

Après les nombres, pour décrire notre environnement (visible et invisible!) il a fallu conceptualiser la notion d'espace. Cette conceptualisation prend des formes diverses, en fonction des nécessités. Les espaces de la relativité ne sont pas identiques aux espaces de la mécanique quantique ni aux espaces de la mécanique de Newton. Pour construire ces espaces, la notion première est celle d'espace vectoriel.

Définitions: Un espace vectoriel est une ensemble E dans lequel les opérations suivantes ont un sens:

- (i) on peut additionner les éléments de E , $u \in E$, $v \in E$, alors $u + v \in E$.

- (ii) on peut multiplier tout élément de E par un nombre réel:
 $\lambda \in \mathbb{R}, v \in E$ alors $\lambda v \in E$.

E est alors un espace vectoriel réel. Si dans la propriété (ii) on remplace $\mathbb{R}$ par $\mathbb{C}$ on a alors un espace vectoriel complexe.

Une base est donc une famille à la fois libre et génératrice. La dimension N d'un espace vectoriel s'interprète comme le nombre de paramètres à déterminer pour caractériser un vecteur de cet espace. Par exemple l'espace des polynômes de degré au plus N est un espace vectoriel de dimension $N + 1$. On peut aussi interpréter N comme le nombre de degrés de liberté d'un système physique modélisé sur E . Par exemple une particule se déplaçant dans l'espace $\mathbb{R}^3$ a 3 degrés de liberté, $\mathbb{R}^3$ étant de dimension 3. Si l'on ajoute une contrainte on diminue le nombre de degrés de liberté (2 si M se déplace sur un plan, 1 si M se déplace sur une droite).

On a vu précédemment que $\mathbb{C}$ est un espace vectoriel réel. De même l'ensemble $M_2(\mathbb{R})$ des matrices réelles 2×2 .

Pour calculer dans un espace vectoriel on utilise souvent une **base**.

Un base dans un espace vectoriel est une ensemble de vecteurs $\{e_1, \dots, e_N\}$ de telle sorte que tout vecteur v de E s'exprime de

manière **unique** comme combinaison linéaire des e_n , $v = \sum_{n=1}^{n=N} a_n e_n$,

$a_n \in \mathbb{R}$. Le nombre N d'éléments de la base est la dimension de l'espace E (il ne dépend pas de la base).

Il existe 2 autres notions, plus faibles, concernant les familles de vecteurs $\{v_1, v_2, \dots, v_m\}$.

On dit que cette famille est **indépendante**, ou **libre**, si

$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ entraîne $\lambda_1 = \cdots = \lambda_m = 0$. Sinon on dit que cette famille est liée.

On dit que cette est une famille **génératrice** de E si tout $u \in E$ peut se représenter comme combinaison linéaire des v_k , c'est à dire $u = \lambda_1 v_1 + \dots + \lambda_m v_m$.

Pour calculer dans un espace vectoriel, il est commode de disposer également d'un **produit scalaire**. C'est indispensable si on veut avoir une notion de distance et de direction. Ainsi, à tout couple (u, v) de vecteurs de E , on associe un nombre réel noté $\langle u, v \rangle$, vérifiant les propriétés suivantes :

- ▶ $\langle u, v \rangle = \langle v, u \rangle$
- ▶ A u fixé, $v \mapsto \langle u, v \rangle$ est linéaire et à v fixé, $u \mapsto \langle u, v \rangle$ est linéaire.
- ▶ $\langle u, u \rangle \geq 0$
- ▶

$$\{\langle u, u \rangle = 0\} \Leftrightarrow \{u = 0\}$$

Deux vecteurs sont dits orthogonaux si $\langle u, v \rangle = 0$.

L'inégalité de Cauchy-Schwarz

$$|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle \langle v, v \rangle} \quad (1)$$

L'inégalité devient une égalité si et seulement si u et v sont colinéaires.

Preuve: Pour l'inégalité on développe $\langle u + \lambda v, u + \lambda v \rangle$.

$$\langle u + \lambda v, u + \lambda v \rangle = \langle u, u \rangle + 2\lambda \langle u, v \rangle + \lambda^2 \langle v, v \rangle$$

le membre de droite est un trinôme en λ qui ne change pas de signe, on a donc l'inégalité.

Le cas d'égalité correspond au cas d'une racine double λ_0 pour le trinôme. On a donc $u + \lambda_0 v = 0$.

Si on part d'une base de E , on a un produit scalaire "naturel" obtenu en utilisant les coordonnées des vecteurs u, v dans cette base :

$$\langle u, v \rangle = \sum_{j=1}^{j=N} u_j v_j$$

Pour ce produit scalaire la base est orthonormée, c'est à dire que l'on a

$$\|e_j\| = 1 \quad (2)$$

$$\langle e_j, e_k \rangle = 0, \text{ si } j \neq k \quad (3)$$

norme euclidienne

A un produit scalaire on associe une norme, dite euclidienne, par l'égalité

$$\|u\| = \sqrt{\langle u, u \rangle}$$

A cette norme est associée une distance dans l'espace physique (en mathématiques on dirait "espace affine"). Elle vérifie l'inégalité triangulaire: $\|u + v\| \leq \|u\| + \|v\|$.

Exercice: 1) Dédurre l'inégalité triangulaire de l'inégalité de Cauchy-Schwarz.

2) Démontrer l'identité du parallélogramme et l'interpréter géométriquement,

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

Inversement si on part d'un produit scalaire, on peut trouver une base de E qui soit orthonormée pour ce produit scalaire par le procédé d'orthonormalisation de Gram-Schmidt. On part d'une base quelconque de E , $\{a_1, \dots, a_N\}$. On commence par choisir $e_1 = \frac{a_1}{\|a_1\|}$. Puis on cherche $e_2 = \lambda_1 e_1 + \lambda_2 a_2$ ($\lambda_2 \neq 0$) de norme 1, tel que $\langle e_2, e_2 \rangle = 0$. Si $N = 2$ on s'arrête. Sinon on continue: on cherche ensuite $e_3 = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 a_3$ ($\lambda_3 \neq 0$), de norme 1 et orthogonal à e_1 et e_2 . On s'arrête à e_N . On a ainsi obtenu une base orthonormée de E .

Exercice: Soit $\mathcal{P}_N$ l'espace des polynômes de degré au plus N . On le munit du produit scalaire $\langle u, v \rangle = \int_{-1}^1 u(x)v(x)dx$. Déterminer une base orthonormée de $\mathcal{P}_N$ en partant de la base usuelle $\{1, x, x^2, \dots, x^N\}$. On obtient une suite P_n de polynômes de degré n , appelés polynômes de Legendre. Calculer P_n pour $n = 0, 1, 2$.

$$\forall v \in E, \quad v = \sum_{n=1}^{n=N} \langle v, e_n \rangle e_n \quad (4)$$

La 2ième ligne est une forme du théorème de Pythagore.

Les éléments de l'espace vectoriel E n'ont pas nécessairement une interprétation physique immédiate, ils servent essentiellement à décrire **les états du systèmes** et à faire des calculs, qui en fin de compte vont se traduire bien par des nombres (suivant l'affirmation de Maxwell), par exemple en décomposant l'état sur une base, si possible orthonormée. Cette idée est constamment mise en oeuvre en mécanique quantique, comme on le verra.

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

Exercice Déterminer le produit scalaire sur $M_2(\mathbb{R})$ pour lequel cette base est orthonormée.

$$m\ddot{x} + kx^2 = 0$$
$$\boxed{\ddot{x} + \omega^2 x = 0}, \quad \boxed{\omega = \sqrt{\frac{k}{m}}} \quad (6)$$

k étant la constante d'élasticité du ressort. On voit facilement que l'ensemble des solutions de (6) est un espace vectoriel de dimension 2 (bien que le problème physique soit unidimensionnel).

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

◀ ◻ ▶ ◻ ▶ ◻ ▶ ◻ ▶ ◻ ▶ ↺ 🔍 ↻

A^{-1} est un opérateur de F dans E . Il est commode de représenter un opérateur A par une matrice $A_{k,\ell}$.

$$\langle Au, v \rangle = \sum_{k, \ell} A_{k, \ell} u_k v_\ell \quad (12)$$

$$\boxed{BA(u) = B(A(u))}, \quad \forall u \in E$$

$E = F = G = \mathbb{R}^2$. A est la symétrie par rapport à la première bissectrice, B la rotation d'angle $\pi/2$. AB est la symétrie par rapport à l'axe Ox , BA est la symétrie par rapport Oy .

Cela peut-être fastidieux. Il existe des techniques pour simplifier ces calculs, comme la méthode du pivot de Gauss, qui permet de se ramener au cas où la matrice de A est triangulaire ($A_{j,k} = 0$ si $j > k$).

En pratique on dispose les nombres $B_{j,\ell}$, $A_{\ell,k}$ sous forme de tableaux rectangulaires, avec lignes et colonnes, le premier indice indique la ligne, le deuxième la colonne. Les vecteurs sont représentés par des matrices colonnes.

et discuter les conditions sur θ et α .

2) Un projecteur est toujours la projection orthogonal sur un sous-espace F , F étant l'image de Π . Si dans F on choisit une base orthonormée $\{e_1, \dots, e_p\}$ (p étant la dimension de F) alors on

$$\Pi v = \sum_{k=1}^{k=p} \langle v, e_k \rangle e_k$$

$v - \Pi v$ est orthogonal à F .

Inversement si on part d'un sous-espace F de E l'égalité précédente détermine le projecteur orthogonal sur F que l'on note Π_F .

Pour changer, donner une application aux statistiques et au calcul des corrélations linéaires. Lorsque l'on récolte des valeurs observées, $x_i, y_i, 1 \leq i \leq n$, représentées par un nuage de points $\{(x_i, y_i), 1 \leq i \leq n\}$ dans le plan, on cherche la droite approximant au mieux ce nuage de points en calculant ses coefficients $a, b \in \mathbb{R}$ tels que la quantité suivante

$$e(a, b) = \sum_{1 \leq i \leq n} |y_i - (ax_i + b)|^2$$

soit minimale. Cette droite porte le nom de droite des moindres carrés ou droite de régression linéaire de y en x .

Interprétons ce problème géométriquement dans l'espace euclidien $E = \mathbb{R}^n$ muni du produit scalaire

$$\langle x, y \rangle = \frac{1}{n} \sum_{i=1}^{i=n} x_i y_i, \quad x = (x_1, \dots, x_n).$$

Le facteur $\frac{1}{n}$ permet de faire le lien avec les probabilités (moyenne, variance, écart type)

3) Interprétation géométrique

Ce qui suit est évident en géométrie plane mais l'est sans doute moins en dimension quelconque, alors que les applications sont nombreuses.

Théorème: Pour tout $u \in E$, Πu est l'unique vecteur de F réalisant la meilleure approximation de u par un élément de $v \in F$ avec une erreur égale à $\|u\|^2 - \|\Pi u\|^2$. En résumé

$$\|u - v\| \geq \|u - \Pi u\|, \quad \forall v \in F \quad (21)$$

et on on l'égalité si et seulement si $v = \Pi u$. On a alors :

$$\|u\|^2 = \|\Pi u\|^2 + \|u - \Pi u\|^2 \quad (22)$$

x représente un échantillon, de moyenne m_x , d'écart type s_x . Si p est un nombre, on désignera également par p l'échantillon dont toutes les valeurs sont égales à p .

$$m_x = \frac{1}{n} \sum_{1 \leq i \leq n} x_i, \quad (23)$$

$$s_x = \left[\frac{1}{n} \sum_{1 \leq i \leq n} (x_i - m_x)^2 \right]^{1/2} = \|x - m_x\|. \quad (24)$$

En particulier x est constant si et seulement si $s_x = 0$.

Noter que $s_x^2 = \|y\|^2 - m_x^2$ (formule de Koenig).

$$c_{xy} = \frac{1}{n} \sum_{1 \leq i \leq n} (x_i - m_x)(y_i - m_y) = \langle x - m_x, y - m_y \rangle \quad (25)$$

c_{xy} est la covariance observée et r_{xy} est le coefficient de corrélation observé sur l'échantillon.

$$y = ax + b, \quad a = r_{xy} \frac{s_y}{s_x}, \quad b = m_y - am_x \quad (28)$$
$$\delta(a, b)^2 = \|y\|^2 - \|\Pi_F y\|^2 \quad (29)$$
$$\|\Pi_F y\|^2 = \langle y, 1 \rangle^2 + \left\langle y, \frac{x - m_x}{s_x} \right\rangle^2 \quad (30)$$

on obtient :

$$\Pi_F y = \langle y, 1 \rangle 1 + \left\langle y, \frac{x - m_x}{s_x} \right\rangle \frac{x - m_x}{s_x} \quad (27)$$

La droite Δ_{xy} porte aussi le nom de droite des moindres carrés.
On considère que la série y_i a une corrélation linéaire forte avec la série des x_i si $|r_{xy}| \geq 0,87$

On part d'un ensemble à n éléments, $\Omega_n = \{1, 2, \dots, n\}$. Une permutation est une application σ de Ω_n telle que si $j \neq k$ alors $\sigma j \neq \sigma k$. Autrement dit c'est une autre façon de réordonner l'ensemble Ω_n . Une transposition est une permutation qui échange 2 éléments et laisse les autres inchangés.

On en déduit que pour tout opérateur A de E dans E il existe un nombre réel noté $\det A$ tel que pour $f \in \mathcal{A}_n$ on a

$$f(Au_1, Au_2, \dots, Au_n) = \det(A)f(u_1, u_2, \dots, u_n) \quad (33)$$

Ce nombre $\det(A)$ s'appelle le déterminant de A . A quoi sert-il?

Soit E un espace vectoriel, f une application n -linéaire de E^n dans $\mathbb{R}$ est dite alternée si pour toute permutation σ on a

$$f(u_{\sigma(1)}, u_{\sigma(2)}, \dots, u_{\sigma(n)}) = (-1)^{s_\sigma} f(u_1, u_2, \dots, u_n).$$

3) Inversement si $\det(A) \neq 0$ alors A est inversible.

Preuve de la dernière propriété. Si A n'est pas inversible, il existe une base $\{b_j\}$ de E telle que $Ab_1 = 0$. Or on a $f_0(b_1, \dots, b_n) \neq 0$ et $f(Ab_1, \dots, Ab_n) = 0$. Donc $\det(A) = 0$. Ce qui démontre 3). Si la matrice de A dans une base $\{b_j\}$ de E est $\{A_{j,k}\}$, le déterminant se calcule par la formule

$$\det(A) = \sum_{\sigma} (-1)^{s_{\sigma}} A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n} \quad (34)$$

σ décrivant l'ensemble des permutations de Ω_n . On en déduit que A et ${}^t A$ (transposé de A) ont même déterminant:

$$\det({}^t A) = \det(A).$$

Interprétation géométrique du déterminant

On considère le plan euclidien, muni d'un repère orthonormé, et deux vecteurs u, v . L'aire du parallélogramme construit sur ces vecteurs est égal à $|\det(u, v)|$. Cette formule s'étend pour 3 vecteurs de l'espace (l'aire devient un volume) et pour n vecteurs de $\mathbb{R}^n$, on obtient un hypervolume. Ainsi, si $u_1, u_2, \dots, u_n$ sont n vecteurs indépendants de $\mathbb{R}^n$, $|\det(u_1, u_2, \dots, u_n)|$ est le volume du parallélogramme n -dimensionnel $P = \{t_1 u_1 + t_2 u_2 + \dots + t_n u_n, t_j \in [0, 1]\}$ (faire un dessin si $n = 2, 3$).

Exercice.

Donner l'expression du déterminant d'une matrice 2×2 puis d'une matrice 3×3 .

Maintenant fixons une base $\{b_j\}$ de E considérons n vecteurs de E , $u_1, u_2, \dots, u_n$. On désigne alors par $\det(u_1, u_2, \dots, u_n)$ le déterminant de l'opérateur A défini par $Ab_j = u_j$, $1 \leq j \leq n$. On déduit alors de ce qui précède : le système $\{b_1, \dots, b_n\}$ est une base si et seulement si $\det(u_1, u_2, \dots, u_n) \neq 0$.

Le déterminant sert donc comme critère pour l'inversibilité d'un opérateur ou pour caractériser les bases.

Voici une méthode pour démontrer le résultat précédent.

On désigne par A l'opérateur défini par $Ae_j = u_j$, $u_1, u_2, \dots, u_n$ étant n vecteurs indépendants de $\mathbb{R}^n$ et e_j la base canonique de $\mathbb{R}^n$. On verra dans un prochain cours que $A = U|A|$, U étant une isométrie et $|A|$ est positif. Or $U^{-1} = {}^t U$ entraîne que $|\det(U)| = 1$. On cherche donc le volume de l'hyperparallélogramme défini par $P = \{t_1 u_1 + t_2 u_2 + \dots + t_n u_n, t_j \in [0, 1]\}$ qui est donc l'image par A de l'hypercube $[0, 1]^n$. On commence par se ramener au cas où U est l'identité, puis au cas où la matrice $|A|$ est diagonale, $Ae_j = \lambda_j e_j$. Alors $\det(A) = \lambda_1 \cdots \lambda_n$ représente bien le volume de $[0, \lambda_1] \times \dots \times [0, \lambda_n]$. On en déduit le résultat souhaité.