a_1)} - a_2\overline{a_1}$$

$$4Re(z\overline{w}) = |z + w|^2 - |z - w|^2$$

3. Méthodes

Voici la liste des méthodes présentées dans ce texte.

Notations

Les ensembles des nombres entiers, entiers relatifs, rationnels, réels et complexes sont notés respectivement $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ et $\mathbf{C}$.

A, P .... matrices
$a_{i,j}$, $b_{i,j}$ .... coefficients de matrices
acos .... fonction arccosinus
φ .... forme bilinéaire
B .... matrice d'une forme bilinéaire
$C^0([0,1],\mathbf{R})$ .... espace des applications continues sur [0,1]
$\mathcal{C}$ .... conique
det .... déterminant
Δ .... discriminant
dim .... dimension
E .... espace vectoriel
F .... sous-espace vectoriel de E
$F^\perp$ .... orthogonal de F
H .... hyperplan
f .... application linéaire
f^* .... application adjointe de f
id .... application identité
im .... image
ker .... noyau
l .... forme linéaire
$M_n(\mathbf{R})$ .... ensemble des matrices de taille $n \times n$
$\|x\|$ .... norme de x
$\emptyset$ .... ensemble vide
$O_n(\mathbf{R})$ .... ensemble des matrices orthogonales
$SO_n(\mathbf{R})$ .... matrices orthogonales de déterminant un
$P_c(X)$ .... polynôme caractéristique
(p,q) .... signature
Q .... forme quadratique
$\mathcal{Q}$ .... quadrique
S .... matrice symétrique
θ .... angle
x, y .... vecteurs
X, Y .... nombres réels
$\langle x,y\rangle$ .... produit scalaire
$\in$ .... appartenance
$\subset$ .... inclusion
$\cap$ .... intersection
$\oplus$ .... somme directe
$\circ$ .... composition

Index

Table des matières