

ESPACES DE HILBERT PROBABILITES¹

J. Mellac

Juillet 1998

1. Merci à E. L'Haridon qui a proposé nombre d'exercices figurant dans ce polycopié

Première partie

ESPACES DE HILBERT

Chapitre 1

ESPACES METRIQUES

1.1 Définitions

Un espace métrique est un couple (E, d) où E est un ensemble et d une application de $E \times E$ dans \mathbb{R}^+ vérifiant :

- $d(x, y) = 0 \Leftrightarrow x = y$.
- $\forall x, y \in E, d(x, y) = d(y, x)$.
- $\forall x, y, z \in E, d(x, z) \leq d(x, y) + d(y, z)$. (Inégalité triangulaire).

Si $x_0 \in E$ et $r \in \mathbb{R}^{+*}$ on appelle boule ouverte de centre x_0 et de rayon r , l'ensemble noté $B(x_0, r) = \{x \in E \mid d(x_0, x) < r\}$.

On appelle ouvert de E toute réunion de boules ouvertes. Un fermé de E est une partie de E dont le complémentaire est un ouvert.

Remarque : Une partie A de E est ouverte si et seulement si tout $x_0 \in A$ est centre d'une boule ouverte incluse dans A .

1.2 Exemples d'espaces métriques

Exemple 1 :

$E = \mathbb{R}, d(x, y) = |x - y|$. Une boule ouverte est dans ce cas un intervalle ouvert. Les trois propriétés d'une distance son aisément vérifiées.

Exemple 2 :

$E = \mathbb{R}^2, d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ Pour démontrer l'inégalité triangulaire il suffit d'utiliser l'inégalité

$$a_1 b_1 + a_2 b_2 \leq \sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2)} \text{ avec}$$
$$a_i = x_i - y_i, b_i = y_i - z_i, i = 1, 2.$$

Les boules ouvertes de cet espace sont les disques ouverts. On peut définir sur cet ensemble d'autres métriques.

$d_1(x, y) = \max(|x_1 - y_1|, |x_2 - y_2|)$ dont les boules ouvertes sont formés de l'intérieur des carrés de côtés parallèles aux axes.

$d_2(x, y) = |x_1 - y_1| + |x_2 - y_2|$ dont les boules ouvertes sont formés de l'intérieur des carrés dont les diagonales sont parallèles aux axes.

1.3 Propriétés.

Théorème 1.1 *Un espace métrique E est séparé, ce qui signifie que si x et y sont des points distincts de E , il existe des boules ouvertes contenant x et y respectivement et d'intersection vide.*

Démonstration

La distance $r = d(x, y) > 0$ car $x \neq y$. Les boules $B_x = B(x, r/4)$ et $B_y = B(y, r/4)$ sont telles que $B_x \cap B_y = \emptyset$ et elles contiennent respectivement x et y .

Définition 1.1 *Un point $\lambda \in E$ est un point d'accumulation d'une partie $A \subset E$ si toute boule ouverte de centre λ contient au moins un point de A différent de λ .*

Exemple

Si $E = \mathbb{R}$ et si $d(x, y) = |x - y|$, 0 est point d'accumulation de $A = \{1, 1/2, 1/3, \dots\}$ mais aussi de $A' = A \cup \{0\}$. Ce qui montre qu'un point d'accumulation peut appartenir ou non à l'ensemble.

Théorème 1.2 *Un ensemble $A \subset E$ est fermé si et seulement si il contient tous ses points d'accumulation.*

Démonstration

Dans le cas où $x \in E - A$ qui est un ouvert, il existe une boule $B(x, \epsilon)$ incluse dans $E - A$, x ne peut donc être un point d'accumulation de A .

Réciproquement, si A contient tous ses points d'accumulation et si $x \in E - A$, il existe une boule ouverte de centre x ne contenant aucun point de A . $E - A$ est donc ouvert et A est fermé.

Définition 1.2 *On appelle fermeture ou adhérence de $A \subset E$ l'ensemble $\bar{A} = A \cup A_a$ où A_a est l'ensemble des points d'accumulation de A . C'est donc le plus petit fermé contenant A .*

Remarque : A est fermé $\Leftrightarrow A = \bar{A}$

1.4 Espaces métriques complets

Définition 1.3 — *Une suite (x_n) de points de l'espace métrique converge vers $x \in E$ si*

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0.$$

— *Une suite (x_n) est de Cauchy si*

$$\lim_{n, m \rightarrow \infty} d(x_n, x_m) = 0.$$

— *un espace métrique (E, d) est complet si toute suite de Cauchy de E converge.*

Remarques

1) Une partie $A \subset E$ est fermée si et seulement si elle contient les limites de ses suites convergentes.

2) $(\mathbb{R}, |\cdot|)$ est complet mais \mathbb{Q} muni de la même distance ne l'est pas. La première affirmation provient de la construction même de \mathbb{R} . Pour montrer la deuxième notons x_n la représentation décimale par défaut de $\sqrt{2}$ contenant n décimales, c'est une suite de \mathbb{Q} , mais la limite $\sqrt{2} \notin \mathbb{Q}$ (c'est une suite de Cauchy convergent dans \mathbb{R} mais pas dans \mathbb{Q}).

1.5 Espaces vectoriels normés

Dans la suite le corps de base des espaces vectoriel considérés est noté K . Il s'agira de \mathbb{R} ou \mathbb{C} .

Définition 1.4 *Un espace vectoriel E est normé s'il est muni d'une application*

$$\|\cdot\| : E \rightarrow \mathbb{R}^+$$

vérifiant

- $\|x\| = 0 \Leftrightarrow x = 0$.
- $\forall \lambda \in K, \forall x \in E \|\lambda x\| = |\lambda| \|x\|$.
- $\forall x, y \in E \|x + y\| \leq \|x\| + \|y\|$.

Il en découle immédiatement que $d(x, y) = \|x - y\|$ définit une distance sur E .

Exemples

a) $E = \mathbb{R}^n$, $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ est un \mathbb{R} -espace vectoriel.

b) $E = C([a, b])$ (fonctions réelles ou complexes continues sur l'intervalle $[a, b]$)
 $\|f\| = \max_{a \leq x \leq b} |f(x)|$.

Une suite (f_n) converge alors vers f pour la distance d associée à $\|\cdot\|$ si et seulement si (f_n) converge uniformément vers f .

c) $E = \mathbb{C}^n$, $\|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}$ est un \mathbb{C} -espace vectoriel normé.

d) $E = \left\{ (x_n)_{n \in \mathbb{N}} \mid \sum_{n \in \mathbb{N}} |x_n|^2 < +\infty \right\}$ où $x_n \in K$ pour tout n est un K -espace vectoriel normé pour la norme

$$N_2(x) = \left(\sum_{n \in \mathbb{N}} |x_n|^2 \right)^{1/2}.$$

Remarque.

En notant μ la mesure de décompte sur \mathbb{N} et $x = (x_n)$ on a

$$N_2(x) = \left(\int_{\mathbb{N}} |x|^2 d\mu \right)^{1/2}$$

cet espace est donc un espace de Riesz particulier étudié dans le cours d'intégration, il est noté $l^2(K)$.

1.6 Espaces vectoriels normés complets

Définition 1.5 On appelle espace de Banach tout espace vectoriel normé complet.

Exemples.

a) \mathbb{R}^n et \mathbb{C}^n sont complets pour la norme précisée précédemment. Ceci provient du fait que \mathbb{R} et \mathbb{C} sont complets et qu'une suite de terme général $x_k = (x_k^1, x_k^2, \dots, x_k^n)$ converge vers $x = (x_1, x_2, \dots, x_n)$ dans K^n si et seulement si pour tout i x_k^i converge vers x_i .

b) L'espace $l^2(K)$ est complet. Ceci a été démontré dans le cours d'intégration (tous les espaces L^p sont complets). Cet exemple est très important car cet espace est isomorphe à tous les espaces de Hilbert que l'on traitera dans ce cours.

c) L'espace (X, \mathcal{B}, μ) étant un espace mesuré (\mathcal{B} est une tribu sur X , μ une mesure sur cet espace) on définit

$$L_\mu^p(X) = \left\{ f : X \rightarrow K \mid \int_X |f|^p d\mu < +\infty \right\}.$$

Tous ces espaces correspondants à $p \in \mathbb{R}$, $p \geq 1$ sont des espaces de Banach (voir cours d'intégration). L'exemple précédent est un cas particulier où $X = \mathbb{N}$.

1.7 Applications linéaires continues

Rappel.

— Les espaces (E, d) et (F, δ) étant métriques et f une application de E vers F , on dit que f est continue en $a \in E$ si

$$\forall \epsilon > 0, \exists \eta > 0, \forall x \in E, d(x, a) < \eta \Rightarrow \delta(f(x), f(a)) < \epsilon.$$

f est continue sur E si elle est continue en tout point de E . Il revient au même de dire que l'image réciproque de tout ouvert est un ouvert (ou l'image réciproque de tout fermé est un fermé).

— On dit que f est lipschitzienne s'il existe $k \in \mathbb{R}^+$ tel que

$$\forall x, y \in E, \delta(f(x), f(y)) \leq kd(x, y).$$

Une application lipschitzienne est continue sur E .

— Une application f est continue en $a \in A$ si et seulement si pour toute suite (x_n) convergeant vers a dans (E, d) , la suite $(f(x_n))$ converge vers $f(a)$ dans (F, δ) .

Dans la suite E et F seront deux espaces vectoriels normés, munis des distances associées à ces normes, qui seront toutes deux notées $\|\cdot\|$.

Théorème 1.3 Soit u une application linéaire de E dans F , les trois propriétés suivantes sont équivalentes

- 1) u est continue en o .
- 2) u est bornée sur toute partie bornée de E .

3) u est lipschitzienne.

Démonstration

(1) \Rightarrow (2) : Si u est continue en 0, il existe $r > 0$ tel que $\|x\| \leq r \Rightarrow \|u(x)\| \leq 1$. u est donc bornée sur la boule $B(0, r)$ donc par homothétie sur toute boule de centre 0 et donc sur toute partie bornée de E .

(2) \Rightarrow (3) : Si (2) est vérifiée on a $\|x\| = 1 \Rightarrow \|u(x)\| \leq M$.

Si $x \neq 0$, $\|u(x/\|x\|)\| \leq M \Rightarrow \|u(x)\| \leq M\|x\|$ et ceci est vérifié si $x = 0$. On a alors $\|u(x) - u(y)\| = \|u(x - y)\| \leq M\|x - y\|$ et u est bien lipschitzienne.

(3) \Rightarrow (1) est immédiat.

Remarque. Soit u une application linéaire de E vers F , la plus petite constante de Lipschitz associée à E est le nombre

$$\|u\| = \sup_{x \neq 0} \frac{\|u(x)\|}{\|x\|} = \sup_{\|x\| \leq 1} \|u(x)\| = \sup_{\|x\|=1} \|u(x)\|.$$

Ce nombre est appelé norme de u . On a donc

$\|u(x)\| \leq \|u\| \|x\|$, le signe $\|\cdot\|$ ayant trois significations différentes dans cette inégalité.

Théoreme 1.4 *L'espace vectoriel $\mathcal{L}(E, F)$ des applications linéaires continues de E dans F est normé par $u \rightarrow \|u\|$. Si F est complet $\mathcal{L}(E, F)$ est complet pour la norme qui vient d'être définie.*

Démonstration Soient $u, v \in \mathcal{L}(E, F)$

$$\|u + v\| = \sup_{\|x\|=1} \|u(x) + v(x)\| \leq \sup_{\|x\|=1} \|u(x)\| + \sup_{\|x\|=1} \|v(x)\| = \|u\| + \|v\|$$

Pour $\lambda \in K$, on a immédiatement $\|\lambda u\| = |\lambda| \|u\|$

$\|u\| = 0 \Rightarrow \forall x \in E \ u(x) = 0$ c'est à dire $u = 0$.

L'espace vectoriel $\mathcal{L}(E, F)$ est donc bien normé à l'aide de l'application définie plus haut.

Supposons F complet et considérons une suite de Cauchy (u_n) dans $\mathcal{L}(E, F)$.

L'inégalité (1) $\|u_p(x) - u_q(x)\| \leq \|u_p - u_q\| \|x\|$ entraîne que $(u_n(x))$ est une suite de Cauchy dans F pour tout $x \in E$, elle converge donc vers un élément de F que l'on notera $u(x)$.

$$u(\lambda x + \mu y) = \lim_{n \rightarrow \infty} u_n(\lambda x + \mu y) = \lim_{n \rightarrow \infty} (\lambda u_n(x) + \mu u_n(y)) = \lambda u(x) + \mu u(y)$$

u est donc bien linéaire, montrons qu'elle est continue.

Soit $\epsilon > 0$ et choisissons $n \in \mathbb{N}$ tel que $p, q \geq n \Rightarrow \|u_p - u_q\| \leq \epsilon$

en faisant tendre q vers l'infini (1) donne $\|u_p(x) - u(x)\| \leq \epsilon \|x\|$, $u_p - u$ est donc Lipschitzienne, donc continue et $u = (u - u_p) + u_p$ est continue.

$\|u_p - u_q\| \leq \epsilon \Rightarrow \|u_p - u\| \leq \epsilon$ ($q \rightarrow \infty$), ce qui exprime que $\lim_{p \rightarrow \infty} u_p = u$ dans $\mathcal{L}(E, F)$ qui est donc complet.

Remarque.

a) Si $E = F$, $\mathcal{L}(E, F)$ est noté $\mathcal{L}(E)$ et on a $\|u \circ v\| \leq \|u\| \|v\|$.

b) Si $F = \mathbb{R}$ ou \mathbb{C} , $\mathcal{L}(E, F)$ est appelé dual topologique de E et noté E^* . C'est un espace vectoriel normé complet.

Définition 1.6 Les espaces vectoriels E et F étant normés, l'application linéaire $u : E \rightarrow F$ est un isomorphisme d'espaces normés si elle est bijective et bicontinue, c'est à dire que u et u^{-1} sont continues.

Théoreme 1.5 L'application linéaire bijective u est un isomorphisme si et seulement il existe deux réels strictement positifs A et B tels que

$$\forall x \in E \quad A\|x\| \leq \|u(x)\| \leq B\|x\|.$$

La première inégalité est équivalente à la continuité de u^{-1} et la deuxième à la continuité de u .

Remarque.

L'espace vectoriel E étant muni des deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$, elles donnent les mêmes ouverts (on dit que les topologies sont identiques) si et seulement si l'application identité $Id : (E, \|\cdot\|_1) \rightarrow (E, \|\cdot\|_2)$ est un isomorphisme l'image réciproque d'un ouvert (par Id ou par $Id^{-1} = Id$) étant alors un ouvert.

Corollaire 1.5.1 $Id : (E, \|\cdot\|_1) \rightarrow (E, \|\cdot\|_2)$ est un isomorphisme d'espaces vectoriels normés si et seulement si les normes sont équivalentes c'est à dire si elles vérifient

$$\forall x \in E, \quad A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1$$

Ce résultat est une conséquence immédiate du résultat précédent.

Théoreme 1.6 Sur un espace vectoriel de dimension finie toutes les normes sont équivalentes, ce qui entraîne qu'un tel espace n'a qu'une seule topologie.

Démontrons d'abord un lemme préliminaire.

lemme 1.6.1 L'espace vectoriel K^n étant muni de la norme définie par $p(x_1, x_2, \dots, x_n) = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$, toute autre norme sur cet espace est continue pour p .

Démonstration

Soit $(e_i)_{1 \leq i \leq n}$ la base canonique de K^n et $x \in K^n$, $x = \sum_{i=1}^n x_i e_i$ Si $\|\cdot\|$ est une norme quelconque sur K^n

$$\|x\| = \left\| \sum_{i=1}^n x_i e_i \right\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq A p(x) \text{ où } A = \left(\sum_{i=1}^n \|e_i\|^2 \right)^{1/2}$$

$\|\cdot\|$ est donc continue en 0. L'inégalité $|\|x\| - \|a\|| \leq \|x - a\|$ entraîne la continuité en tout point a .

Démontrons maintenant le théorème. Appelons q et r deux normes sur K^n , elles sont continues par rapport à p . Sur $S = \{x \in K^n | p(x) = 1\}$, sphère unité de K^n , q et r ne s'annulent pas, S étant compacte (fermée et bornée) r/q a un maximum et un minimum sur S

$$\forall x \in S, \quad \exists(A, B) \in \mathbb{R}^2 \quad A \leq \frac{r(x)}{q(x)} \leq B \Leftrightarrow Aq(x) \leq r(x) \leq Bq(x)$$

Ces inégalités se généralisent à K^n , ce qui montre que sur K^n toutes les normes sont équivalentes. Considérons un espace vectoriel E quelconque de dimension n sur K et appelons f un isomorphisme algébrique de K^n sur E (toute application linéaire envoyant la base canonique de K^n sur une base de E est un tel isomorphisme algébrique). Dans le cas où q et r sont deux normes sur E , $q \circ f$ et $r \circ f$ sont des normes sur K^n et à l'aide de ce qui précède

$$\forall x \in K^n, Aq(f(x)) \leq r(f(x)) \leq Bq(f(x))$$

or $f(x)$ est un vecteur quelconque de E , ce qui montre que q et r sont équivalentes. Remarque. Ceci est faux en dimension infinie.

Soit $E = C([0, 1], \mathbb{R})$, $\|f\|_u = \sup_{t \in [0, 1]} |f(t)|$, $\|f\|_1 = \int_0^1 |f(t)| dt$. La suite de fonctions définie par $f_n(t) = t^n$ vérifie $\|f_n\|_u = 1$ et $\|f_n\|_1 = 1/(n+1)$. Il est facile de montrer que $\|f\|_1 \leq \|f\|_u$ pour tout $f \in E$, mais il n'est pas possible de trouver une constante k telle que $\forall f \in E, \|f\|_u \leq k\|f\|_1$ car $\frac{\|f_n\|_u}{\|f_n\|_1} = n+1$ et ce quotient n'est pas borné.

Théoreme 1.7 *Tout espace normé de dimension n est isomorphe à K^n , il est donc complet.*

Démonstration

Soit $f : K^n \rightarrow E$ un isomorphisme algébrique. Si q est la norme sur E , $r = q \circ f$ est une norme sur K^n , elle est donc équivalente à la norme euclidienne p

$$\forall x \in K^n, Ap(x) \leq q(f(x)) \leq Bp(x)$$

f est donc non seulement un isomorphisme algébrique mais aussi un isomorphisme d'espace normé.

Théoreme 1.8 *Toute application linéaire f d'un espace vectoriel E de dimension finie n vers un espace vectoriel quelconque F est continue.*

Démonstration

Soit $f : K^n \rightarrow E$ un isomorphisme et $g : E \rightarrow F$ une application linéaire. On pose $h = g \circ f$, $h(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i h(e_i)$

$$\|h(x_1, x_2, \dots, x_n)\| \leq \sum_{i=1}^n |x_i| \|h(e_i)\| \leq \left(\sum_{i=1}^n \|h(e_i)\|^2 \right)^{1/2} p(x_1, x_2, \dots, x_n)$$

ce qui entraîne la continuité de h et donc également la continuité de $g = h \circ f^{-1}$.

Chapitre 2

ESPACES DE HILBERT

2.1 Formes Hermitiennes

Dans toute la suite on prendra $K = \mathbb{C}$.

Définition 2.1 — E et F étant deux espaces vectoriels complexes, une application $u : E \rightarrow F$ est dite semi-linéaire si

$$\forall x, y \in E \quad u(x + y) = u(x) + u(y), \quad \forall \lambda \in \mathbb{C}, \quad \forall x \in E \quad u(\lambda x) = \bar{\lambda}u(x).$$

- E, F, G étant trois espaces vectoriels complexes, une application $u : E \times F \rightarrow G$ est dite sesquilinéaire si $x \rightarrow u(x, y)$ est linéaire et $y \rightarrow u(x, y)$ est semi-linéaire. Si $G = \mathbb{C}$ u est une forme sesquilinéaire.
- Une forme sesquilinéaire sur l'espace vectoriel complexe $E \times E$ est hermitienne si $\forall x, y \in E \quad f(y, x) = \overline{f(x, y)}$. Dans le cas $K = \mathbb{R}$ l'analogue des formes hermitiennes sont les formes bilinéaires symétriques.

Exemples

1) Sur $\mathbb{C}^n \times \mathbb{C}^n$ la forme hermitienne la plus générale s'exprime à l'aide des coordonnées x_i et y_i par

$$f(x, y) = \sum_{i,j=1}^n \alpha_{ij} x_i \bar{y}_j, \quad \alpha_{ij} = f(e_i, e_j) \in \mathbb{C}.$$

f est hermitienne si et seulement si $\alpha_{ji} = \bar{\alpha}_{ij}$ $i, j = 1, 2, \dots, n$. Dans ce cas on dit que la matrice $A = (\alpha_{ij})_{i,j}$ est hermitienne.

2) On peut munir l'espace

$$l^2 = \{(x_n)_{n \in \mathbb{N}} \mid \forall n, x_n \in \mathbb{C}, \sum_{n=0}^{\infty} |x_n|^2 < +\infty\}$$

de la forme hermitienne définie par $f((x_n), (y_n)) = \sum_{i=0}^{\infty} x_i \bar{y}_i$. Cette série est bien définie sur l^2 car $|x_i \bar{y}_i| \leq 1/2(|x_i|^2 + |y_i|^2)$. Il est alors facile de vérifier que f est une forme hermitienne.

3) Sur $E = C([0, 1], \mathbb{C})$ la forme définie par

$$f(x, y) = \int_0^1 x(t) \overline{y(t)} dt$$

est hermitienne.

4) $E = L^2_\mu(X)$ où (X, \mathcal{A}, μ) est un espace mesuré et E l'espace vectoriel constitué des classes d'équivalence des fonctions mesurables de carré intégrable, à valeurs dans \mathbb{C} . Soit $f, g \in E$, l'application $\phi(f, g) = \int_X f \cdot \bar{g} d\mu$ est une forme hermitienne sur E , l'intégrale étant bien définie d'après l'inégalité de Hölder.

Théoreme 2.1 Soit f une forme sesquilinéaire sur E , on a

$$(1) f(x, y) = \frac{1}{4} (f(x + y, x + y) - f(x - y, x - y) + if(x + iy, x + iy) - f(x - iy, x - iy))$$

Démonstration

$$f(x + y, x + y) = f(x, x) + f(y, y) + f(x, y) + f(y, x)$$

$$f(x + iy, x + iy) = f(x, x) + f(y, y) - if(x, y) + if(y, x)$$

En remplaçant y par $-y$ dans ces expressions et en effectuant les opérations précisées on obtient le résultat.

f est donc déterminée par les valeurs prises sur la diagonale.

Corollaire 2.1.1 Une forme sesquilinéaire est hermitienne si et seulement si elle prend des valeurs réelles sur la diagonale de $E \times E$.

Si f est hermitienne on a $f(x, x) = \overline{f(x, x)} \in \mathbb{R}$.

Réciproquement la formule (1) s'écrit $f(x, y) = 1/4(A - B + iC - iD)$ où $A, B, C, D \in \mathbb{R}$, donc $f(y, x) = \overline{f(x, y)}$.

Définition 2.2 Une forme hermitienne est définie positive si $\forall x \in E, x \neq 0$ $f(x, x) > 0$.

Théoreme 2.2 Si f est une forme hermitienne définie positive sur E on a

$$\forall (x, y) \in E \times E, |f(x, y)|^2 \leq f(x, x) f(y, y)$$

, c'est l'inégalité de Cauchy-Schwarz.

Démonstration

Si $x = y = 0$, l'inégalité est immédiate. Supposons par exemple $f(x, y) = a > 0$ ($x \neq 0$). Posons $f(x, y) = b$, $f(y, y) = c$.

$$\forall \lambda \in \mathbb{C}, f(\lambda x + y, \lambda x + y) = a\lambda\bar{\lambda} + b\lambda + \bar{b}\bar{\lambda} + c = a \left[|\lambda + \bar{b}/a|^2 + \frac{ac - |b|^2}{a^2} \right] \geq 0$$

La valeur $\lambda = -\bar{b}/a$ entraîne $ac - |b|^2 \geq 0$, ce qui est l'inégalité à démontrer.

Théoreme 2.3 L'inégalité ci-dessus devient une égalité si et seulement si x et y sont linéairement dépendants.

Démonstration

Si $\lambda x + y = 0$, $f(\lambda x + y, \lambda x + y) = 0 \Rightarrow |b|^2 = ac$.

Réciproquement si $|b|^2 = ac$ dans le cas $a = 0$ on a $x = 0$ et si $a > 0$

$f(\lambda x + y, \lambda x + y) = a|\lambda + \bar{b}/a|^2$ ce qui entraîne $\lambda x + y = 0$ si $\lambda = -\bar{b}/a$.

Corollaire 2.3.1 Si f est une forme hermitienne définie positive sur $E \times E$, l'application $p : x \rightarrow \sqrt{f(x, x)}$ est une norme sur E .

Démonstration

- $p(x) = 0 \Leftrightarrow x = 0$.
- $p(\lambda x) = |\lambda|p(x)$.
- $p^2(x + y) = f(x + y, x + y) = f(x, x) + f(y, y) + 2\Re(f(x, y))$
 $\geq p^2(x) + p^2(y) + 2p(x)p(y) = (p(x) + p(y))^2$
 d'où $p(x + y) \leq p(x) + p(y)$.

2.2 Espaces de Hilbert

Définition 2.3 Un espace de Hilbert (ou espace Hilbertien) est un espace vectoriel complexe muni d'une forme hermitienne définie positive et complet pour la norme associée à cette forme.

Cette forme définie sur E (on dit aussi produit scalaire) sera notée $(x|y)$ et la norme associée $\|x\|$.

Avec ces notations on obtient

$$|(x|y)| \leq \|x\| \|y\| \text{ (Inégalité de Cauchy-Schwarz) et}$$

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

Exemples. 1) l^2 est un espace de Hilbert associé au produit scalaire

$$(x|y) = \sum_{i=0}^{\infty} x_i \bar{y}_i$$

où $x = (x_i)_{i \in \mathbb{N}}$, $y = (y_i)_{i \in \mathbb{N}} \in l^2$.

2) L'espace $C([0, 1], \mathbb{C})$ peut être muni de la forme hermitienne définie positive

$$(x|y) = \int_0^1 x(t) \overline{y(t)} dt$$

mais il n'est pas complet pour la norme associée.

Posons $x_n(t) = \sin(\pi/t)$ pour $1/n \leq t \leq 1$, $x_n(t) = 0$ pour $0 \leq t \leq 1/n$

$$\|x_n - x_m\|^2 = \int_{1/n}^{1/m} \sin^2(\pi/t) dt \quad (n > m), \quad \lim_{n, m \rightarrow \infty} \|x_n - x_m\|^2 = 0$$

Ceci montre que (x_n) est une suite de Cauchy dans cet espace. Si elle admettait une limite $x \in E$ on aurait pour $a > 0$

$$\lim_{n \rightarrow \infty} \int_a^1 |x(t) - x_n(t)|^2 dt = \int_a^1 |x(t) - \sin(\pi/t)|^2 dt = 0$$

ce qui entrainerait $x(t) = \sin(\pi/t)$ pour $0 < a \leq t \leq 1$ car $x(t) - \sin(\pi/t)$ est continue, d'où $x(t) = \sin(\pi/t)$ pour $0 < t \leq 1$ or une telle fonction ne peut être continue en 0 quelle que soit la valeur prise par $x(0)$.

3) $E = L^2_\mu(X)$ muni du produit scalaire $(f|g) = \int_X f \bar{g} d\mu$ est un espace de Hilbert. La norme $\|\cdot\|_2$ en fait un espace vectoriel normé complet (cours d'intégration).

Théorème 2.4 Dans un espace de Hilbert on a les relations suivantes

- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re(x|y)$
- $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$. Identité du parallélogramme.
- $\|x\|^2 + \|y\|^2 = 2\|\frac{x+y}{2}\|^2 + \frac{1}{2}\|x - y\|^2$. Théorème de la médiane.

La vérification est immédiate.

Notons que si $(x|y) = 0$ la première relation est le théorème de Pythagore, mais la réciproque n'est vraie que si l'espace vectoriel est réel.

2.3 Théorème de projection

Définition 2.4 — Soit F une partie d'un espace de Hilbert E et $a \in E$. On appelle projection de a sur F le point $\alpha \in F$ (s'il existe) tel que

$$\|a - \alpha\| = \inf_{x \in F} \|a - x\|.$$

C'est le point réalisant le minimum de la distance de a à F (si ce point existe et dans ce cas il peut y en avoir plusieurs).

- Une partie F d'un espace vectoriel E est convexe si lorsqu'elle contient deux points elle contient le segment joignant ces deux points, où le segment $[a, b]$ est l'ensemble des points $\{x \in E | x = \alpha a + (1 - \alpha)b, 0 \leq \alpha \leq 1\}$

Théorème 2.5 Soit E un espace Hilbertien et F une partie fermée convexe non vide de E . Chaque point de E a une projection unique sur F .

Démonstration

On pose $d = \inf_{x \in F} \|a - x\|$. Les propriétés de la borne inférieure entraîne l'existence d'une suite (x_n) de points de F telle que $\lim_{n \rightarrow \infty} \|a - x_n\| = d$. En appliquant le théorème de la médiane à $x = a - x_n$ et $y = a - x_m$ on obtient

$$(1) \|a - x_n\|^2 + \|a - x_m\|^2 = 2\|a - \frac{x_n + x_m}{2}\|^2 + 1/2\|x_n - x_m\|^2$$

$\lim_{n \rightarrow \infty} \|a - x_n\|^2 = \lim_{m \rightarrow \infty} \|a - x_m\|^2 = d^2$, F étant convexe $\frac{x_n + x_m}{2} \in F$. La définition de d entraîne $\|a - \frac{x_n + x_m}{2}\| \geq d$. On a $\limsup_{n, m \rightarrow \infty} \|x_n - x_m\| \leq 0$ ce qui donne $\lim_{n, m \rightarrow \infty} \|x_n - x_m\| = 0$. La suite (x_n) est donc une suite de Cauchy de F , E étant complet, elle converge vers un élément $\alpha \in E$ et F étant fermé $\alpha \in F$. Ceci montre l'existence d'un point $\alpha \in F$ vérifiant $d = \|a - \alpha\|$, α est donc bien une projection de a sur F .

S'il existait une autre projection β on pourrait écrire

$$\|\beta - \alpha\|^2 = 2\|\beta - a\|^2 + 2\|\alpha - a\|^2 - 4\|a - \frac{\alpha + \beta}{2}\|^2$$

d'où $\|\beta - \alpha\|^2 \leq 2d^2 + 2d^2 - 4d^2 \leq 0 \Rightarrow \|\beta - \alpha\| = 0 \Rightarrow \alpha = \beta$.

Théorème 2.6 (Caractérisation de la projection) Soit F partie fermée convexe non vide de l'espace de Hilbert E , et $a \in E$ on a

$$\alpha \text{ est la projection de } a \text{ sur } F \Leftrightarrow \forall x \in F, \Re(a - \alpha | x - \alpha) \leq 0$$

Démonstration

Si $\Re(a - \alpha|x - \alpha) \leq 0$ pour $x \in F$ on a

$$\|a - x\|^2 = \|a - \alpha\|^2 + \|x - \alpha\|^2 - 2\Re(a - \alpha|x - \alpha) \geq \|a - \alpha\|^2$$

et donc α est bien la projection de a sur F .

Réciproquement, si α est la projection de a sur F on a pour $x \in F$
 $(1 - t)\alpha + tx \in F$, $t \in [0, 1]$ d'après la convexité de F .

$$\forall x \in F, \forall t \in [0, 1], \|a - ((1 - t)\alpha + tx)\|^2 = \|(a - \alpha) - t(x - \alpha)\|^2 \geq \|a - \alpha\|^2$$

Après développement et simplification on obtient

$-2t\Re(a - \alpha|x - \alpha) + t^2\|x - \alpha\|^2 \geq 0$ en divisant par $t \in]0, 1]$ et en faisant tendre t vers 0 on obtient $\Re(a - \alpha|x - \alpha) \leq 0$ pour tout $x \in F$.

Remarque.

Si le corps de base était \mathbb{R} , en posant $\cos \theta = \frac{(x|y)}{\|x\|\|y\|}$ ceci exprime que l'angle θ est obtus ou droit.

2.4 Projection sur un sous-espace vectoriel fermé

Un sous-espace vectoriel étant convexe, ce qui précède s'applique aux sous-espaces fermés d'un espace de Hilbert E .

Définition 2.5 — Les vecteurs x et y de E sont orthogonaux si $(x|y) = 0$.

— F étant une partie de E , x est orthogonal à F s'il est orthogonal à tout élément de F .

On note F^\perp l'ensemble des éléments orthogonaux à F . Cet ensemble, appelé orthogonal de F est un sous-espace vectoriel fermé de E car $F^\perp = \bigcap_{a \in F} \{a\}^\perp$ et l'orthogonal d'un élément $a \in E$ est le noyau de la forme linéaire continue $\phi_a : x \rightarrow (a|x)$, c'est donc $\phi_a^{-1}(\{0\})$ qui est fermé.

— Deux parties F et G sont orthogonales si tout vecteur de l'une est orthogonal à tout vecteur de l'autre.

Remarque.

En appelant G l'espace vectoriel engendré par F on a $F^\perp = G^\perp = \overline{G}^\perp$. La première égalité est vraie car un vecteur orthogonal à d'autres vecteurs est orthogonal à leurs combinaisons linéaires, la deuxième s'obtient par passage à la limite en utilisant la continuité des ϕ_a .

Théorème 2.7 (de projection sur un sous espace fermé) La projection d'un point a d'un espace Hilbertien E sur un sous-espace fermé F est l'unique point α de F tel que $a - \alpha$ soit orthogonal à F .

Démonstration Soit $x \in F$, le vecteur $y = x + \alpha$ est alors élément de F qui est espace vectoriel.

$\Re(a - \alpha|y - \alpha) = \Re(a - \alpha|x) \leq 0$, en remplaçant x par $-x$ on obtient $\Re(a - \alpha|x) = 0$, puis en remplaçant x par ix on obtient $\Im(a - \alpha|x) = 0$ ce qui entraîne $(a - \alpha|x) = 0$ pour tout $x \in F$, ce qui exprime que $a - \alpha \perp F$.

Si $\beta \in F$ est tel que $a - \beta \perp F$ on a $(a - \alpha|x) = (a - \beta|x) = 0$ pour tout $x \in F$, en faisant la différence on obtient $(\alpha - \beta|x) = 0$ et si $x = \alpha - \beta$

$$(x|x) = 0 \Rightarrow x = \alpha - \beta = 0 \Rightarrow \alpha = \beta$$

il existe donc un seul point α vérifiant $a - \alpha \perp F$.

Dans la suite on notera P_F la projection de E sur F .

Théoreme 2.8 *Si F n'est pas réduit au sous-espace $\{0\}$ P_F est linéaire, continue, de norme 1. $\ker P_F = F^\perp$ et $E = F \oplus F^\perp$ (les deux sous-espaces sont supplémentaires).*

Démonstration

Soient $x, y \in E, \forall z \in F$ on a $(P_F x - x|z) = (P_F y - y|z) = 0$

d'où $(P_F x + P_F y - (x + y)|z) = 0$ ce qui entraîne $P_F(x + y) = P_F(x) + P_F(y)$.

De même si $\lambda \in \mathbb{C}$ $(\lambda P_F(x) - \lambda x|z) = 0 \Rightarrow P_F(\lambda x) = \lambda P_F(x)$. P_F est donc linéaire.

L'égalité $(P_F(x) - x|P_F(x)) = 0$ entraîne $\|P_F x\|^2 = (x|P_F x) \leq \|x\| \|P_F x\|$ donc $\|P_F x\| \leq \|x\|$ et $\|P_F\| \leq 1$.

Si $x \in F$ on a $\|P_F x\| = \|x\|$ d'où $\|P_F\| \geq 1$ et donc $\|P_F\| = 1$.

$$P_F x = 0 \Leftrightarrow \forall z \in F (x|z) = 0 \Leftrightarrow x \in F^\perp$$

ce qui démontre que $\ker P_F = F^\perp$.

Un vecteur $x \in E$ s'écrit $x = (x - P_F x) + P_F x$, le premier vecteur est un élément de F^\perp et le second un élément de F , on a de plus $F \cap F^\perp = \{0\}$ ces deux résultats établissent que $E = F \oplus F^\perp$.

lemme 2.8.1 *Soit v un vecteur d'un espace de Hilbert E . $f_v : x \rightarrow (x|v)$ est une forme linéaire continue sur E de norme $\|v\|$. L'application $v \rightarrow f_v$ est une isométrie semi linéaire de E vers E^* .*

Démonstration

En appliquant l'inégalité de Cauchy-Schwarz on a

$|(x|v)| \leq \|x\| \|v\| \Rightarrow \|f_v\| \leq \|v\|$. De plus $f_v(v) = \|v\|^2$, on déduit donc

$\|f_v\| \geq \|v\|$ d'où $\|f_v\| = \|v\|$.

$v \rightarrow f_v$ est bien une isométrie semi-linéaire de E dans E^* , le théorème suivant montre qu'elle est surjective.

Théoreme 2.9 *Si f est une forme linéaire continue sur l'espace de Hilbert E , il existe un vecteur unique v tel que $\forall x \in E, f(x) = (x|v)$.*

Démonstration Si $f = 0$ $v = 0$ convient. Sinon $F = f^{-1}(\{0\})$ est un hyperplan fermé, f étant continue. Le théorème de projection entraîne l'existence d'un vecteur $u \perp F$. Posons $g(x) = (x|u)$, cette forme linéaire a le même noyau que f , on a donc $f = \lambda g$ pour un $\lambda \in \mathbb{C}$, soit $\forall x \in E f(x) = \lambda(x|u)$, on prend alors $v = \bar{\lambda}u$

Application.

Pour toute forme linéaire continue ϕ sur $E = L^2_\mu(X)$ on a :

$$\exists g \in E, \forall f \in E, \phi(f) = \int_X f \bar{g} d\mu \text{ et } \|\phi\| = \|g\|_2.$$

On peut donc identifier E et son dual topologique.

2.5 Espaces de Hilbert et systèmes orthogonaux

Définition 2.6 — Un espace de Hilbert E est séparable s'il existe un sous-ensemble dénombrable $A = \{x_1, x_2, \dots\}$ dense dans E , c'est à dire tel que $\overline{A} = E$. Il est équivalent de dire que toute boule ouverte non vide de E contient au moins un élément de A .

— Une famille $(e_i)_{i \in I}$ de vecteurs de E est orthonormée si $(e_i | e_j) = 0$ pour $i \neq j$ et $\|e_i\| = 1$ pour tout $i \in I$.

Théorème 2.10 — Toute famille orthonormée est libre, ce qui signifie que toute combinaison linéaire composée d'un nombre fini de vecteurs donne le vecteur nul seulement si tous les coefficients scalaires sont nuls.

— Dans un espace de Hilbert séparable toute famille orthonormée est finie ou dénombrable.

Démonstration

Soit J une partie finie quelconque de I et $(e_i)_{i \in I}$ une famille orthonormée.

$$\sum_{i \in J} \alpha_i e_i = 0 \Rightarrow \forall j \in J (e_j | \sum_{i \in J} \alpha_i e_i) = 0 = \sum_{i \in J} \overline{\alpha_i} (e_j | e_i) = \overline{\alpha_j}$$

ce qui démontre qu'une famille orthonormée est une famille libre.

Soit $(e_i)_{i \in I}$ une famille orthonormée dans un espace de Hilbert séparable E . Si $e_j \neq e_i$ $\|e_j - e_i\|^2 = \|e_j\|^2 + \|e_i\|^2 - 2\Re(e_j | e_i) = 2$ donc $\|e_j - e_i\| = \sqrt{2}$.

Soit $A = \{x_1, x_2, \dots\}$ un ensemble dénombrable dense dans E . Pour chaque $i \in I$ il existe un entier $n(i)$ tel que $\|x_{n(i)} - e_i\| < 1/4$ car A est dense dans E . Montrons que l'application $i \rightarrow n(i)$ est injective. Soit $n(i) = n(j)$

$$\|e_i - e_j\| \leq \|e_i - x_{n(i)}\| + \|x_{n(i)} - e_j\| < 1/2 \Rightarrow e_i = e_j.$$

Cette application de I dans \mathbb{N} étant injective, l'ensemble I est fini ou dénombrable. Dans la suite l'espace E sera toujours supposé séparable

Théorème 2.11 Soit $(e_i)_{1 \leq i \leq n}$ un système orthonormé fini de E et F le sous-espace à n dimensions engendré par ces vecteurs. Pour tout x de E on a

- 1) $P_F x = \sum_{i=1}^n (x | e_i) e_i$.
- 2) $d^2(x, F) = \|x\|^2 - \sum_{i=1}^n |(x | e_i)|^2$ où $d(x, F) = \inf_{y \in F} \|x - y\|$.

Démonstration

- 1) On a $\forall j, 1 \leq j \leq n (x - \sum_{i=1}^n (x | e_i) e_i | e_j) = 0$, donc $x - \sum_{i=1}^n (x | e_i) e_i \perp F$.
- 2) $d^2(x, F) = \|x - \sum_{i=1}^n (x | e_i) e_i\|^2 = \|x\|^2 - \sum_{i=1}^n |(x | e_i)|^2 - \sum_{i=1}^n |(x | e_i)|^2 + \sum_{i=1}^n |(x | e_i)|^2$. Ceci démontre que

$$\forall (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{C}^n, \|x - \sum_{i=1}^n \lambda_i e_i\| \geq \|x - \sum_{i=1}^n (x | e_i) e_i\|$$

Corollaire 2.11.1 *Inégalité de Bessel* Soit $(e_i)_{i \in I}$, L 'ensemble I étant fini ou dénombrable, un système orthonormé quelconque et $x \in E$, on a

$$\sum_{i \in I} |(x|e_i)|^2 \leq \|x\|^2.$$

La procédure qui suit, qui permet de construire des systèmes orthonormés en dimension finie s'étend aux espaces de Hilbert.

Théoreme 2.12 (Procédé d'orthogonalisation de Gram-Schmidt) Soit E un espace de Hilbert et $(x_n, n \geq 1)$ des vecteurs linéairement indépendants de E . On définit

$$\begin{aligned} w_1 &= x_1 & , v_1 &= w_1 / \|w_1\| \\ w_2 &= x_2 - (v_1|x_2)v_1 & , v_2 &= w_2 / \|w_2\| \\ &\vdots & & \\ w_n &= x_n - \sum_{k=1}^{n-1} (v_k|x_n)v_k & , v_n &= w_n / \|w_n\| \\ &\vdots & & \end{aligned}$$

La suite $(v_n, n \geq 1)$ est un système orthonormé et pour tout $n \geq 1$, (v_1, v_2, \dots, v_n) et (x_1, x_2, \dots, x_n) engendrent le même sous espace vectoriel de E .

La démonstration de ce théorème est aisée et laissée en exercice. Application.

Déterminer

$$\min_{a,b,c} \int_{-1}^1 (x^3 - a - bx - cx^2)^2 dx.$$

Considérons $E = L^2([-1, 1], \mathbb{R})$ et F le sous espace vectoriel de dimension 3 engendré par f_0, f_1, f_2 où $f_i : x \rightarrow x^i$. F est complet car c'est un sous espace de dimension finie. Le problème revient à calculer la distance d de $f : x \rightarrow x^3$ à ce sous espace (l'intégrale est égale à d^2). Déterminons tout d'abord $P_F f$, l'intégrale cherchée sera $\|f - P_F f\|^2$.

Construisons une base orthonormée de F par la méthode de Gram-Schmidt.

$$v_0(x) = \frac{f_0(x)}{\|f_0\|} = \sqrt{2}/2$$

$$w_1(x) = x - (v_0|f_1)v_0 = x, \quad v_1(x) = (\sqrt{6}/2)x$$

$$w_2(x) = x^2 - (v_0|f_2)v_0 - (v_1|f_2)v_1 = x^2 - 1/3, \quad v_2 = (2\sqrt{2}/3)(x^2 - 1/3).$$

$$P_F f = (f|v_0)v_0 + (f|v_1)v_1 + (f|v_2)v_2 = (3/5)x \quad \|f - P_F f\|^2 = \int_{-1}^1 (x^3 - 3/5x)^2 dx = 8/175$$

Définition 2.7 — Un système orthonormé $A = (e_i, i \geq 1)$ est complet si

$$\forall x \in E, \forall \epsilon > 0, \exists n_0 \in \mathbb{N} \|x - \sum_{i=1}^{n_0} (x|e_i)e_i\| \leq \epsilon.$$

- Un système orthonormé est fermé si pour tout $x \neq 0$ de E on peut trouver $n \in \mathbb{N}^*$ tel que $(e_n|x) \neq 0$.
- Un système orthonormé est maximal s'il n'est pas sous ensemble propre d'un autre système orthonormé.

Théoreme 2.13 Soit $A = (e_i, i \in I)$, $I = \{1, 2, \dots, n\}$ ou $I = \mathbb{N}^*$, un système orthonormé dans un espace de Hilbert E . Les propriétés suivantes sont équivalentes.

- 1) A est complet.
- 2) A est maximal dans E .
- 3) A est fermé.
- 4) Si $x \in E$, $(x|x) = \|x\|^2 = \sum_{i \in I} |(x|e_i)|^2$. (Identité de Parseval).
- 5) Si $x, y \in E$, $(x|y) = \sum_{i \in I} (x|e_i)(y|e_i)$.

Démonstration

1 \Rightarrow 2

Si A n'est pas maximal dans E $\exists e_0 \in E$ tel que $(e_i, i \in \mathbb{N})$ est un système orthonormé. Soit $c_i = (e_0|e_i)$

$$\forall n \in \mathbb{N}, \|e_0 - \sum_{i=1}^n c_i e_i\| = \|e_0\| = 1$$

A n'est donc pas complet.

2 \Rightarrow 3

Si A n'est pas fermé, il existe $x \neq 0$, $x \in E$ tel que quel que soit $i \in I$, $(x|e_i) = 0$. En posant $e_0 = x/\|x\|$ le système $(e_i, i \in \mathbb{N})$ est orthonormé et A n'est donc pas maximal.

3 \Rightarrow 4

L'inégalité de Bessel entraîne $\sum_{i \in I} |(x|e_i)|^2 \leq \|x\|^2$, montrons l'inégalité inverse dans le cas le plus général $I = \mathbb{N}^*$. Posons $s_n = \sum_{i=1}^n (x|e_i)e_i$

$$\|s_n - s_m\|^2 = \left(\sum_{i=m+1}^n (x|e_i)e_i \mid \sum_{i=m+1}^n (x|e_i)e_i \right) = \sum_{i=m+1}^n |(x|e_i)|^2 \leq \sum_{i=m+1}^{\infty} |(x|e_i)|^2 \quad (n > m).$$

Le dernier terme est le reste d'une série convergente (inégalité de Bessel) et tend donc vers zéro quand n et m tendent vers zéro, (s_n) est donc une suite de Cauchy de E , convergeant vers un élément $y \in E$ car E est complet.

$$\forall n \in \mathbb{N}^* (e_n|x) = (e_n|y) \Rightarrow \forall n (e_n|x - y) = 0 \Rightarrow x = y$$

car le système est fermé.

4 \Rightarrow 5

$(x|y)$ et $\sum_{i \in I} (x|e_i)\overline{(y|e_i)}$ sont deux formes hermitiennes prenant les mêmes valeurs sur la diagonale, donc identiques.

5 \Rightarrow 4 est immédiat.

4 \Rightarrow 1

Soit n_0 tel que

$$n > n_0 \Rightarrow \|x\|^2 - \sum_{i=1}^n |(x|e_i)|^2 < \epsilon, \text{ alors } \|x - \sum_{i=1}^n (x|e_i)e_i\|^2 = \|x\|^2 - \sum_{i=1}^n |(x|e_i)|^2 < \epsilon.$$

Théoreme 2.14 *Un espace de Hilbert est séparable si et seulement il contient un système orthonormé complet fini ou dénombrable.*

Démonstration

Supposons E séparable et soit (x_n) un sous ensemble dénombrable qui est dense.

On peut en extraire une sous suite (x_{n_k}) de vecteurs linéairement indépendants telle que tout x_n soit une combinaison linéaire finie des x_{n_k} . Le procédé de Gram-Schmidt appliqué à cette suite permet de construire un système orthonormé (y_n) tel que tout x_n soit combinaison linéaire finie des y_n . Un vecteur $x \in E$ orthogonal à tous les y_n est donc orthogonal à tous les x_n donc à $E = \overline{(x_n)}$ ce qui entraîne que le système (y_n) est fermé.

Nous admettrons la réciproque de ce théorème.

Théorème 2.15 *Un espace de Hilbert séparable est isométrique à l'espace $l^2 = l^2(\mathbb{N})$ s'il est de dimension infinie, à un espace \mathbb{C}^n s'il est de dimension finie.*

Démonstration

Si $\dim E = n$ l'application $x = \sum_{i=1}^n (x|e_i)e_i \rightarrow ((x|e_1), (x|e_2), \dots, (x|e_n))$, où $(e_i, 1 \leq i \leq n)$ est une base orthonormée de E , est une isométrie car $\|x\|^2 = \sum_{i=1}^n |(x|e_i)|^2$.

Si E est de dimension infinie choisissons $A = (e_i, i \in \mathbb{N}^*)$ un système orthonormé complet et posons $x_i = (x|e_i)$. On a démontré que $\|x\|^2 = \sum_{i=1}^{\infty} |x_i|^2$, l'application

$$\phi : E \rightarrow l^2 \quad x \rightarrow (x_i, i \in \mathbb{N}^*)$$

est donc une isométrie, montrons qu'elle est surjective. On peut remarquer que la famille $B = (f_i, i \in \mathbb{N}^*)$, définie par $f_i(j) = \delta_i^j$, est un système orthonormé complet de l^2 car $(f_i|f_j) = \delta_i^j$ et $\|(a_i)\|_{l^2}^2 = \sum_{i=1}^{\infty} |(a_i|f_i)|^2$.

$B \subset \phi(E)$, $\phi(E)$ contenant les combinaisons linéaires finies d'éléments de B est donc dense dans l^2 , or $\phi(E)$ est complet comme image isométrique d'un espace de Hilbert, il est donc fermé dans l^2 , on a donc $\phi(E) = l^2$. Nous allons maintenant étudier un cas particulier important comprenant la convergence en moyenne quadratique (dans L^2) des séries de Fourier.

2.6 Bases orthonormales de $L^2([a, b])$

Théorème 2.16 (Critère de Dazzell) *Une famille orthonormée $(f_n, n \geq 1)$ de $L_m^2([a, b])$, m étant la mesure de Lebesgue de $[a, b]$, est un système orthonormé complet si et seulement si*

$$\frac{2}{(a-b)^2} \sum_{n \geq 1} \int_a^b \left| \int_a^u f_n(t) dt \right|^2 du = 1$$

Démonstration

Rappelons que si A est un sous ensemble de E la fonction indicatrice χ_A notée aussi $\mathbf{1}_A$ est définie par $\chi_A(x) = 0$ si $x \notin A$ et $\chi_A(x) = 1$ si $x \in A$.

Supposons le système $(f_n, n \geq 1)$ orthonormé complet. $\forall u \in [a, b]$, $\chi_{[a, u]} \in L_m^2$ on peut donc écrire

$$\chi_{[a, u]} = \sum_{n=1}^{\infty} c_n f_n, \quad c_n = (\chi_{[a, u]}|f_n) = \int_a^u \overline{f_n(t)} dt$$

L'égalité de Parseval entraîne

$$\sum_{n=1}^{\infty} \left| \int_a^u f_n(t) dt \right|^2 = u - a \Rightarrow \sum_{n=1}^{\infty} \int_a^b \left| \int_a^u f_n(t) dt \right|^2 du = \frac{(b-a)^2}{2}.$$

Nous admettrons la réciproque de ce théorème.

2.6.1 Application

Corollaire 2.16.1 *La famille de fonctions*

$$e_n : x \rightarrow \frac{e^{inx}}{\sqrt{2\pi}}, \quad n \in \mathbb{Z}$$

est un système orthonormé complet de l'espace de Hilbert $E = L_m^2([0, 2\pi])$.

Démonstration

Montrons que cette famille de fonctions vérifient le critère de Dalzell.

$$\begin{aligned} \frac{2}{4\pi^2} \sum_{n \in \mathbb{Z}} \int_0^{2\pi} \frac{1}{2\pi} \left| \int_0^u e^{int} dt \right|^2 du &= \frac{2}{8\pi^3} \left(\frac{8\pi^3}{3} + \sum_{n \in \mathbb{Z}^*} \frac{4\pi}{n^2} \right) \\ &= 2/3 + \frac{2}{8\pi^3} \left(2 \sum_{n=1}^{\infty} \frac{4\pi}{n^2} \right) = 2/3 + \frac{2}{8\pi^3} \cdot \frac{8\pi^3}{6} = 1. \end{aligned}$$

Une fonction f étant élément de E on a

$$(f|e_n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx} dx \text{ et } f = \sum_{n \in \mathbb{Z}} (f|e_n) e_n \text{ dans } E$$

$$f = \sum_{n \in \mathbb{Z}} \left(\frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx \right) f_n \text{ où } f_n(x) = e^{inx}.$$

On note $c_n(f) = (1/2\pi) \int_0^{2\pi} f(x) e^{-inx} dx$, ce qui entraîne $f = \sum_{n \in \mathbb{Z}} c_n(f) f_n$ dans E , ce qui signifie

$$\lim_{N \rightarrow \infty} \left\| f - \sum_{n=-N}^N c_n(f) f_n \right\|_2 = 0.$$

La série $\sum_{n \in \mathbb{Z}} c_n(f) e^{inx}$ est appelée série de Fourier de f . En écrivant $e^{inx} = \cos(nx) + i \sin(nx)$ la série précédente peut également s'exprimer sous la forme $\sum_{n=0}^{\infty} a_n \cos(nx) + b_n \sin(nx)$. On a les relations $a_0 = c_0$, $a_n = c_n + c_{-n}$, $b_n = ic_n - ic_{-n}$.

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \quad a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad (n > 0), \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \quad (n > 0).$$

Remarque.

- On peut supposer $f(0) = f(2\pi)$ car les coefficients de la série de Fourier sont inchangés en modifiant un nombre fini de valeurs prises par f , ceci permet de considérer f comme la restriction à $[0, 2\pi]$ d'une fonction 2π -périodique.
- Le calcul des coefficients peut dans ces conditions être effectué sur $[a, a + 2\pi]$ avec a quelconque.

- Si f est une fonction paire, $\forall n > 0$ $b_n = 0$ et si elle est impaire $\forall n \in \mathbb{N}$ $a_n = 0$.

Les résultats du théorème 2.13 se traduisent de la façon suivante pour les séries de Fourier.

4) Identité de Parseval

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{k \in \mathbb{Z}} |c_k|^2 = |c_0|^2 + \sum_{k=1}^{\infty} (|c_k|^2 + |c_{-k}|^2)$$

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = |a_0|^2 + \frac{1}{2} \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2)$$

pour une fonction $f \in L_m^2([0, 2\pi])$

$$5) \forall f, g \in L_m^2([0, 2\pi]), \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} dt = \sum_{n \in \mathbb{Z}} c_n(f) \overline{c_n(g)}.$$

L'isométrie entre $L_m^2([0, 2\pi])$ et l^2 se traduit par

pour toute suite $(c_n, n \in \mathbb{Z})$ vérifiant $\sum_{n \in \mathbb{Z}} |c_n|^2 < +\infty$ il existe $f \in L_m^2([0, 2\pi])$ telle que la série de Fourier de f soit $\sum_{n \in \mathbb{Z}} c_n e^{inx}$.

2.6.2 Convergence ponctuelle d'une série de Fourier

Les coefficients a_n, b_n, c_n sont également définis si $f \in L_m^1([0, 2\pi])$ car $|f(t)e^{int}| = |f(t)|$. Dans le cas où f est de plus 2π -périodique le problème de la convergence ponctuelle (convergence simple ou uniforme) de $S_N(f) = \sum_{n=-N}^N c_n e^{inx}$ se pose. Le théorème de Dirichlet, que nous allons démontrer donne des conditions pour la convergence ponctuelle de cette série. Commençons par transformer $S_N(f)$.

L'expression $D_N(t) = \sum_{n=-N}^N c_n e^{inx}$ est appelé noyau de Dirichlet.

$$S_N(f)(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \left(\sum_{n=-N}^N e^{in(x-t)} \right) dt = \frac{1}{2\pi} \int_0^{2\pi} f(t) D_N(x-t) dt$$

$$\begin{aligned} S_N(f)(x) &= \frac{1}{2\pi} \int_0^{2\pi} f(x+t) D_N(t) dt = \frac{1}{2\pi} \int_0^{2\pi} f(x-t) D_N(t) dt \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} (f(x+t) + f(x-t)) D_N(t) dt. \end{aligned}$$

La deuxième égalité provient de la parité de la fonction D_N .

Théorème 2.17 (Lemme de Riemann-Lebesgue.) *On considère une fonction $f \in L_m^1([0, 2\pi])$, les coefficients $c_n = (1/(2\pi)) \int_0^{2\pi} f(x) e^{-inx} dx$ de la série de Fourier de f tendent vers zéro quand $|n|$ tend vers zéro.*

Si $f \in L_m^2([0, 2\pi])$ c'est immédiat car le terme général d'une série convergente (identité de Parseval) tend vers zéro quand $n \rightarrow \infty$ sinon $L_m^2([0, 2\pi])$ est dense dans $L_m^1([0, 2\pi])$ et si $f \in L^1$ il existe $g \in L^2$ telle que $\|f - g\|_1 < \epsilon$.

$$|c_n(f)| \leq |c_n(f) - c_n(g)| + |c_n(g)| \leq \|f - g\|_1 + |c_n(g)|$$

et $c_n(f)$ tend bien vers zéro quand $|n|$ tend vers l'infini.

Remarque.

Ce résultat peut être démontré pour un intervalle $[a, b]$ quelconque.

Théoreme 2.18 (Critère de Dini) Si $f \in L^1_m([-\pi, \pi])$, $A \in \mathbb{R}$, f 2π -périodique et si

$$\int_0^\pi \frac{|f(x+t) + f(x-t) - 2A|}{t} dt < +\infty$$

la série de Fourier de f converge en x et $\lim_{n \rightarrow \infty} S_N(f)(x) = A$

On peut remarquer que

$$D_N(t) = e^{-2iNx}(1 + e^{ix} + \dots + e^{2iNx}) = e^{-2iNx} \frac{1 - e^{i(2N+1)x}}{1 - e^{ix}} = \frac{\sin(2N+1)t/2}{\sin(t/2)}$$

($x \neq 0 \pmod{2\pi}$)

de plus $\int_{-\pi}^\pi D_N(t) dt = 2\pi$ et $\int_0^\pi D_N(t) dt = \pi$ cette dernière valeur étant obtenue par parité de D_N .

$$S_N(f)(x) - A = \frac{1}{2\pi} \int_0^\pi \frac{f(x+t) + f(x-t) - 2A}{t} \frac{t}{\sin(t/2)} \sin((N+1/2)t) dt$$

ceci tend vers zéro d'après le lemme de Riemann-Lebesgue car les deux premiers termes de l'intégrale sont intégrables (le premier d'après l'hypothèse) et le troisième est la partie imaginaire de $e^{i(N+1/2)t}$.

Théoreme 2.19 (Théorème de Dirichlet) Soit f une fonction 2π -périodique et intégrable sur $[0, 2\pi]$ admettant une limite à gauche et à droite en $x \in [0, 2\pi]$ et dérivable à gauche et à droite en ce point. La série de Fourier de f converge en x vers la demi-somme des limites des limites à gauche et à droite.

$$\lim_{N \rightarrow \infty} S_N(f)(x) = \frac{f(x+) + f(x-)}{2}.$$

Ceci est une conséquence immédiate du critère de Dini avec

$A = (f(x+) + f(x-))/2$ de plus

$$t \rightarrow \frac{|f(x+t) + f(x-t) - f(x+) - f(x-)|}{t}$$

est intégrable sur $[0, \pi]$, f étant dérivable à gauche et à droite de x .

Remarque.

Si les conditions du théorème sont vérifiées et si f est continue en x la série de Fourier de f converge alors vers $f(x)$ en x .

2.7 Opérateur linéaire sur un espace de Hilbert

Un opérateur linéaire sur E est une application linéaire de E dans E , sauf mention contraire ils seront continus.

Théoreme 2.20 Une forme sesquilinéaire est continue sur $E \times E$ si et seulement si

$$\exists M \in \mathbb{R}^+, \forall (x, y) \in E \times E, |f(x, y)| \leq M \|x\| \|y\|.$$

Démonstration

Supposons f continue, la continuité en 0 se traduit par

$$\exists a, b \in \mathbb{R}_+^* , \|x\| < a , \|y\| < b \Rightarrow |f(x, y)| \leq 1$$

Par homogénéité on a

$$\forall x, y \in E , |f(x, y)| \leq \frac{1}{ab} \|x\| \|y\|.$$

Réciproquement, si la condition est vérifiée on a $x, y, \xi, \eta \in E$

$$|f(x + \xi, y + \eta) - f(x, y)| \leq M (\|\xi\| \|y + \eta\| + \|\eta\| \|x\|).$$

Ceci montre la continuité de f .

Théorème 2.21 — Si u est un opérateur linéaire sur E ,

$\phi_u : (x, y) \rightarrow (ux|y)$ définit une forme sesquilinéaire continue sur $E \times E$ telle que $\|\phi_u\| = \|u\|$.

— Réciproquement si ϕ est une forme sesquilinéaire continue sur $E \times E$ il existe $u \in \mathcal{L}(E)$ tel que $\phi = \phi_u$.

Remarque. On définit

$$\|\phi\| = \sup_{x \neq 0, y \neq 0} \frac{\phi(x, y)}{\|x\| \|y\|}$$

et donc

$$|\phi(x, y)| \leq \|\phi\| \|x\| \|y\|$$

et le M du théorème est tel que $M \geq \|\phi\|$.

Démonstration

La sesquilinearité de ϕ_u est immédiate. L'inégalité de Cauchy-Schwarz entraîne

$$|(ux|y)| \leq \|ux\| \|y\| \leq \|u\| \|x\| \|y\| \Rightarrow \|\phi_u\| \leq \|u\|.$$

En prenant $y = ux$ on obtient

$$\|ux\|^2 = \phi_u(x, ux) \leq \|\phi_u\| \|x\| \|ux\| \Rightarrow \|ux\| \leq \|\phi_u\| \|x\| \Rightarrow \|u\| \leq \|\phi_u\|$$

ces deux inégalités permettant d'écrire $\|\phi_u\| = \|u\|$.

Réciproquement soit ϕ une forme sesquilinéaire continue sur $E \times E$. L'application $g : y \rightarrow \phi(x, y)$ est linéaire continue (car $|g(y)| \leq (\|\phi\| \|x\|) \|y\|$). D'après le théorème 2.9 il existe $v \in E$ tel que $g(y) = (y|v)$ pour tout $y \in E$, ou encore $(v|y) = \phi(x, y)$. Posons $v = ux$, on vérifie que $u(\lambda x + \mu y) = \lambda u(x) + \mu u(y)$ si $x, y \in E$, $\lambda, \mu \in \mathbb{C}$, u est donc linéaire.

$$\|ux\|^2 = \phi(x, ux) \leq \|\phi\| \|x\| \|ux\| \Rightarrow \|ux\| \leq \|\phi\| \|x\|$$

l'application linéaire u est donc continue.

Théorème 2.22 *A tout opérateur $u \in \mathcal{L}(E)$ (opérateur linéaire continu sur E) il existe un élément unique $u^* \in \mathcal{L}(E)$ tel que*

$$\forall x, y \in E, (ux|y) = (x|u^*y).$$

u^* est appelé opérateur adjoint de u .

Démonstration

Posons $\psi(y, x) = (y|ux)$. C'est une forme sesquilinéaire continue de norme $\|u\|$, d'après le théorème précédent il existe $u^* \in \mathcal{L}(E)$ tel que $\psi = \phi_{u^*} : (y, x) \rightarrow (u^*y|x)$, d'où

$$\forall (x, y) \in E \times E, (x|u^*y) = (ux|y) \text{ et } \|\phi_{u^*}\| = \|u^*\| = \|\psi\| = \|u\|$$

d'où $\|u^*\| = \|u\|$.

Propriétés.

- 1) $u^{**} = u$.
 - 2) $(uv)^* = v^*u^*$.
 - 3) $\|u^*\| = \|u\|$.
 - 4) $\|uu^*\| = \|u^*u\| = \|u\|^2 = \|u^*\|^2$.
 - 5) $(\lambda u)^* = \bar{\lambda}u^*$.
 - 6) $(u + v)^* = u^* + v^*$.
 - 7) Si u est inversible, u^* est inversible et $(u^*)^{-1} = (u^{-1})^*$.
- 1,2,3,5,6 sont immédiats, 4 sera montré plus loin. Montrons 7. Posons $v = u^{-1}$.

$$uv = vu = I \Rightarrow v^*u^* = u^*v^* = I \Rightarrow u^* \text{ est inversible et } (u^*)^{-1} = (u^{-1})^*.$$

Définition 2.8 — $u \in \mathcal{L}(E)$ est normal si $u^*u = uu^*$.

- u est unitaire si $u^*u = uu^* = I$.
- u est autoadjoint ou hermitien si $u = u^*$.

Exemple.

$E = L^2_\mu(X)$, $u : E \rightarrow E$ $f \rightarrow hf$ où h est mesurable bornée.

$$\int_X (hf)\bar{g}d\mu = \int_X f\bar{h}gd\mu \Rightarrow u^*(g) = \bar{h}g$$

donc $uu^* = u^*u$, u est donc un opérateur normal. Si h est réelle u est autoadjoint.

Remarque.

Si u est autoadjoint $\phi_u : (x, y) \rightarrow (ux|y)$ est sesquilinéaire hermitienne (la réciproque est vraie), c'est pourquoi on dit aussi que u est hermitien.

Théorème 2.23 —

$$\frac{1}{2}\|u\| \leq \sup_{\|x\| \leq 1} |(ux|x)| \leq \|u\|.$$

— si u est hermitien $\|u\| = \sup_{\|x\| \leq 1} |(ux|x)|$.

Démonstration

$$\|u\| = \|\phi_u\| \Rightarrow \|u\| = \sup_{\|x\| \leq 1} \|ux\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |(ux|y)| \geq \sup_{\|x\| \leq 1} |(ux|x)|.$$

Posons $\alpha = \sup_{\|x\| \leq 1} |(ux|x)|$, on a $\forall x \in E$, $|(ux|x)| \leq \alpha \|x\|^2$ $\|u\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |(ux|y)|$, en utilisant 2.1 pour la forme sesquilinéaire ϕ_u on obtient

$$\begin{aligned} \|u\| &\leq \frac{1}{4} \left(\alpha \|x+y\|^2 + \alpha \|x-y\|^2 + \alpha \|x+iy\|^2 + \alpha \|x-iy\|^2 \right) \\ &= \frac{\alpha}{4} \left(2(\|x\|^2 + \|y\|^2) + 2(\|x\|^2 + \|y\|^2) \right) \leq 2\alpha. \end{aligned}$$

on a donc $(1/2)\|u\| \leq \sup_{\|x\| \leq 1} |(ux|x)|$.
Supposons u hermitienne. Pour $z \in \mathcal{C}$

$$\begin{aligned} |z| &= \sup_{0 \leq \theta \leq 2\pi} |\Re(e^{i\theta} z)| \Rightarrow \|u\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} |(ux|y)| = \sup_{\|x\| \leq 1, \|y\| \leq 1, 0 \leq \theta \leq 2\pi} |\Re(e^{i\theta} (ux|y))| \\ &= \sup_{\|x'\| \leq 1, \|y\| \leq 1} |\Re(u(x')|y)| \leq \frac{1}{4} \sup_{\|x'\| \leq 1, \|y\| \leq 1} \left(\alpha \|x'+y\|^2 + \alpha \|x'-y\|^2 \right) \\ &\leq \frac{\alpha}{4} \sup_{\|x'\| \leq 1, \|y\| \leq 1} (2(\|x'\|^2 + \|y\|^2)) = \alpha \end{aligned}$$

on a donc $\|u\| \leq \alpha$ et finalement $\|u\| = \sup_{\|x\| \leq 1} |(ux|x)|$.

Corollaire 2.23.1 Si $u \in \mathcal{L}(E)$ u^*u est autoadjoint et $\|u^*u\| = \|u\|^2$.

Démonstration

$(u^*u)^* = u^*u$, il est donc bien autoadjoint.

$$\|u^*u\| = \sup_{\|x\| \leq 1} |(u^*ux|x)| = \sup_{\|x\| \leq 1} (ux|ux) = \sup_{\|x\| \leq 1} \|ux\|^2 = \|u\|^2.$$

La propriété 4 est donc démontrée.

2.8 Eléments propres d'un opérateur hermitien

Si $\lambda \in \mathcal{C}$, $E_\lambda(u) = \ker(u - \lambda I) = \{x \in E | ux = \lambda x\}$ est le sous espace propre pour la valeur propre λ si $E_\lambda(u)$ n'est pas réduit au vecteur nul. Dans ce cas un élément $x \in E_\lambda(u)$ est appelé vecteur propre associé à u (si $ux = \lambda x$ alors $|\lambda| \leq \|u\|$). En dimension finie l'ensemble Λ des valeurs propres de u est constitué des racines de $\det(u - \lambda I)$, il est non vide (le corps est \mathcal{C}) mais E n'est somme des $E_\lambda(u)$ que si u est diagonalisable.

En dimension infinie plus rien de cela ne subsiste, Λ peut être infini

$$(x_i)_{i \in \mathbb{N}} \rightarrow (\lambda_i x_i)_{i \in \mathbb{N}} \text{ de } l^2 \text{ dans lui même } \sup_{i \in \mathbb{N}} |\lambda_i| < +\infty$$

admet tous les λ_i comme valeurs propres, mais Λ peut être vide

$$E = L_m^2([0, 1]), u : f \rightarrow hf \text{ } h \text{ mesurable bornée } (u - \lambda I)f = 0 \Leftrightarrow hf = \lambda f$$

ceci signifie $f = 0$ p.p. dans $K = \{x \in [0, 1] | h(x) \neq \lambda\}$, si \overline{K} est négligeable $f = 0$ p.p. or si $h(x) = x$ $\{x \in [0, 1] | h(x) = \lambda\}$ est \emptyset si $\lambda \notin [0, 1]$ et $\{\lambda\}$ si $\lambda \in [0, 1]$, u n'a donc pas de valeur propre.

Théoreme 2.24 *Les valeurs propres d'un opérateur hermitien sont réelles et les sous espaces propres sont deux à deux orthogonaux.*

Démonstration

Si $ux = \lambda x$ $(ux|x) = \lambda \|x\|^2 = (x|ux) \in \mathbb{R} \Rightarrow \lambda \in \mathbb{R}$.

Soient x et y deux vecteurs propres associés respectivement aux valeurs propres distinctes λ et μ .

$$\lambda(x|y) = (\lambda x|y) = (ux|y) = (x|uy) = \mu(x|y) \text{ d'où } (x|y) = 0$$

. On va terminer en montrant que sous certaines un opérateur hermitien est diagonalisable.

Définition 2.9 *L'opérateur u sur E est dit compact si $\overline{u(B)}$ est compacte, où B est la boule unité fermée de E . Il revient au même de dire que pour toute suite (x_n) bornée, on peut extraire une sous suite (x_{n_k}) telle que $(u(x_{n_k}))$ converge.*

Théoreme 2.25 *Soit u un opérateur hermitien compact sur E . Pour $\lambda \neq 0$ $E_\lambda(u)$ est de dimension finie. L'ensemble Λ des valeurs propres de u est finie ou peut être rangé en une suite tendant vers zéro. De plus E est l'adhérence du sous espace engendré par tous les $E_\lambda(u)$, $\lambda \in \Lambda$.*

Démonstration

Si Λ est infini, soit $\delta > 0$. Supposons qu'il existe une infinité de valeurs propres telles que $|\lambda| \geq \delta$ ($\delta \leq \lambda \leq \|u\|$), il existe donc une suite de ces valeurs propres convergeant vers $\lambda \neq 0$ car la couronne $\{z \in \mathbb{C} | \delta \leq |z| \leq \|u\|\}$ est compacte dans \mathbb{C} . Soit x_n un vecteur propre associé à la valeur propre λ_n , les x_n sont linéairement indépendants, on appelle E_n le sous espace engendré par x_0, x_1, \dots, x_n . On a $u(E_n) \subset E_n$ et $(u - \lambda_n I)E_n \subset E_{n-1}$ $n \geq 1$, E_{n-1} étant sous espace strict de E_n il existe $y_n \in E_n$, $y_n \perp E_{n-1}$, $\|y_n\| = 1$.

$$\left\| \frac{y_n}{\lambda_n} \right\| = \frac{1}{|\lambda_n|} \leq \frac{1}{\delta},$$

u étant compact on peut extraire une sous suite pour laquelle la suite $(u(y_n/\lambda_n))$ converge. Ceci est contradictoire car

$$n < m, \quad u\left(\frac{y_n}{\lambda_n}\right) - u\left(\frac{y_m}{\lambda_m}\right) = y_n + \underbrace{\frac{1}{\lambda_n}(u - \lambda_n I)y_n - u\left(\frac{y_m}{\lambda_m}\right)}_{\in E_{n-1}}$$

cet élément est donc de norme ≥ 1 et ceci démontre que pour $\delta > 0$ il y a un nombre fini de valeurs propres de module $\geq \delta$.

L'ensemble des valeurs s'il est infini est donc dénombrable car réunion dénombrable d'ensembles finis et de plus d'après ce qui précède $\Lambda = (\lambda_n)$ peut être rangé

en une suite tendant vers zéro. Si $E_\lambda(u)$ était de dimension infinie, on pourrait choisir un système orthonormé dénombrable $(e_n)_{n \in \mathbb{N}}$ dans $E_\lambda(u)$, mais $u(e_n) = \lambda e_n$ et $\|e_n - e_m\| = \sqrt{2}$, ceci contredit la compacité de u , $E_\lambda(u)$ est donc de dimension finie. Montrons qu'il existe une valeur propre λ telle que $|\lambda| = \|u\|$.

D'après le théorème 2.23 il existe une suite (x_n) d'éléments de E telle que

$$\lim_{n \rightarrow \infty} |(ux_n|x_n)| = \|u\|$$

il existe une sous suite de (x_n) telle que $(ux_n|x_n)$ converge vers $\alpha = \|u\|$ ou $-\|u\|$ car $(ux_n|x_n) \in \mathbb{R}$, u étant hermitien.

$$0 \leq \|ux_n - \alpha x_n\|^2 - 2\alpha(ux_n|x_n) + \alpha^2\|x_n\|^2 \xrightarrow{n \rightarrow \infty} 0$$

u étant compact il existe une sous suite de (x_n) convergeant vers un élément $y \in E$ et ceci entraîne pour cette même sous suite $\lim_{n \rightarrow \infty} \alpha x_n = y$, si $u \neq 0$ $\alpha \neq 0$ et $\lim_{n \rightarrow \infty} x_n = y/\alpha = x$ et $ux = \alpha x$, α est donc valeur propre pour u . Soit $F = \overline{\cup_{\lambda \in \Lambda} E_\lambda}$ et $H = F^\perp$. H est u -invariant ($u(E) \subset E$) et u_H est un opérateur hermitien compact sans valeur propre, d'après ce qui précède ceci n'est possible que si $H = \{0\}$, on a donc $E = F \oplus H = F$.

Après avoir choisi un système orthonormé dans chaque $E_\lambda(u)$ on obtient par leur réunion un système orthonormé complet dans E constitué de vecteurs de u et quel que soit $\delta > 0$ il y a un nombre fini de λ_n , valeurs propres de u , de module supérieur ou égal à δ .

2.9 Exercices

Exercice 1

Soit $(F_n)_{n \in \mathbb{N}}$ une suite croissante de convexes fermés d'un espace de Hilbert E ,

$$F = \overline{\cup_{n \in \mathbb{N}} F_n}$$

a étant un point de E on appelle α_n et α les projections orthogonales de a sur F_n et F respectivement.

- 1) Montrer que la suite $(d_n)_{n \in \mathbb{N}}$ définie par $d_n = \|a - \alpha_n\|$ est une suite convergente dans \mathbb{R} .
- 2) En utilisant l'identité de la médiane, montrer que la suite $(\alpha_n)_{n \in \mathbb{N}}$ est une suite de Cauchy de E .
- 3) Montrer que la suite $(\alpha_n)_{n \in \mathbb{N}}$ converge vers α .

Exercice 2

- 1) Soit V un sous espace fermé de l'espace de Hilbert E . Montrer que $P_V^2 = P_V$ et que $P_V^* = P_V$.
- 2) Montrer réciproquement que si $T \in \mathcal{L}(E)$ vérifie $T^2 = T$ et $T^* = T$, T est l'opérateur de projection sur V , sous espace vectoriel fermé engendré par $T(E)$.

Exercice 3

Soient x_1, x_2, \dots, x_n n points d'un espace de Hilbert E . Le déterminant de Gram de ces n points est défini par

$$G(x_1, x_2, \dots, x_n) = |(x_i | x_j)|_{1 \leq i \leq n, 1 \leq j \leq n}$$

1) Montrer que $G(x_1, x_2, \dots, x_n) \geq 0$ et que $G(x_1, x_2, \dots, x_n) > 0$ équivaut à dire que les x_i sont linéairement indépendants. (Utiliser une base orthonormée dans le sous espace vectoriel engendré par les x_i).

2) Montrer que si les x_i sont linéairement indépendants et si L est le sous espace engendré par ces vecteurs, on a

$$d^2(x, F) = \frac{G(x, x_1, x_2, \dots, x_n)}{G(x_1, x_2, \dots, x_n)}.$$

Exercice 4

Calculer

$$\min_{a,b,c} \int_{-1}^1 (x^3 - a - bx - cx^2)^2 dx.$$

Exercice 5

1) Soit E un espace de Hilbert et M un sous espace vectoriel fermé de E . $x_0 \in E$, montrer que

$$\min (\|x_0 - x\| \mid x \in M) = \max (|(x_0 | y)| \mid y \in M^\perp, \|y\| = 1).$$

2) Calculer

$$\max_g \int_{-1}^1 x^3 g(x) dx \text{ où } g : \mathbb{R} \rightarrow \mathbb{R}$$

vérifie

$$\int_{-1}^1 g(x) dx = \int_{-1}^1 xg(x) dx = \int_{-1}^1 x^2 g(x) dx = 0, \int_{-1}^1 (g(x))^2 dx = 1.$$

Exercice 6

Calculer

$$\min_{a,b,c} \int_0^{+\infty} (x^3 - a - bx - cx^2)e^{-x} dx.$$

Exercice 7

On appelle E l'espace de Hilbert constitué par les fonctions mesurables à valeurs dans \mathbb{R} vérifiant :

$$\|f\|^2 = (f|f) = \int_0^{+\infty} (f(t))^2 e^{-t} dt < +\infty.$$

Le produit scalaire associé est défini par

$$(f|g) = \int_0^{+\infty} f(t)g(t)e^{-t} dt.$$

Le résultat suivant pourra être utilisé sans démonstration

$$\int_0^{+\infty} x^n e^{-x} dx = n!, \quad n \in \mathbb{N}^*.$$

1) On considère la suite

$$L_n(x) = \frac{1}{n!} e^x \frac{d^n}{dx^n} (e^{-x} x^n)$$

a) Vérifier, en utilisant la formule de Leibnitz que $L_n(x)$ est un polynôme de degré n .

b) Montrer que si f_n est la fonction $x \rightarrow x^n$ on a

$$\forall k \in \{0, 1, \dots, n-1\} \quad (f_k|L_n) = 0 \quad \text{et} \quad (f_n|L_n) = (-1)^n n!$$

NB : Il pourra être utile de remarquer que $x \rightarrow \frac{d^i}{dx^i} (e^{-x} x^n)$ s'annule en 0 si $i < n$ (justifiez le).

c) Déterminer le coefficient de x^n dans $L_n(x)$ et en déduire que $(L_n)_{n \in \mathbb{N}}$ est un système orthonormal de E .

2) a) Montrer que $\forall \lambda \geq 0$ la fonction $u_\lambda : x \rightarrow e^{-\lambda x}$ appartient à E .

b) Montrer que

$$(u_\lambda|L_n) = \frac{\lambda^n}{n!} \int_0^{+\infty} e^{-(\lambda+1)x} x^n dx = \frac{\lambda^n}{(\lambda+1)^{n+1}}$$

et calculer

$$\sum_{n=0}^{\infty} |(u_\lambda|L_n)|^2$$

Comparer ce dernier résultat avec $\|u_\lambda\|^2$.

3) Déterminer L_0, L_1, L_2 . On appelle F le sous espace vectoriel engendré par ces polynômes, déterminer $P_F(f_3)$, projection orthogonale de $f_3 : x \rightarrow x^3$ sur F .

Exercice 8

On appelle H l'espace de Hilbert des fonctions mesurables f de \mathbb{R} dans \mathbb{C} telles que

$$\int_{-\infty}^{+\infty} |f(t)|^2 e^{-t^2} dt < \infty$$

muni du produit scalaire

$$(f|g) = \int_{-\infty}^{+\infty} f(t)\overline{g(t)}e^{-t^2} dt$$

1) On définit

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

Déterminer H_0, H_1, H_2, H_3 .

2) Déterminer une relation entre H_n, H'_n, H_{n+1} . Démontrer par récurrence que H_n est un polynôme de degré n pour tout $n \in \mathbb{N}$. Quel est le coefficient de degré n de H_n ?

3) On appelle f_i la fonction monôme $x \rightarrow x^i$. Montrer, en effectuant n intégrations par parties, que

$$(f_k | H_n) = \begin{cases} 0 & \text{si } k < n \\ n! \sqrt{\pi} & \text{si } k = n \end{cases}$$

4) Déterminer a, b, c pour que l'intégrale

$$I(a, b, c) = \int_{-\infty}^{+\infty} (x^3 - ax^2 - bx - c)^2 e^{-x^2} dx$$

soit minimale.

5) On appelle $(P_n)_{n \in \mathbb{N}}$ la famille orthonormée construite à partir de la famille $(f_n)_{n \in \mathbb{N}}$ (question 3) par la méthode de Gram-Schmidt. Montrer qu'il existe un scalaire α tel que $P_n = \alpha H_n$ (on pourra écrire P_n comme combinaison linéaire de H_0, H_1, \dots, H_n).

Exercice 9

On appelle H l'espace complexe des fonctions f de \mathbb{R} dans \mathbb{C} telles que

$$\int_{-\infty}^{+\infty} |f(t)|^2 e^{-t^2} dt < \infty$$

muni du produit scalaire

$$(f | g) = \int_{-\infty}^{+\infty} f(t) \overline{g(t)} e^{-t^2} dt$$

On appelle (P_n) la suite orthonormée où P_n est un polynôme réel de degré n , dont le coefficient du terme de plus haut degré est positif. (c'est la suite obtenue par la méthode de Gram-Schmidt à partir de la base canonique de $\mathbb{R}[X]$). Expliciter P_0 et P_1 .

1) Soit Q_{n+1} un polynôme de degré $n+1$, montrer que l'on peut le mettre sous la forme

$$Q_{n+1} = \sum_{k=0}^{n+1} a_k P_k$$

En déduire qu'il existe des constantes $\alpha_n, \beta_n, \gamma_n$ telles que

$$\forall t \in \mathbb{R}, tP_n(t) = \alpha_n P_{n+1}(t) + \beta_n P_n(t) + \gamma_n P_{n-1}(t)$$

2) Montrer que $\beta_n = 0$ On admettra que $\alpha_n = \sqrt{\frac{n+1}{2}}$ et que $\gamma_n = \alpha_{n-1}$. Expliciter les polynômes P_2 et P_3 .

3) Déterminer a, b, c pour que l'intégrale

$$I(a, b, c) = \int_{-\infty}^{+\infty} (t^3 - (at^2 + bt + c))^2 e^{-t^2} dt$$

soit minimale. Calculer alors la valeur de cette intégrale.

Exercice 10

On considère un réel $\alpha > -\frac{1}{2}$. Toutes les fonctions considérées sont à valeurs dans \mathcal{C} . On pose $X = [-1, 1]$ et on note dx la mesure de Lebesgue sur X .

La fonction ρ définie sur $] -1, 1[$ par $\rho(x) = (1-x)^\alpha$ est intégrable sur X et on définit la mesure μ sur X par $d\mu(x) = \rho(x)dx$, c'est la mesure de base ou de densité ρ par rapport à la mesure de Lebesgue. Pour $s \in \mathbb{N}$ on désigne par x^s la fonction $x \rightarrow x^s$ sur X .

On notera E l'espace de Hilbert $L^2(X, \mu)$ pour le produit scalaire

$$(f|g) = \int_{-1}^1 f(x)\overline{g(x)}d\mu(x) = \int_{-1}^1 f(x)\overline{g(x)}\rho(x)dx$$

de norme associée $\|f\| = \sqrt{\int_{-1}^1 |f(x)|^2 \rho(x)dx}$

1) On définit une suite (P_n) de polynômes sur X par

$$P_0(x) = 1 \text{ et } P_n(x) = \frac{(-1)^n}{2^n(\alpha+1)\cdots(\alpha+n)}(1-x^2)^{-\alpha} \frac{d^n}{dx^n} \left((1-x^2)^{\alpha+n} \right) \text{ si } n > 0$$

Montrer que

a) P_n est un polynôme de degré n à coefficients réels.

b) $P_n(1) = 1$.

c) $(x^s|P_n) = 0$ pour $0 \leq s \leq n-1$.

2) En déduire que P_n est une famille orthogonale de E . On admettra que c'est une base Hilbertienne de E .

3) Pour $f \in E$ et $n \in \mathbb{N}$ on pose $\hat{f}(n) = (f|P_n)$. Montrer que

$$\sum_{n=0}^{\infty} \frac{|\hat{f}(n)|^2}{\|P_n\|^2} < \infty$$

et que la série

$$\sum_{n=0}^{\infty} \frac{\hat{f}(n)}{\|P_n\|^2} P_n$$

converge vers f dans $(E, \|\cdot\|)$.

Exercice 11

Soit E un espace de Hilbert et λ un réel strictement positif.

1) Soit $x \in E$, on pose

$$F(y) = \|y\|^2 + \lambda\|y-x\|^2, \quad y \in E$$

Montrer en utilisant le théorème de la médiane que

$$F(y) + F(z) - 2F\left(\frac{y+z}{2}\right) = \frac{1}{2}\|y-z\|^2 + \frac{1}{2}\lambda\|y-z\|^2$$

2) On pose $m(x) = \inf_{y \in E} F(y)$.

a) En utilisant la question précédente, montrer qu'il existe au plus un point $y \in E$ tel que $F(y) = m(x)$.

b) Soit (y_n) une suite de points de E tels que $\lim_{n \rightarrow \infty} F(y_n) = m(x)$. (L'existence de cette suite résulte de la définition de la borne inférieure). Montrer que (y_n) est une suite de Cauchy et que si y est sa limite $F(y) = m(x)$. On notera $y = u(x)$ ce point, on a donc $F(u(x)) = m(x)$.

3) En utilisant l'inégalité $F(u(x) + ty) \geq F(u(x))$ vérifiée pour tout $y \in E$ et tout $t \in \mathbb{R}$ montrer que

$$\operatorname{Re}((u(x)|y) + \lambda(u(x) - x|y)) = 0, \text{ en déduire en remplaçant } y \text{ par } iy$$

$$\operatorname{Im}((u(x)|y) + \lambda(u(x) - x|y)) = 0 \text{ puis que } (u(x)|y) = \frac{\lambda}{1 + \lambda}(x|y)$$

4) Déduire de la question précédente que u est linéaire, continue (on trouvera une majoration de $\|u\|$ en faisant $y = u(x)$) et que

$$\forall x, y \in E \quad (u(x)|y) = (x|u(y))$$

Exercice 12

On appelle μ la mesure de densité $f(x) = \frac{1}{x}$ par rapport à la mesure de Lebesgue sur $]0, 1]$ et E est l'espace de Hilbert réel $L^2_\mu(]0, 1])$, c'est à dire $E = \left\{ f :]0, 1] \rightarrow \mathbb{R} \mid \int_0^1 |f(x)|^2 \frac{dx}{x} < +\infty \right\}$.

1) Montrer que tout polynôme P tel que $P(0) = 0$ est élément de E .

2) Le procédé d'orthogonalisation de Gram-Schmidt, appliqué à la suite x^n fournit les polynômes $P_1, P_2, P_3, \dots, P_n, \dots$ obtenus de la façon suivante

a) P_k est de degré k et le coefficient de x^k est 1.

b) $(P_i|P_j) = 0$ pour $i \neq j$.

Montrer que $(P_n|P_n) = (P_n|x^n)$ Calculer explicitement P_1, P_2, P_3 .

3) Calculer $\|P_1\|$ et $\|P_2\|$.

4) On considère $I(a, b) = \|\sqrt{x} - aP_1 - bP_2\|$. Déterminer les nombres réels α et β tels que $I(\alpha, \beta) = \inf_{a, b \in \mathbb{R}} I(a, b) = I$.

5) Calculer la valeur de I .

Exercice 13

On considère un nombre réel tel que $0 < a < 1$.

1) Calculer les sommes

$$S_1(x) = \sum_{n=0}^{\infty} a^n e^{inx} \text{ et } S_2(x) = \sum_{n=1}^{\infty} a^n e^{-inx}$$

En déduire le développement en série de Fourier de $f(x) = S_1(x) + S_2(x)$

2) En déduire la valeur de l'intégrale

$$I_n = \int_0^{2\pi} \frac{\cos nt}{1 - 2a \cos t + a^2} dt$$

Exercice 14

1) Développer en série de Fourier sur l'intervalle $[-1,1]$ la fonction f définie par

$$f(x) = \begin{cases} -1 - x & \text{si } -1 \leq x < 0 \\ 1 - x & \text{si } 0 < x \leq 1 \end{cases}$$

2) En déduire la valeur des sommes

$$S_1 = \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} \text{ et } S_2 = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

Exercice 15

On considère la fonction impaire et 2π -périodique f définie par $f(x) = x(\pi - x)$ pour $x \in [0, \pi]$. Montrer que

$$\forall x \in [0, \pi] \quad x(\pi - x) = \frac{8}{\pi} \sum_{p=0}^{\infty} \frac{\sin(2p+1)x}{(2p+1)^3}$$

En déduire les valeurs de

$$\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)^3} \text{ et } \sum_{p=0}^{\infty} \frac{1}{(2p+1)^6}$$

Exercice 16

1) On considère

$$S_{n,p}(x) = \sum_{k=n}^{n+p} \frac{\sin kx}{k} \text{ et } V_{n,l}(x) = \sum_{k=n}^{n+l} \sin kx$$

Montrer que

$$\forall x \neq k\pi \quad |V_{n,l}(x)| \leq \frac{1}{\sin(x/2)} \text{ et } |S_{n,p}(x)| \leq \frac{|a_n|}{|\sin(x/2)|}$$

En déduire que la série $\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ converge uniformément sur tout intervalle $[2k\pi + \epsilon, 2(k+1)\pi - \epsilon]$ où $0 < \epsilon < \pi$

2) Montrer que la série

$$S(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$$

est continue sur \mathbb{R} et dérivable sur $\mathbb{R} - 2\pi\mathbb{Z}$

3) on choisit $0 < a < \pi$ et on considère la fonction 2π -périodique

$$f(x) = \begin{cases} 1 & \text{si } -a \leq x \leq a \\ 0 & \text{si } x \in [-\pi, -a[\cup]a, \pi] \end{cases}$$

Déterminer la série de Fourier de f et montrer que $\forall x \in]0, 2\pi[\quad S'(x) = \frac{x}{2} - \frac{\pi}{2}$.

Montrer que $\int_0^\pi S(x) dx = 0$ en déduire $S(x)$ et $\sum_{n=1}^{\infty} \frac{1}{n^2}$ et $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$.

Exercice 17

L'espace Hilbertien E est l'ensemble des fonctions mesurables sur \mathbb{R}^+ telles que $|f|^2$ soit intégrable pour la mesure de Lebesgue.

1) On définit l'opérateur u sur E par

$$uf(x) = f\chi_{[a, +\infty[}(x), \quad f \in E, \quad a \geq 0$$

Montrer que u est une projection.

2) On définit l'opérateur V sur E par

$$Vf(x) = f(x-a)\chi_{[a, +\infty[}(x), \quad f \in E, \quad a \geq 0$$

Montrer que V est une isométrie.

3) Déterminer V^* adjoint de V et calculer V^*V .

4) Soit $W = VV^*$. Déterminer W^2 . Calculer Wf pour $f \in E$.

Exercice 18

$E = l^2$, $A = (e_n)_{n \in \mathbb{N}^*}$ base Hilbertienne canonique de E . On définit l'opérateur u sur E par $ue_n = \lambda_n e_n$, $n \in \mathbb{N}^*$ où $(\lambda_n)_{n \in \mathbb{N}^*}$ est une suite vérifiant

$$\forall n \in \mathbb{N}^*, \lambda_n \neq a, \quad \lim_{n \rightarrow \infty} \lambda_n = a$$

Déterminer les valeurs propres de u .

Exercice 19

1) $E = L^2([0, 1])$, K fonction continue sur $[0, 1] \times [0, 1]$, pour $f \in E$ on définit

$$uf(x) = \int_0^1 K(x, y)f(y)dy$$

Montrer que u est un opérateur Hermitien sur E et que sa norme est majorée par

$$\sqrt{\iint_{[0,1] \times [0,1]} |K(x, y)|^2 dx dy}$$

Déterminer les valeurs propres et les vecteurs propres associés dans les deux cas suivants.

a) $K(x, y) = xy$.

b) $K(x, y) = e^{x+y}$.

2) Résoudre dans les deux cas précédents l'équation intégrale

$$uf(x) - \lambda f(x) = g(x)$$

avec

a) $g(x) = 3x - 2$.

b) $g(x) = x$.

Exercice 20

Soit E l'espace de Hilbert des fonctions réelles de carré intégrable. Soit A l'opérateur défini sur E par

$$Ax(t) = \int_0^{2\pi} x(s) \cos(t - s) ds$$

1) Montrer que A est Hermitien.

2) Donner l'expression de la série de Fourier de Ax en fonction de celle de x .

3) Calculer $\|A\|^2$.

4) Déterminer les valeurs propres et les vecteurs propres de A .

Deuxième partie

PROBABILITES

Chapitre 3

NOTIONS FONDAMENTALES

3.1 Ensemble fondamental. Evènements.

3.1.1 Exemples

Exemple 1

On lance un dé à 6 faces numérotées de 1 à 6. On cherche à décrire cette expérience aléatoire, c'est à dire due au hasard. Chaque réalisation possible sera notée sous forme de singleton, $\{1\}, \{2\}, \dots, \{6\}$. L'ensemble des réalisations sera notée $\Omega = \{1, 2, 3, 4, 5, 6\}$ et appelé ensemble des états ou ensemble fondamental. Un observateur peut s'intéresser à un aspect partiel de l'expérience, par exemple au cas où le nombre qui sort est pair. On dira que l'évènement "nombre pair" est réalisé si le nombre obtenu est élément du sous ensemble $\{2, 4, 6\}$. Les évènements seront décrits par des sous ensembles de Ω . Les réalisations seront appelées évènements élémentaires ou éventualités.

Exemple 2

On lance 2 pièces de monnaie, les éventualités pourront être notées $(P, P), (P, F), (F, P), (F, F)$ si les 2 pièces peuvent être distinguées. L'évènement "On obtient au moins un pile" sera représenté par le sous ensemble $\{(P, F), (F, P), (P, P)\}$.

Exemple 3

On lance une pièce de monnaie et on s'arrête lorsqu'on obtient pile pour la première fois. On peut choisir

$$\Omega = \{P, FP, FFP, \dots, FF \dots FP, \dots\}$$

c'est un ensemble infini. L'évènement "on a joué au plus 3 fois" s'écrit $\{P, FP, FFP\}$.

Exemple 4

La physique conduit à considérer dans certains cas le déplacement d'une particule comme un phénomène aléatoire. Si le déplacement se fait dans le plan, l'ensemble Ω des positions possibles de la particule pourra être le plan ou une

partie de celui-ci. Les évènements seront des parties du plan, mais pour des raisons mathématiques il s'agira d'un sous ensemble strict de l'ensemble des parties du plan. L'ensemble des évènements sera une tribu de parties de \mathbb{R}^2 , par exemple $\mathcal{B}(\mathbb{R}^2)$ qui est la tribu borélienne de \mathbb{R}^2 ou encore la plus petite tribu contenant les ouverts de \mathbb{R}^2 .

3.1.2 Définition

Définition 3.1 *Le couple (Ω, \mathcal{E}) constitué par un ensemble Ω et une tribu de parties \mathcal{E} de Ω est appelé espace probabilisable. Les éléments de Ω sont appelés éventualités ou évènements élémentaires. Les éléments de \mathcal{E} sont appelés évènements.*

3.2 Probabilité. Espace Probabilisé

3.2.1 Définition

Définition 3.2 *Une probabilité P sur un espace probabilisable (Ω, \mathcal{E}) est une mesure sur cet espace (qui est mesurable) telle que $P(\Omega) = 1$.*

Exemples

L'ensemble Ω est fini ou infini dénombrable.

Dans ce cas on choisit en général $\mathcal{E} = \mathcal{P}(\Omega)$. Une probabilité sur Ω est définie par la donnée d'une famille

$\{P(x) = \alpha_x, x \in \Omega\}$ de nombres réels vérifiant la condition

$$\sum_{x \in \Omega} \alpha_x = 1.$$

Dans le cas particulier où Ω est fini on appelle probabilité uniforme sur Ω la probabilité correspondant à $\alpha_x = 1/\text{Card}\Omega$.

Pour $A \subset \Omega, P(A) = \text{Card}A/\text{Card}\Omega$. cette probabilité correspond au tirage d'un point au hasard dans Ω .

Exemple

On lance n fois une pièce équilibrée. On prend

$\Omega = \{P, F\}^n, \alpha_x = 1/2^n$. L'évènement $A =$ "On obtient 3 piles en n lancers" a pour probabilité $P(A) = C_n^3/2^n$ pour $n \geq 3$.

L'ensemble (Ω, \mathcal{E}) est $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

On considère une fonction mesurable positive h vérifiant $\int_{-\infty}^{+\infty} h(x)dx = 1$. Pour tout élément $A \in \mathcal{B}(\mathbb{R})$ on définit une probabilité par $P(A) = \int_A h(x)dx$.

3.2.2 Propriétés des probabilités

$$\forall A \in \mathcal{E} P(\bar{A}) = 1 - P(A) .$$

$$.P(\emptyset) = 0.$$

$$\forall A, B \in \mathcal{E} A \subset B \Rightarrow P(A) \leq P(B).$$

$$\forall A, B \in \mathcal{E} P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Pour démontrer ce dernier résultat il suffit d'écrire

$$A \cup B = (A - A \cap B) \cup (A \cap B) \cup (B - A \cap B) \text{ et de remarquer que si}$$

$D, C \in \mathcal{E}, C \subset D$ on a $P(C - D) = P(C) - P(D)$. Ce résultat admet la généralisation suivante, dont la démonstration se fait par récurrence.

Théoreme 3.1 Pour n évènements A_1, A_2, \dots, A_n on a

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k-1} S_n^k$$

avec

$$S_n^k = \sum_{1 \leq i_1 < i_2 < \dots < i_k} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$$

Exemple.

On dispose de 2 jeux de 52 cartes et on tire une carte de chaque jeu. Quelle est la probabilité d'avoir tiré au moins un roi ?

On choisit $\Omega = \{(a, b) | a \text{ carte du premier jeu}, b \text{ carte du deuxième jeu}\}$, $\mathcal{E} = \mathcal{P}(\Omega)$ et P = probabilité uniforme sur Ω .

Première méthode

Définissons les évènements suivants :

A_1 = "On tire un roi du premier jeu et une carte autre qu'un roi du deuxième jeu".

A_2 = "On tire un roi du deuxième jeu et une carte autre qu'un roi du premier jeu".

A_3 = "On tire un roi du premier jeu et un roi du deuxième jeu".

$P(A_1) = P(A_2) = (4.48)/52^2 = 192/2704$ $P(A_3) = 16/2704$. $A =$ "au moins un roi" = $A_1 \cup A_2 \cup A_3$, ces 3 événements étant 2 à 2 disjoints on a $P(A) = 400/2704$.

Deuxième méthode

$$P(\bar{A}) = P(\text{"aucun roi"}) = 48^2/2704 = 2304/2704$$

$$P(A) = 1 - P(\bar{A}) = 400/2704.$$

3.3 Probabilité conditionnelle.

3.3.1 Exemple.

Une entreprise achète 1000 pièces détachées à un fournisseur A_1 dont 5 pour cent sont défectueuses et 4000 à un fournisseur A_2 dont 3 pour cent sont défectueuses. Une pièce choisie au hasard est défectueuse, quelle est la probabilité qu'elle provienne du fournisseur A_1 ?

Définissons les évènements suivants :

Ω = "ensemble des pièces détachées".

D = "ensemble des pièces défectueuses".

A_i = "ensemble des pièces provenant du fournisseur A_i , $i = 1, 2$."

L'ensemble D devient l'évènement certain.

$\text{Card } D = 0,05.1000 + 0,03.4000 = 170$. A_1 contenant 50 pièces défectueuses, la probabilité cherchée est

$$50/170 = \frac{50/5000}{170/5000} = P(A_1 \cap D)/P(D).$$

3.3.2 Définition.

Définition 3.3 Soit (Ω, \mathcal{E}, P) un espace probabilisé et A un évènement de probabilité non nulle, $F \in \mathcal{E}$, on appelle probabilité conditionnelle de F par rapport à A ou probabilité conditionnelle de F sachant A le nombre

$$P(F|A) = \frac{P(F \cap A)}{P(A)}.$$

Conséquence. Sous les hypothèses précédentes on a $P(F \cap A) = P(F|A)P(A)$. Ce résultat peut être généralisé par récurrence à n évènements A_1, A_2, \dots, A_n .

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

Exemples.

On tire successivement et sans remise 3 cartes d'un jeu de 52 cartes. Quelle est la probabilité d'avoir tiré 3 rois ?

Posons $A_i =$ "on tire un roi au i ème tirage", $i = 1, 2, 3$.

La probabilité cherchée est

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) = 4/52 \cdot 3/51 \cdot 2/50.$$

On considère la composition en garçons et filles d'une famille de 2 enfants. Un ensemble fondamental est $\Omega = \{(G, G), (G, F), (F, G), (F, F)\}$. Calculons la probabilité des évènements suivants :

$H =$ "la famille a un garçon".

$A =$ "l'aîné est un garçon".

$B =$ "les 2 enfants sont des garçons".

$$P(B|H) = \frac{P(B \cap H)}{P(H)} = 1/3, \quad P(B|A) = \frac{P(B \cap A)}{P(A)} = 1/2.$$

Cet exemple montre l'importance de bien poser le problème et de se méfier de l'intuition dans les calculs de probabilité conditionnelle.

3.3.3 Théorème de Bayes

Théorème 3.2 On considère (A_1, A_2, \dots, A_n) n évènements constituant une partition de Ω tels que $P(A_i) > 0$, $i = 1, 2, \dots, n$. Pour tout évènement A on a

$$P(A) = \sum_{i=1}^n P(A|A_i)P(A_i) \quad \text{et si } P(A) \neq 0$$

$$P(A_j|A) = \frac{P(A|A_j)P(A_j)}{\sum_{i=1}^n P(A|A_i)P(A_i)}.$$

Ce résultat est immédiat en remarquant que

$A = \cup_{i=1}^n (A \cap A_i)$ et que ces évènements sont deux à deux disjoints.

Exemple.

Une urne A_1 contient 70 pour cent de boules blanches et une urne A_2 80 pour

cent de boules blanches. L'urne A_1 contient trois fois plus de boules que l'urne A_2 . On mélange les contenus de A_1 et de A_2 et on choisit une boule au hasard, elle est blanche, quelle est la probabilité qu'elle provienne de A_1 ?

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} = \frac{0,7 \cdot 0,75}{0,7 \cdot 0,75 + 0,8 \cdot 0,25} = 21/29.$$

3.3.4 Evènements indépendants.

Définition 3.4 Soit (Ω, \mathcal{E}, P) un espace probabilisé et $A, B \in \mathcal{E}$. Ces 2 évènements sont indépendants si $P(A \cap B) = P(A)P(B)$.

Conséquence.

Dans le cas où ces probabilités sont définies on a $P(A|B) = P(A)$, $P(B|A) = P(B)$ quand A et B sont indépendants. Autrement dit la réalisation de l'un des évènements n'influe pas sur la probabilité qu'a l'autre de se réaliser.

Définition 3.5 Deux tribus \mathcal{A} et \mathcal{B} , sous ensembles de \mathcal{E} , sont indépendantes si

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{B} P(A \cap B) = P(A)P(B)$$

Théorème 3.3 L'indépendance des évènements A et B entraîne l'indépendance de \bar{A} et B , de A et \bar{B} , de \bar{A} et \bar{B} , c'est à dire des tribus $\mathcal{A} = \{\emptyset, A, \bar{A}, \Omega\}$ et $\mathcal{B} = \{\emptyset, B, \bar{B}, \Omega\}$ engendrées par A et B respectivement.

Montrons l'indépendance de \bar{A} et B , les autres cas se traitant de la même manière.

$$(\bar{A} \cup A) \cap B = (\bar{A} \cap B) \cup (A \cap B) = B.$$

$$P(B) = P(\bar{A} \cap B) + P(A \cap B).$$

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) = p(B) - p(A)P(B).$$

$$P(\bar{A} \cap B) = P(B)(1 - P(A)) = P(B)P(\bar{A}).$$

Définition 3.6 Les évènements A_1, A_2, \dots, A_n sont indépendants si les tribus engendrées respectivement par A_1, A_2, \dots, A_n sont indépendantes. Il revient au même de dire

$$\forall k \in \{1, 2, \dots, n\}, P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j})$$

pour toute suite $1 \leq i_1 < i_2 < \dots < i_k \leq n$.

Chapitre 4

VARIABLES ALEATOIRES DISCRETES

4.1 Loi de probabilité.

4.1.1 Mesure image et mesure définie par une densité

Mesure image.

Définition 4.1 *On considère un espace mesuré (X, \mathcal{A}, μ) et un espace mesurable (Y, \mathcal{B}) , h étant une application mesurable de X vers Y , on appelle mesure image de μ par h la mesure notée $\nu = h(\mu)$ définie sur (Y, \mathcal{B}) par $\nu(B) = \mu(h^{-1}(B))$, $B \in \mathcal{B}$.*

Théorème 4.1 *Pour toute fonction f positive et ν -mesurable sur Y on a*

$$\int_Y f d\nu = \int_X (f \circ h) d\mu$$

Démonstration

Le théorème est vrai pour les fonctions indicatrices d'ensembles χ_B avec $B \in \mathcal{B}$, donc également pour toute fonction étagée, puis pour toute fonction mesurable positive en utilisant une suite croissante de fonctions étagées convergeant vers cette fonction.

Corollaire 4.1.1 *Une fonction f réelle ou complexe \mathcal{B} -mesurable sur Y est ν -intégrable si et seulement si $f \circ h$ est μ -intégrable et dans ce cas*

$$\int_Y f d\nu = \int_X (f \circ h) d\mu.$$

Démonstration

Le résultat est vrai pour les fonctions positives d'après le théorème précédent, si f est à valeurs dans \mathbb{R} on applique le résultat à f^+ et f^- , si f est à valeurs dans \mathcal{C} on utilise $Re f$ et $Im f$.

Mesures de base donnée

Définition 4.2 Soit (X, \mathcal{A}, μ) un espace mesuré et g une fonction mesurable positive sur X . On appelle mesure de base (ou de poids) g par rapport à la mesure μ , la fonction d'ensembles définie sur \mathcal{A} par $\nu(E) = \int_E g d\mu$, $E \in \mathcal{A}$.

La fonction ν est bien une mesure car si les ensembles mesurables A_n sont deux à deux disjoints on obtient

$$\nu(\cup_n E_n) = \int (\chi_{\cup_n E_n}) g d\mu = \int \left(\sum_n \chi_{E_n} \right) g d\mu = \sum_n \int_{E_n} g d\mu = \sum_n \nu(E_n)$$

L'avant dernière égalité est une conséquence du théorème de Beppo-Levi. La mesure ν est notée $\nu = g\mu$.

Théorème 4.2 On a pour toute fonction mesurable positive f

$$\int f d\nu = \int f g d\mu$$

où ν est la mesure de base g par rapport à μ .

La démonstration est similaire à celle du théorème précédent.

Corollaire 4.2.1 La fonction f est ν -intégrable si et seulement si $f g$ est μ -intégrable et

$$\int f d\nu = \int f g d\mu$$

Remarque.

Dans le cas d'un ensemble mesurable E tel que $\mu(E) = 0$ on a alors $\nu(E) = 0$. On exprime ceci en disant que ν est absolument continue par rapport à μ .

4.1.2 Variable aléatoire

Définition 4.3 Soit (Ω, \mathcal{E}, P) un espace probabilisé et (E, \mathcal{B}) un espace mesurable. Une application mesurable $X : \Omega \rightarrow E$ est appelée variable aléatoire. Si E est dénombrable (en général $E = \mathbb{Z}$ ou $E = \mathbb{N}$) on dit que X est une variable aléatoire discrète. Dans ce cas on prend en général $\mathcal{B} = \mathcal{P}(E)$.

Notations.

$I \subset E$, $X^{-1}(I) = \{\omega \in \Omega | X(\omega) \in I\}$ est noté $[X \in I]$
 $n \in E$, $X^{-1}(\{n\}) = \{\omega \in \Omega | X(\omega) = n\}$ est noté $[X = n]$.

4.1.3 Loi de probabilité.

Définition 4.4 La loi de probabilité de la variable aléatoire discrète X est la mesure image de P par X . C'est une probabilité sur (E, \mathcal{B}) notée p_X . L'espace (E, \mathcal{B}, p_X) est un espace probabilisé.

Remarque.

$B \in \mathcal{B}$, $p_X(B) = P(X^{-1}(B)) = P([X \in B])$.

Dans le cas où $\mathcal{B} = \mathcal{P}(E)$ et $n \in E$, $p_X(\{n\})$ est noté $p_X(n) = P([X = n])$.

Dans ce cas

$B \subset E$, $p_X(B) = \sum_{n \in B} p_X(n)$.

Exemple.

On lance 3 dés de couleurs différentes et on appelle X le nombre de 6 obtenus. Modélisons cette expérience.

$\Omega = \{1, 2, 3, 4, 5, 6\}^3$, $\mathcal{E} = \mathcal{P}(\Omega)$, $P =$ probabilité uniforme sur Ω

$X((a, b, c)) =$ "nombre de six parmi a, b, c ", $E = \{0, 1, 2, 3\}$, $\mathcal{B} = \mathcal{P}(E)$

$p_X(0) = 125/216$, $p_X(1) = 75/216$, $p_X(2) = 15/216$, $p_X(3) = 1/216$.

4.1.4 Fonction de répartition.

Définition 4.5 La fonction de répartition de la variable aléatoire X est la fonction définie par $F_X(x) = P([X < x]) = P(\{\omega \in \Omega | X(\omega) < x\})$.

Propriétés.

- $\lim_{x \rightarrow -\infty} F_X(x) = 0$, $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
- F_X est croissante.
- $\cup_{n=1}^{\infty} [X < x - 1/n] = [X < x]$ entraîne $\lim_{n \rightarrow \infty} F_X(x - 1/n) = F_X(x)$, la fonction F_X est donc continue à gauche en tout point.
- En faisant un raisonnement analogue au précédent on obtient $\lim_{n \rightarrow \infty} F_X(x + 1/n) = F_X(x) + P(\{x\})$. La fonction F_X est continue à droite en x si et seulement si $P(\{x\}) = 0$.

Exercice.

Représenter graphiquement la fonction de répartition de la variable aléatoire définie dans l'exemple précédent.

4.2 Espérance mathématique. Variance.

4.2.1 Espérance mathématique.

Définition 4.6 L'espérance mathématique de la variable aléatoire X est le nombre $E(X) = \int_{\Omega} X dP$ si X est intégrable.

Conséquence.

p_X étant la loi image de P par X on a $E(X) = \int_E x dp_X(x)$. Si X est une variable aléatoire discrète, à valeurs dans \mathbb{Z} ou \mathbb{N} on a $E(X) = \sum_{x \in E} x p_X(x)$, ceci correspond à la moyenne pondérée (ou barycentre) des valeurs prises par X .

4.2.2 Fonction génératrice.

Nous allons dans ce paragraphe introduire un outil qui facilitera le calcul de l'espérance mathématique et de la variance pour une variable aléatoire discrète.

Définition 4.7 La variable aléatoire X étant à valeurs dans \mathbb{N} , la fonction génératrice de X est la série entière

$$G_X(z) = \sum_{k \in E} p_X(k)z^k$$

le rayon de convergence de cette série est supérieur ou égal à 1 car $G_X(1) = 1$. Si $R > 1$ $G'_X(1) = \sum_{k \in E} kp_X(k) = E(X)$.

Exemple.

$E = \mathbb{N}^*$, $p_X(k) = q^{k-1}p$, ($p + q = 1$).

$$G_X(z) = \sum_{k=1}^{\infty} q^{k-1}pz^k = pz \sum_{k=1}^{\infty} (qz)^{k-1} = \frac{pz}{1-qz}$$

$$E(X) = G'_X(1) = \frac{p}{(1-q)^2}.$$

4.2.3 Propriétés de l'espérance mathématique.

— Si X et Y sont deux variables aléatoires on a

$$E(X + Y) = E(X) + E(Y), \quad E(aX) = aE(X) \quad (a \in \mathbb{R})$$

— E et F étant \mathbb{Z} ou \mathbb{N} , $g : E \rightarrow F$ une application et X une variable aléatoire discrète on a

$$E(g \circ X) = E(g(X)) = \sum_{k \in E} g(k)p_X(k)$$

Le premier résultat provient de la linéarité de l'intégrale et le deuxième des égalités

$$E(g(X)) = \int_{\Omega} g(X)dP = \int_E g(x)dp_X(x)$$

car p_X est la mesure image de P par X .

Exemples.

$$E(X^n) = \sum_{k \in E} k^n p_X(k).$$

$$E(e^X) = \sum_{k \in E} e^k p_X(k).$$

4.2.4 Variance.

Définition 4.8 Une variable aléatoire X vérifiant $E(X^2) < +\infty$ on définit la variance de X par $\text{Var}(X) = E((X - E(X))^2)$. Cette quantité mesure la dispersion (en fait le carré de la dispersion) moyenne des valeurs de X par rapport à l'espérance mathématique $E(X)$. L'écart type de X est la quantité $\sigma(X) = \sqrt{\text{var}(X)}$.

La fonction génératrice peut permettre de calculer la variance car $\phi_X''(1) = \sum_{k \geq 2} k(k-1) = E(X(X-1)) = E(X^2) - E(X)$. Ceci suppose que le rayon de convergence R soit supérieur strictement à 1.

Covariance de deux variables aléatoires.

Définition 4.9 $X, Y \in L^2(\Omega, \mathcal{E}, P) \Rightarrow XY \in L^1(\Omega, \mathcal{E}, P)$. Ceci permet de définir la covariance des variables aléatoires X et Y par

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

L'inégalité de Cauchy-Schwarz entraîne

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y).$$

Définition 4.10 Le coefficient de corrélation de deux variables aléatoires X et Y qui ne sont pas presque sûrement constantes est

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}, \quad |\rho(X, Y)| \leq 1$$

Ce coefficient mesure le degré de dépendance linéaire de X et Y .

Théorème 4.3 On considère deux variables aléatoires X et Y telles que $E(X^2) < +\infty$ et $E(Y^2) < +\infty$.

1. $\text{Var}(X) = 0 \Leftrightarrow X$ est constante P presque sûrement.
2. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$.
3. $\rho(X, Y) = \pm 1 \Leftrightarrow (\exists (a, b, c) \in \mathbb{R}^3 - (0, 0, 0) | aX + bY + c = 0 \text{ P p.s.})$.

4.3 Lois de probabilités discrètes usuelles

4.3.1 Loi de Bernoulli.

Une épreuve de Bernoulli est une expérience aléatoire à deux issues, par exemple jeu de pile ou face, tirer un roi ou une autre carte dans un jeu de carte... Ces deux possibilités sont en général notées 0 et 1 et appelées échec ou succès. La variable aléatoire $X : \Omega \rightarrow \{0, 1\}$ est appelée variable de Bernoulli. Sa loi est définie par $p = P([X = 0])$, $q = 1 - p = P([X = 1])$. On dit que p est le paramètre associé à X .

$$E(X) = p, \quad \text{Var}(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p) = pq.$$

4.3.2 Loi Binomiale.

On répète n épreuves de Bernoulli indépendantes de paramètre p , X est le nombre de 1 obtenus ($0 \leq X \leq n$). On appelle A l'évènement "On obtient 1". La loi suivie par X est appelée loi binomiale de paramètres n , p et notée $B(n, p)$.

L'écriture $X \sim B(n, p)$ signifie que X suit la loi binomiale $B(n, p)$.

$p_X(k) = P([X = k]) = P(\text{“réaliser } k \text{ fois } A) = C_n^k p^k q^{n-k}$
 car les répétitions sont indépendantes. $k \in \{0, 1, \dots, n\}$.

$$\begin{aligned} \phi_X(z) &= \sum_{k=0}^n C_n^k p^k q^{n-k} z^k = (pz + q)^n. \\ E(X) &= \phi'_X(1) = np, \quad E(X^2) = \phi''_X(1) + E(X) = n(n-1)p^2 + np, \\ \text{Var}(X) &= E(X^2) - (E(X))^2 = np(1-p) = npq. \end{aligned}$$

4.3.3 Loi géométrique.

On répète une épreuve de Bernoulli, de paramètre p et on s'arrête dès que l'on obtient 1. On appelle X le nombre de répétitions nécessaires. La loi de X est appelée loi géométrique de paramètre p . Les valeurs prises par X sont les éléments de \mathbb{N}^* .

$k \in \mathbb{N}^*$, $p_X(k) = q^{k-1}p$
 car ceci revient à faire $k-1$ répétitions de \bar{A} la k -ième répétitions donnant A .

On a montré dans un exemple précédent que
 $\phi_X(z) = pz/(1-qz)$ d'où
 $\phi'_X(z) = p/(1-qz)^2$, $\phi''_X(z) = 2pq/(1-qz)^3$.
 $E(X) = \phi'_X(1) = 1/p$, $E(X^2) = \phi''_X(1) + E(X) = 2q/p^2 + 1/p$
 $\text{Var}(X) = q/p^2$.

4.3.4 Loi hypergéométrique.

On choisit n objets parmi $N \geq n$ objets de deux types notés I et II. X représente le nombre d'objets de type I parmi les n objets. Il existe d objets de type I.

$$p_X(k) = \frac{C_n^k C_{N-d}^{n-k}}{C_N^n}, \quad \max(0, n - (N - d)) \leq k \leq \min(n, d)$$

4.3.5 Loi de Poisson.

C'est une loi utile pour l'étude des phénomènes de files d'attente. L'ensemble des valeurs prises par X est \mathbb{N} et

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}, \quad \lambda > 0$$

$\phi_X(z) = e^{-\lambda(z-1)}$ d'où $E(X) = \text{Var}(X) = \lambda$.

4.4 Couple de variables aléatoires.

Les variables aléatoires discrètes X et Y étant définies sur (Ω, \mathcal{E}, P) et à valeurs dans les ensembles dénombrables E et F respectivement, on définit une variable aléatoire discrète $Z = (X, Y)$ à valeurs dans $E \times F$ en posant $Z(\omega) = (X(\omega), Y(\omega))$. On a en effet
 $\{\omega \in \Omega | Z(\omega) = (x, y)\} = \{\omega \in \Omega | X(\omega) = x\} \cap \{\omega \in \Omega | Y(\omega) = y\} \in \mathcal{E}$.

Définition 4.11 La loi de probabilité de Z est appelée loi conjointe du couple (X, Y) et les lois des variables X et Y sont appelées lois marginales du couple de variables aléatoires.

Il est possible de connaître les lois marginales lorsque l'on connaît la loi conjointe.

Théorème 4.4

$$p_X(x) = \sum_{y \in F} p_Z(x, y), \quad x \in E, \quad p_Y(y) = \sum_{x \in E} p_Z(x, y), \quad y \in F$$

Ce résultat est obtenu en remarquant que

$$p_X(x) = p_Z(\{x\} \times F) = \sum_{x=x} \sum_{y \in F} p_Z(x, y), \quad p_Y(y) = p_Z(E \times \{y\}) = \sum_{x \in E} \sum_{y=y} p_Z(x, y)$$

X et Y étant deux variables aléatoires discrètes, $X^{-1}(\mathcal{P}(E))$ et $Y^{-1}(\mathcal{P}(F))$ sont deux sous tribus de \mathcal{E} appelées respectivement tribus engendrées par X et par Y .

Définition 4.12 Les variables aléatoires X_1, X_2, \dots, X_n sont indépendantes si les n tribus qu'elles engendrent sont indépendantes.

Théorème 4.5 Les variables aléatoires discrètes X et Y sont indépendantes si et seulement si

$$\forall (x, y) \in E \times F \quad P([Z = (x, y)]) = P([X = x])P([Y = y]), \quad \text{ou encore } p_Z(x, y) = p_X(x)p_Y(y)$$

La vérification de ce théorème est simple et laissée en exercice. Ce résultat exprime le fait que la mesure p_Z sur $E \times F$ est le produit des mesures p_X sur E et p_Y sur F .

Théorème 4.6 Les variables aléatoires X et Y étant indépendantes on a $E(XY) = E(X)E(Y)$ si X et Y sont intégrables.

Les variables aléatoires étant indépendantes on a $p_Z = p_X \otimes p_Y$.

$E(XY) = \int_{E \times F} xy dp_X(x) dp_Y(y) = \int_E x dp_X(x) \int_F y dp_Y(y)$ d'après le théorème de Fubini. Ce résultat peut être redémontré directement en écrivant la première intégrale sous forme d'une série double.

Remarque.

La réciproque de ce théorème est fautive comme le montre l'exemple suivant. $E = F = \{-1, 0, 1\}$, $p_Z(-1, 0) = p_Z(1, 0) = p_Z(0, -1) = p_Z(0, 1) = 1/4$, les 5 autres valeurs étant nulles.

$E(XY) = \int_{x \in E} \int_{y \in F} xy dp_Z(x, y) = \sum_{x \in E} \sum_{y \in F} xy p_Z(x, y) = 0$ car dans chaque terme de la somme un des facteurs est nul.

Par symétrie les lois marginales de X et de Y sont identiques.

$$p_X(-1) = p_Z(-1, -1) + p_Z(-1, 0) + p_Z(-1, 1) = 1/4$$

$$p_X(0) = p_Z(0, -1) + p_Z(0, 0) + p_Z(0, 1) = 1/2$$

$$p_X(1) = 1 - (1/4 + 1/2) = 1/4.$$

$$E(X) = -1/4 + 0 + 1/4 = E(Y) = 0. \text{ On a bien}$$

$E(XY) = E(X)E(Y)$ et pourtant $p_Z(0, 0) = 0 \neq p_X(0)p_Y(0)$, ce qui montre

que X et Y ne sont pas indépendantes.

Remarque.

Si les variables aléatoires discrètes X et Y sont indépendantes, il en est de même de $f \circ X$ et $g \circ Y$ où f et g sont des applications de E et F respectivement vers des ensembles dénombrables. Ceci provient de l'égalité $(f \circ X)^{-1}(A) = X^{-1}(f^{-1}(A))$ qui est donc élément de la tribu engendrée par X . En particulier, dans ce cas

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(X - E(X))E(Y - E(Y)) = 0$$

d'après le théorème précédent. Deux variables aléatoires indépendantes ayant une covariance nulle (on dit qu'elles sont non corrélées) on a aussi

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \text{ quand elles sont indépendantes.}$$

Théorème 4.7 *Soit X et Y deux variables aléatoires discrètes indépendantes. On a la relation suivante entre fonctions génératrices*

$$\phi_{X+Y}(z) = \phi_X(z)\phi_Y(z)$$

La réciproque de ce théorème est fautive.

On a en effet $\phi_{X+Y}(z) = E(z^{X+Y}) = E(z^X)E(z^Y)$ d'après une remarque précédente concernant $f(X)$ et $g(Y)$, X et Y étant indépendants.

Exemple.

X et Y étant deux variables aléatoires indépendantes suivant une même loi de Poisson de paramètre λ , $\phi_{X+Y}(z) = e^{2\lambda(z-1)}$, la variable aléatoire $X + Y$ suit donc une loi de Poisson de paramètre 2λ car une fonction génératrice est caractéristique d'une loi.

Chapitre 5

VARIABLES ALEATOIRES CONTINUES

5.1 Loi de probabilité.

5.1.1 Généralités.

Définition 5.1 Une variable aléatoire continue X est une application mesurable d'un espace probabilisé (Ω, \mathcal{E}, P) vers $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ où $\mathcal{B}(\mathbb{R})$ est la tribu des boréliens de \mathbb{R} . On appelle loi de probabilité de X la mesure image p_X de P par X . Cette mesure admet en général une densité g par rapport à la mesure de Lebesgue sur \mathbb{R} , appelée densité de X .

Remarques.

- Si X admet une densité g on a alors
 $\forall A \in \mathcal{B}(\mathbb{R}), p_X(A) = \int_A g(x)dx.$
- Si F_X est la fonction de répartition de X on a
 $F_X(t) = P([X < t]) = p_X(]-\infty, t]) = \int_{-\infty}^t g(t)dt$ d'où $F'_X(t) = g(t)$ si g est continue.
- Une variable aléatoire continue X peut être définie à partir de sa fonction de répartition ou de sa densité, mais par contre $P([X = x]) = 0$ en général (c'est toujours le cas quand on a une densité continue).
- une fonction g est une densité si et seulement si
 $\forall x \in \mathbb{R} f(x) \geq 0, \int_{-\infty}^{+\infty} g(x)dx = 1.$

Les propriétés de la fonction de répartition sont les mêmes que dans le cas discret.

5.1.2 Espérance mathématique et variance

Les définitions des espérance mathématiques et variance sont identiques à celles présentées dans le chapitre précédent. Dans le cas continu on obtient pour une variable intégrable

$$E(X) = \int_{\Omega} X(\omega)dP(\omega) = \int_{\mathbb{R}} xdp_X(x) = \int_{\mathbb{R}} xg(x)dx$$

d'après les propriétés d'une mesure image et d'une mesure à densité. Si h est une fonction réelle continue par morceaux on a de même

$$E(h \circ X) = E(h(X)) = \int_{\Omega} h \circ X(\omega) dP(\omega) = \int_{\mathbb{R}} h(x)g(x)dx$$

sous réserve que $h \circ X$ soit intégrable.

5.2 Lois continues usuelles

5.2.1 Loi Exponentielle.

On appelle loi exponentielle de paramètre λ la loi de densité

$$f(x) = \lambda e^{-\lambda x} \chi_{\mathbb{R}^+}, \lambda > 0$$

elle est notée $\text{Exp}(\lambda)$.

$$E(X) = \int_{\mathbb{R}} x f(x) dx = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = 1/\lambda$$

$$E(X^2) = \int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx = 2/\lambda^2, \text{Var}(X) = E(X^2) - (E(X))^2 = 1/\lambda^2$$

5.2.2 Loi uniforme sur $[a, b]$.

La densité de cette loi est

$$f(x) = \frac{1}{b-a} \chi_{[a,b]}(x), P([a \leq X \leq b]) = \frac{d-c}{b-a}, a \leq c < d \leq b$$

autrement dit cette probabilité est proportionnelle à la longueur de l'intervalle et ne dépend pas de la position des points dans $[a, b]$.

5.2.3 Loi normale ou loi de Gauss.

Une loi normale ou loi de Gauss de paramètres μ, σ est une loi de densité

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

où $\sigma > 0$.

Dans le cas particulier où $\mu = 0, \sigma = 1$, on dit que c'est une loi normale centrée réduite.

Théorème 5.1 Si X suit une loi normale de paramètres $\mu, \sigma, Y = (X - \mu)/\sigma$ suit une loi normale centrée réduite.

En effet $F_Y(t) = P(Y < t) = P(X < \sigma t + \mu) = F_X(\sigma t + \mu)$. En dérivant on obtient $f_X(t) = \sigma f_X(\sigma t + \mu) = 1/\sqrt{2\pi} e^{-t^2/2}$ ce qui est la densité de la loi normale centrée réduite.

Remarque : un calcul rapide donne $E(Y) = 0$ et $\text{Var}(Y) = 1$, ce qui entraîne $E(X) = E(\sigma Y + \mu) = \mu$ et $\text{Var}(X) = \sigma^2 \text{Var}(Y) = \sigma^2$.

5.2.4 Loi Gamma.

La fonction d'Euler de seconde espèce est définie par

$$\Gamma(u) = \int_0^{+\infty} x^{u-1} e^{-x} dx$$

On vérifie que l'intégrale est convergente si et seulement si $u > 0$. Une intégration par partie donne $\Gamma(u+1) = u\Gamma(u)$, ce qui entraîne $\Gamma(n+1) = n!$, $n \in \mathbb{N}^*$. On dit qu'une variable aléatoire X suit une loi Gamma de paramètres $\lambda > 0$, $\alpha > 0$ si sa densité est

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \chi_{\mathbb{R}^+}(x)$$

Un calcul simple donne $E(X) = \alpha/\lambda$, $Var(X) = \alpha/\lambda^2$. La loi exponentielle est un cas particulier de loi Gamma correspondant à $\alpha = 1$.

5.3 Variables aléatoires à valeurs dans \mathbb{R}^n

5.3.1 Loïs de probabilités.

Une variable aléatoire X définie sur l'espace probabilisable (Ω, \mathcal{E}, P) et à valeurs dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ peut être représentée par $X = (X_1, X_2, \dots, X_n)$. La mesure image de P par X est appelée loi de X ou loi conjointe du n -uplet (X_1, X_2, \dots, X_n) . Les lois des variables aléatoires X_i , à valeurs dans \mathbb{R} , sont appelées lois marginales. La loi de X est notée p_X et la loi de X_i p_{X_i} . On traitera dans la suite le cas $n = 2$ pour simplifier les notations et on supposera que p_X admet une densité f par rapport à la mesure de Lebesgue de \mathbb{R}^2

$$\forall B \in \mathcal{B}(\mathbb{R}^2), p_X(B) = \iint_B f(x_1, x_2) dx_1 dx_2$$

Théoreme 5.2 La densité de la variable aléatoire X_1 est $f_{X_1}(x_1) = \int_{\mathbb{R}} f(x_1, x_2) dx_2$

On a en effet pour B_1 borélien de \mathbb{R}

$$p_{X_1}(B_1) = p_X(B_1 \times \mathbb{R}) = \int_{B_1} \left[\int_{\mathbb{R}} f(x_1, x_2) dx_2 \right] dx_1$$

On a un résultat analogue pour p_{X_2} .

5.3.2 Fonction de répartition.

Définition 5.2 On appelle fonction de répartition de $X = (X_1, X_2)$ la fonction définie par $F_X(t_1, t_2) = P(X_1 < t_1, X_2 < t_2)$

La fonction F étant deux fois continument différentiable, on vérifie, en calculant l'intégrale, que

$$P(a \leq X_1 < b, c \leq X_2 < d) = \iint_{[a,b] \times [c,d]} \frac{\partial^2 F}{\partial x_1 \partial x_2} dx_1 dx_2$$

Les rectangles $[a, b[\times]c, d[$ engendrant la tribu boélienne de \mathbb{R}^2 ceci permet d'obtenir la densité f de X par la formule

$$f(x_1, x_2) = \frac{\partial^2 F}{\partial x_1 \partial x_2}.$$

5.3.3 Somme de variables aléatoires indépendantes.

Deux variables aléatoires réelles X et Y sont indépendantes, comme dans le cas discret, si les tribus engendrées par X et Y sont indépendantes, ou encore si $p_X = p_{X_1} \otimes p_{X_2}$ ce qui est équivalent à $F_Z(x, y) = F_X(x)F_Y(y)$, $Z = (X, Y)$ ou à $f_Z(x, y) = f_X(x)f_Y(y)$.

Théorème 5.3 *Les variables aléatoires X et Y étant indépendantes, on a :*

- $E(XY) = E(X)E(Y)$.
- $Var(X + Y) = Var(X) + Var(Y)$.
- $Cov(X, Y) = 0$.

la deuxième et la troisième égalité sont des conséquences de la première qui se démontre à l'aide du théorème de Fubini.

Théorème 5.4 *En posant $S = X + Y$ on a*

$$f_S(t) = \int_{-\infty}^t \left(\int_{-\infty}^{\infty} f_X(x)f_Y(u-x)dx \right) du$$

ce qui exprime que f_S est le produit de convolution de f_X et f_Y .

Démonstration.

$$F_S(t) = P(S < t) = \iint_{(x,y), x+y < t} f_X(x)f_Y(y)dxdy = \int_{\mathbb{R}} \left(\int_{-\infty}^{t-x} f_X(x)f_Y(y)dy \right) dx$$

En posant $y = u - x$ on obtient

$$F_S(t) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^t f_X(x)f_Y(u-x)du \right) dx = \int_{-\infty}^t \left(\int_{\mathbb{R}} f_X(x)f_Y(u-x)dx \right) du.$$

5.4 Fonction caractéristique d'une variable aléatoire.

Définition 5.3 — *X étant une variable aléatoire réelle, sa fonction caractéristique est définie par*

$$\phi_X(t) = E(e^{itX}) = \int eitxf_X(x)dx.$$

— *$Z = (X, Y)$ étant une variable aléatoire à valeurs dans \mathbb{R}^2 , sa fonction caractéristique est définie par*

$$\phi_Z(t, v) = E(e^{i(tX+vY)}) = \iint e^{i(tx+vy)} f_Z(x, y)dxdy.$$

Exemple.

Loi uniforme sur $[-A, A]$. un calcul simple donne

$$\phi_X(t) = \frac{\sin At}{At}$$

Loi normale centrée réduite.

Un calcul utilisant les propriétés des fonctions holomorphes donne

$$\phi_X(t) = e^{-\frac{t^2}{2}}$$

Propriétés.

- Si $Y = aX + b$, $\phi_Y(t) = e^{ibt} \phi_X(at)$.
- $\phi_{X+Y}(t) = \phi_{(X,Y)}(t, t)$.
- Si $X \in L^1$ le théorème de dérivation sous le signe somme donne $\phi'_X(0) = iE(X)$.
- Si $X \in L^2$, une nouvelle application de ce théorème donne $\phi''_X(0) = -E(X^2)$.

Le premier résultat permet de calculer la fonction caractéristique d'une variable aléatoire X suivant une loi normale de moyenne μ et d'écart type σ , car dans ce cas $X = \sigma Y + \mu$ avec Y variable aléatoire suivant une loi normale centrée réduite.

$$\phi_X(t) = e^{i\mu t - (\sigma^2 t^2)/2}.$$

Le théorème qui suit, qui est donné sans démonstration, montre que la fonction caractéristique d'une variable aléatoire détermine totalement sa loi.

Théoreme 5.5 *Si ϕ_X , fonction caractéristique de la variable aléatoire réelle X est intégrable sur \mathbb{R} , on a*

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_X(t) e^{-itx} dt.$$

Théoreme 5.6 *Les variables aléatoires X et Y étant indépendantes, on a*

$$\phi_{X+Y}(t) = \phi_X(t) \phi_Y(t).$$

Ceci provient de l'égalité $E(e^{it(X+Y)}) = E(e^{itX})E(e^{itY})$ provenant de l'indépendance des variables aléatoires e^{itX} et e^{itY} .

Il faut noter que la réciproque de ce théorème est fautive. En effet, si X suit la loi de Cauchy de densité $f(x) = 1/(\pi(1+x^2))$, $\phi_X(t) = e^{-|t|}$ et $\phi_{2X}(t) = \phi_X(2t) = e^{-2|t|} = (\phi_X(t))^2$.

5.5 Théorèmes limites.

5.5.1 Inégalité de Markov.

Théoreme 5.7 *La variable aléatoire X étant à valeurs dans \mathbb{R}^+ , on a l'inégalité*

$$P(X \geq a) \leq \frac{1}{a} E(X), a > 0.$$

Démonstration La densité de probabilité de X étant f ,

$$E(X) = \int_0^{+\infty} xf(x)dx \geq \int_a^{+\infty} xf(x)dx \geq \int_a^{+\infty} af(x)dx = aP(X \geq a).$$

5.5.2 Inégalité de Tchebychev.

Théoreme 5.8 Une variable aléatoire X ayant une variance finie vérifie

$$P(|X - E(X)| \geq \lambda) \leq \frac{Var(X)}{\lambda^2}.$$

C'est une conséquence immédiate de l'inégalité de Markov.

5.5.3 Lois des grands nombres.

Loi faible des grands nombres.

Théoreme 5.9 Soient (X_n) une suite de variables aléatoires indépendantes et de même loi, d'espérance mathématique μ . On a

$$\forall k > 0, \lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq k\right) = 0$$

Ce théorème est vrai même lorsque la variance n'existe pas, toutefois, en cas d'existence de $Var(X)$ ce théorème se déduit immédiatement de l'inégalité de Tchebychev.

Loi forte des grands nombres.

En fait, sous les hypothèses précédentes, on a le résultat plus fort suivant : la suite S_n/n converge presque sûrement vers μ .

5.5.4 Théorème central limite.

Définition 5.4 Une suite de variables aléatoires réelles (X_n) converge en loi vers la variable aléatoire réelle X si la suite de fonctions ϕ_{X_n} caractéristiques converge simplement vers ϕ_X .

Ceci entraîne en particulier la convergence simple de la suite de fonctions de répartition (F_n) vers F fonction de répartition de X en tout point de continuité de F . Ceci revient en quelque sorte à dire qu'à la limite la loi de X_n est la loi de X .

Théoreme 5.10 (Moivre-Laplace) Les (X_n) étant des variables aléatoires indépendantes, de même loi, admettant μ et σ respectivement comme espérance mathématique et variance

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \text{ converge en loi vers la loi normale centrée réduite.}$$

Démonstration

$$\phi_{\frac{X_k - \mu}{\sigma}}(t) = E\left(e^{it\frac{X_k - \mu}{\sigma}}\right)$$

est deux fois dérivable car X_n est de carré intégrable (on utilise à deux reprises le théorème de dérivation sous le signe somme).

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \sum_{i=1}^n \frac{X_i - \mu}{\sigma}, \quad \phi_{\frac{S_n - n\mu}{\sigma\sqrt{n}}}(t) = \prod_{i=1}^n \phi_{\frac{X_i - \mu}{\sigma}}(t) = \left(\phi_{\frac{X_1 - \mu}{\sigma}}(t)\right)^n$$

$$\phi'_{\frac{X_1 - \mu}{\sigma}}(0) = E\left(\frac{X_1 - \mu}{\sigma}\right) = 0, \quad \phi''_{\frac{X_1 - \mu}{\sigma}} = \text{Var}\left(\frac{X_1 - \mu}{\sigma}\right) = 1$$

on a donc

$$\phi_{\frac{X_1 - \mu}{\sigma}}(t) = 1 - \frac{t^2}{2} + o(t^2)$$

ce qui entraîne

$$\lim_{n \rightarrow \infty} \phi_{\frac{S_n - n\mu}{\sigma\sqrt{n}}}(t) = e^{-\frac{t^2}{2}}$$

qui est la fonction caractéristique de la loi normale centrée réduite.

5.6 Exercices.

Exercice 1

Une urne contient r boules blanches et s boules noires. On extrait simultanément, au hasard, n boules, $1 \leq n \leq r + s$ et on s'intéresse au nombre de boules blanches tirées.

- 1) Comment modéliser cette expérience? (Choix de l'ensemble fondamental Ω et de la probabilité P).
- 2) Quel est en fonction de r , s , et n , l'ensemble \mathcal{V} des valeurs possibles prises par le nombre de boules blanches tirées?
- 3) On note A_k l'évènement "Parmi les n boules tirées figurent k boules blanches". Calculer $P(A_k)$, $k \in \mathcal{V}$.
- 4) Quelle relation obtient-on sur les coefficients binomiaux?

Exercice 2

Le jeu du Loto consiste à deviner les 6 entiers qui vont être tirés au hasard parmi les entiers $\{1, 2, \dots, 49\}$. On ne tient pas compte ici du numéro complémentaire.

- 1) Comment modéliser le tirage de 6 entiers parmi les 49 premiers entiers naturels non nuls?
- 2) Le joueur propose un tirage. On considère les évènements suivant :
 - A_k = "avoir deviné exactement k bons numéros", $k \in \{0, 1, 2, 3, 4, 5, 6\}$.
 - P = "perdre" = "avoir deviné 0, 1 ou 2 bon(s) numéro(s)".
 - G = "gagner" = "avoir au moins 3 bons numéros".

Calculer la probabilité de ces événements.

Exercice 3

On dispose de 10 cartes numérotées de 1 à 10. On tire successivement, au hasard, chacune des 10 cartes.

Modéliser cette expérience. On appelle A_k l'évènement "la k ème carte tirée porte le numéro k ", $k \in \{0, 1, \dots, 10\}$. Calculer la probabilité de A_k .

Calculer $P(A_i \cap A_j)$, $i \neq j$, puis $P(\overline{A_i} \cap A_j)$ et $P(\overline{A_i} \cap \overline{A_j})$.

Exercice 4

Une urne contient $n \in \mathbb{N}^*$ boules numérotées de 1 à n et indiscernables au toucher. On tire une boule au hasard. p_1, p_2, \dots, p_r étant les diviseurs premiers de n , on note A_{p_i} l'évènement le numéro porté par la boule tirée est divisible par p_i .

- 1) Calculer $P(A_{p_i})$, $1 \leq i \leq r$. Montrer que les évènements A_{p_i} sont indépendants dans leur ensemble.
- 2) Calculer la probabilité de l'évènement A = "le numéro porté par la boule tirée est premier avec n " (il n'est divisible par aucun p_i). En déduire le nombre $\phi(n)$ des entiers plus petits que n et premiers avec n (la fonction $\phi(n)$ est appelée fonction d'Euler et joue un rôle important en arithmétique).

Exercice 5

Une urne contient des boules blanches ($1/3$) et des boules noires ($2/3$). On tire une boule. Si elle est blanche, le jeu est terminé. Si elle est noire, on la remet et on procède à un nouveau tirage et ceci jusqu'à l'obtention d'une boule blanche, le jeu est alors terminé. On appelle X le nombre de tirages à effectuer pour voir le jeu se terminer.

1) Quelle est la loi suivie par X . déterminer son espérance mathématique et calculer $P([X \leq 3])$ et $P([X > 4])$

2) Le gain étant de trente francs si la boule blanche sort au premier ou au deuxième coup, vingt francs au troisième, dix francs si elle sort au quatrième et cinq francs ensuite et sachant qu'il faut payer dix francs pour procéder à un tirage, le jeu est-il équitable? (c'est à dire a-t-on $E(Z) = 0$ en appelant Z le bilan global) .

N.B. : En appelant Y le gain obtenu, déterminer la loi de Y et écrire une relation entre X , Y et Z .

Exercice 6

La loi de probabilité d'une variable aléatoire continue est donnée par

$$f(x) = \begin{cases} r\lambda^r \frac{1}{x^{r+1}} & \text{pour } x > \lambda \\ 0 & \text{ailleurs} \end{cases}$$

r et λ étant deux paramètres strictement positifs.

1) Vérifier que f est une densité de probabilité.

2) Calculer la fonction de répartition dans le cas où $r = \lambda = 2$.

3) Calculer $E(X^k)$ dans le cas général. Cette quantité existe-t-elle toujours?

4) En déduire l'espérance mathématique et la variance de X .

Exercice 7

On considère 2 variables aléatoires indépendantes X et Y , de même loi, dont la densité est la fonction

$$f(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

1) Déterminer $F_{X^2}(x) = P([X^2 < x])$ en fonction de $F_X(\sqrt{x}) = P([X < \sqrt{x}])$. En déduire la densité de X^2 .

2) Déterminer la densité de la variable aléatoire $Z = X + Y$.

3) Déterminer la densité du couple (X, Y) En déduire la probabilité $P(X < 2Y)$.

Exercice 8

Soit $n \in \mathbb{N}^*$, on rappelle les identités

$$\sum_{i=1}^n i = \frac{n(n+1)}{2}, \quad \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

Une urne contient r , $r > 1$ boules numérotées par $M = \{1, 2, \dots, r\}$, on y prélève au hasard, successivement et sans remise 2 boules.

Soient X_1 , X_2 les variables aléatoires égales aux numéros des boules extraites de l'urne au premier et au second tirage.

1) a) Déterminer la loi du couple (X_1, X_2) c'est à dire les probabilités $P([X_1 = k_1, X_2 = k_2])$. Déterminer les lois marginales de X_1 et X_2 . Ces variables sont-elles indépendantes ?

b) Calculer $E(X_1)$, $E(X_2)$, $\text{Var}(X_1)$, $\text{Var}(X_2)$.

c) Pour $k \in M$, $k > 1$ calculer $P([X_1 = k] \cap [X_2 < k])$ et $P([X_1 < k] \cap [X_2 = k])$.

On considère maintenant la variable aléatoire U , $U(\omega) = \max\{X_1(\omega), X_2(\omega)\}$ et la variable aléatoire V définie par

$V(\omega) = 1$ si $X_1(\omega) > X_2(\omega)$, $V(\omega) = 2$ si $X_1(\omega) < X_2(\omega)$.

3) a) Déterminer la loi du couple (U, V) .

b) Déterminer la loi de U et la loi de V . Ces variables sont-elles indépendantes ?

Exercice 9

Le plan est rapporté à un repère orthonormé $R = (O, \vec{i}, \vec{j})$ et on considère le couple de variables aléatoires continues (X, Y) dont la densité est uniforme dans le disque D de centre O et de rayon R .

$$f_{(X,Y)}(x, y) = \frac{1}{\pi R^2} \chi_D(x, y)$$

où χ_D est la fonction qui prend la valeur 1 dans D et 0 à l'extérieur de D .

1) On note $M(\omega) = (X(\omega), Y(\omega))$, $r(\omega) = OM(\omega) = \sqrt{X^2(\omega) + Y^2(\omega)}$ et $\theta(\omega)$ est la mesure de l'angle (\vec{i}, \vec{OM}) prise entre 0 et 2π .

Déterminer la fonction de répartition de la variable aléatoire r , c'est à dire calculer $F_r(x) = P([r < x])$.

2) Déterminer $F_{r,\theta}(x, y) = P([r < x] \cap [\theta < y])$ fonction de répartition conjointe du couple (r, θ) .

En déduire les densités de r et θ . Ces deux variables aléatoires sont-elles indépendantes ?

Exercice 10

Trois usines A, B, C fournissent respectivement 20 pour cent, 30 pour cent et 50 pour cent des pièces détachées à une entreprise. 5 pour cent des pièces fournies par A sont défectueuses de même que 3 pour cent de celles fournies par B et 2 pour cent de celles fournies par C. Une pièce choisie au hasard est défectueuse quelle est la probabilité qu'elle provienne de l'usine A ?

Exercice 11

une urne contient 4 boules blanches, 3 boules noires et 3 boules rouges. On tire simultanément trois boules de l'urne déterminer les probabilités des évènements suivants

1) A = " Les trois boules tirées sont blanches "

2) B = " On tire deux boules noires et une boule rouge "

- 3) C = " On tire au moins une boule blanche "
- 4) D = " On tire une boule de chaque couleur "

Exercice 12

Un étudiant estime avoir 2,5 chances sur 100 d'arriver en retard au cours chaque matin. Ce retard est imputable à un grand nombre de facteurs (réveil-matin, bus, embouteillages ...) et il l'attribue au hasard. on considère une période de 40 jours ouvrables numérotés de 1 à 40.

- 1) On appelle X le numéro du premier jour de retard. Quelle est la loi suivie par X? Quelle est la probabilité qu'il arrive en retard pour la première fois le dixième jour?
- 2) On appelle Y le nombre de retards observables sur la période. Quelle est la loi suivie par Y? Quelle est la probabilité qu'il arrive en retard une seule fois?

Exercice 13

Un amphitheâtre comprend 200 élèves des écoles de Coetquidan. On demande à 30 d'entre-eux, choisis au hasard, leur avis sur une nouvelle loi concernant la réforme des écoles. Quelle est la probabilité que, sur les 30, on observe 10 réponses favorables, sachant que dans l'ensemble 48 pour cent sont favorables

Remarque : on pourra appeler X le nombre de réponses favorables parmi les 30 élèves choisis (ici $X = 10$) et déterminer la loi suivie par X.

Exercice 14

Une usine désire fabriquer des pièces de longueurs l avec $12 < l < 18$. Pour cela on dispose d'une machine qui réglée en m , avec $10 < m < 20$, coupe les pièces de façon que la longueur soit une variable aléatoire L de densité f_m définie par

$$f_m(t) = \begin{cases} k_m(m+1-t)(t-(m-1)) & \text{si } t \in [m-1, m+1] \\ 0 & \text{si } t \notin [m-1, m+1] \end{cases}$$

où $k_m \in \mathbb{R}$. On notera F_m la fonction de répartition de L .

- 1) Déterminer k_m pour que f_m soit une densité de probabilité.
- 2) La machine étant réglée en m , déterminer la probabilité qu'une pièce ait une longueur supérieure à $m - 0,5$.
- 3) La machine étant réglée en m , déterminer la probabilité qu'une pièce ait une longueur inférieure à $m + 0,5$ sachant que la longueur est supérieure à $m - 0,5$.
- 4) Une pièce est acceptée si sa longueur est comprise entre $l - 0,5$ et $l + 0,5$, elle est perdue si sa longueur est inférieure à $l - 0,5$, ce qui engendre une perte de 10F, si elle est de taille supérieure à $l + 0,5$ on peut la retailler, ce qui engendre une perte de 3F. On appelle X la variable aléatoire égale à la perte subie (0 si elle est acceptée). Déterminer la loi de X en fonction de $F_m(l - 0,5)$ et de $F_m(l + 0,5)$.
- 5) Déterminer le réglage m à adopter pour que l'espérance de X soit minimum.

Exercice 15

Etant donné un réel $a > 0$ soit la fonction f_a définie par

$$f_a(x) = \frac{\lambda_a}{a^2 + x^2}$$

- 1) Déterminer λ_a pour que f_a soit la densité d'une variable aléatoire X_a .
- 2) Montrer que la fonction caractéristique de la variable aléatoire X_a est $\phi_X(t) = e^{-a|t|}$.
- 3) Les réels a et b étant strictement positifs on appelle X_a et X_b deux variables aléatoires indépendantes de densités respectives f_a et f_b . Déterminer les fonctions caractéristiques des variables aléatoires $Y = X_a + X_b$ et $Z = X_a - X_b$. En déduire les densités de Y et Z .

Exercice 16

Une urne contient un tiers de boules blanches et deux tiers de boules noires. On tire une boule, si elle est blanche le jeu est terminé, si elle est noire on la remet dans l'urne et on procède à un nouveau tirage. On répète cette opération jusqu'à l'obtention d'une boule blanche. On appelle X le nombre de tirages à effectuer pour voir le jeu se terminer.

- 1) Déterminer la loi de X et calculer l'espérance mathématique de X .
- 2) Le gain Y est de 30 F si la boule blanche sort au premier ou au deuxième coup, 20 F si elle sort au troisième, 10 F si elle sort au quatrième et 5 F ensuite. Déterminer la distribution de probabilité de Y . Sachant qu'il faut payer 10 F pour procéder à un tirage écrire le bilan global Z en fonction de X et Y . Le jeu est-il équitable? (c'est à dire a-t-on $E(Z) = 0$?)

Exercice 17

On considère deux variables aléatoires indépendantes X et Y , de même loi dont la densité est la fonction $f(x) = \lambda e^{-\lambda x} \chi_{]0, +\infty[}(x)$, $\lambda > 0$

- 1) Déterminer $F_{X^2}(x) = P([X^2 < x])$ en fonction de $F_X(\sqrt{x}) = P([X < \sqrt{x}])$. En déduire la densité de X^2 .
Remarque : La variable aléatoire X prend des valeurs positives.
- 2) On définit $Z = \max(X, Y)$ et $T = \min(X, Y)$. Montrer que $F_T(x) = F_X(x) + F_Y(x) - F_X(x)F_Y(x)$ et que $F_Z(x) = F_X(x)F_Y(x)$
(Remarque : $F_X = F_Y$). En déduire les densités de Z et de T .
- 3) On note $D = \{(z, t) \in \mathbb{R}^2 | z > 0, t > 0, z \geq t\}$. Montrer que la densité du couple (Z, T) est donnée par $g(z, t) = 2f_{(X, Y)}(z, t)\chi_D(z, t)$ où $f_{(X, Y)}$ est la densité conjointe du couple (X, Y) (remarque que si $C = [a, b] \times [c, d] \in D$ on a

$$P((Z, T) \in C) = \int \int_C 2f_{(Z, T)}(z, t) dz dt$$

Retrouver les densités de Z et de T .

Exercice 18

On dispose de 7 dés à 6 faces. On numérote ces dés de 1 à 7 de telle sorte que le dé numéro i a $i - 1$ faces noires et $7 - i$ faces blanches.

On se propose dans un premier temps de choisir un de ces dés. A cette fin on lance un huitième dé et on regarde le résultat :

- Si ce résultat est i ($2 \leq i \leq 6$) on choisit le dé numéro i .
- Si ce résultat est 1 on lance une deuxième fois le dé.

Le dé choisi sera alors le dé numéro 1 si le second résultat est 1, 2 ou 3 et le dé numéro 7 si le second résultat est 4, 5 ou 6.

Pour $1 \leq i \leq 7$, on note D_i l'évènement "le dé n° i est choisi" et A l'évènement "le premier jet du huitième dé a donné 1".

1)

- a) Quelle est la valeur de $P(D_i)$ pour $2 \leq i \leq 6$.
- b) Quelle est la valeur de $P(D_i|A)$ pour $i = 1$ et $i = 7$. En déduire $P(D_1)$ et $P(D_7)$.

Après avoir sélectionné un dé on le lance plusieurs fois et on s'intéresse aux événements :

N_k : "On a obtenu une face noire au k -ième lancer du dé sélectionné" ($k \geq 1$).

$S_n = \cap_{k=1}^n N_k$, $n \geq 1$.

On admettra que

$$P(S_n|D_i) = \prod_{k=1}^n P(N_k|D_i), n \geq 1 \text{ et } 1 \leq i \leq 7$$

- 2) a) Que vaut $P(N_k|D_i)$?
- b) Montrer que $P(N_k) = 1/2$.
- 3) a) Calculer $P(S_2)$. Les événements N_1 et N_2 sont-ils indépendants?
- b) Calculer $P(N_2|N_1)$.

Exercice 19

On considère la fonction

$$f(x) = \begin{cases} xe^{-\frac{x^2}{2}} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

- 1) Montrer que f peut être considérée comme une densité de probabilité. On appelle X une variable aléatoire dont la loi de probabilité admet f comme densité.
- 2) Déterminer la fonction de répartition F_Y de la variable aléatoire $Y = X^2$ en fonction de la fonction de répartition de X . En déduire la densité de la variable aléatoire Y .
- 3) Calculer l'espérance mathématique de X ainsi que sa variance.

Exercice 20

Un gardien de nuit doit ouvrir une des portes à contrôler durant sa tournée, dans le noir. Il possède 10 clés d'allure semblable, mais une seule peut ouvrir la porte en question. Le gardien dispose de deux méthodes.

- Méthode A : Il pose les 10 clés devant lui et les essaye l'une après l'autre dans l'ordre dans lequel elles se présentent
- Méthode B : il essaye une clé après avoir agité le trousseau, en recommençant cette opération jusqu'à ce qu'il trouve la bonne clé .

On appelle X_A (resp. X_B) la variable aléatoire qui désigne le nombre de clés essayées (y compris celle qui donne satisfaction) par la méthode A (resp. B).

1) Montrer que X_A suit la loi uniforme sur $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Quelle est la loi suivie par X_B ?

2) Quelle est la probabilité d'essayer plus de huit clés par la méthode A ? Par la méthode B ? On notera H l'événement " essayer plus de huit clés ".

Le gardien utilise la méthode A deux fois sur trois, quelle est la probabilité conditionnelle que le gardien utilise la méthode B sachant que les huit premiers essais ont échoués ?

Exercice 21

On définit

$$f(x, y) = \begin{cases} k(1 + xy(x^2 - y^2)) & \text{si } -1 \leq x \leq 1 \text{ et } -1 \leq y \leq 1 \\ 0 & \text{sinon} \end{cases}$$

1) Déterminer la valeur de k pour que f soit une densité de probabilité associée à un couple (X, Y) de variables aléatoires.

2) Déterminer les densités marginales de X et Y ainsi que leurs fonctions caractéristiques.

3) Déterminer la fonction caractéristique de la variable $X+Y$. Conclusion ? Pour le calcul de cette fonction caractéristique on pourra utiliser sans démonstration le résultat suivant où ϕ_U est la fonction caractéristique de U .

$$\phi_{X+Y}(t) = \phi_{(X,Y)}(t, t).$$

Exercice 22

1) a) Une urne contient une proportion p de boules blanches et q de boules noires ($p+q = 1$). On répète des tirages avec remise d'une boule dans l'urne. On appelle X le nombre de répétitions nécessaires pour obtenir la première boule blanche . Quel type de loi suit X ?

b) Déterminer la fonction génératrice $G_X(t)$ de X . En déduire la fonction génératrice de $S_n = X_1 + X_2 + \dots + X_n$ où les X_i sont des variables aléatoires indépendantes de même loi que X .

2) On appelle Y_r le nombre de répétitions nécessaires pour obtenir exactement r boules blanches. Montrer que

$$P([Y_r = n]) = C_{n-1}^{r-1} p^r q^{n-r}$$

La loi suivie par Y_r est appelée loi binomiale négative.

3) Montrer que

$$\frac{1}{(1-x)^n} = \frac{1}{(n-1)!} \frac{d^{n-1}}{dx^{n-1}} \frac{1}{(1-x)} = \sum_{i=0}^{\infty} C_{n-1+i}^i x^i$$

En déduire la fonction génératrice $G_{Y_r}(t)$ de Y_r . Quelle type de loi suit la variable aléatoire S_n introduite dans la question précédente ?

Table des matières

I	ESPACES DE HILBERT	1
1	ESPACES METRIQUES	2
1.1	Définitions	2
1.2	Exemples d'espaces métriques	2
1.3	Propriétés.	3
1.4	Espaces métriques complets	3
1.5	Espaces vectoriels normés	4
1.6	Espaces vectoriels normés complets	5
1.7	Applications linéaires continues	5
2	ESPACES DE HILBERT	9
2.1	Formes Hermitiennes	9
2.2	Espaces de Hilbert	11
2.3	Théorème de projection	12
2.4	Projection sur un sous-espace vectoriel fermé	13
2.5	Espaces de Hilbert et systèmes orthogonaux	15
2.6	Bases orthonormales de $L^2([a, b])$	18
2.6.1	Application	19
2.6.2	Convergence ponctuelle d'une série de Fourier	20
2.7	Opérateur linéaire sur un espace de Hilbert	21
2.8	Eléments propres d'un opérateur hermitien	24
2.9	Exercices	26
II	PROBABILITES	35
3	NOTIONS FONDAMENTALES	36
3.1	Ensemble fondamental. Evènements.	36
3.1.1	Exemples	36
3.1.2	Définition	37
3.2	Probabilité. Espace Probabilisé	37
3.2.1	Définition	37
3.2.2	Propriétés des probabilités	37
3.3	Probabilité conditionnelle.	38
3.3.1	Exemple.	38
3.3.2	Définition.	39

3.3.3	Théorème de Bayes	39
3.3.4	Evènements indépendants.	40
4	VARIABLES ALEATOIRES DISCRETES	41
4.1	Loi de probabilité.	41
4.1.1	Mesure image et mesure définie par une densité	41
4.1.2	Variable aléatoire	42
4.1.3	Loi de probabilité.	42
4.1.4	Fonction de répartition.	43
4.2	Espérance mathématique. Variance.	43
4.2.1	Espérance mathématique.	43
4.2.2	Fonction génératrice.	44
4.2.3	Propriétés de l'espérance mathématique.	44
4.2.4	Variance.	44
4.3	Lois de probabilités discrètes usuelles	45
4.3.1	Loi de Bernoulli.	45
4.3.2	Loi Binomiale.	45
4.3.3	Loi géométrique.	46
4.3.4	Loi hypergéométrique.	46
4.3.5	Loi de Poisson.	46
4.4	Couple de variables aléatoires.	46
5	VARIABLES ALEATOIRES CONTINUES	49
5.1	Loi de probabilité.	49
5.1.1	Généralités.	49
5.1.2	Espérance mathématique et variance	49
5.2	Lois continues usuelles	50
5.2.1	Loi Exponentielle.	50
5.2.2	Loi uniforme sur $[a, b]$	50
5.2.3	Loi normale ou loi de Gauss.	50
5.2.4	Loi Gamma.	51
5.3	Variables aléatoires à valeurs dans \mathbb{R}^n	51
5.3.1	Lois de probabilités.	51
5.3.2	Fonction de répartition.	51
5.3.3	Somme de variables aléatoires indépendantes.	52
5.4	Fonction caractéristique d'une variable aléatoire.	52
5.5	Théorèmes limites.	53
5.5.1	Inégalité de Markov.	53
5.5.2	Inégalité de Tchebychev.	54
5.5.3	Lois des grands nombres.	54
5.5.4	Théorème central limite.	54
5.6	Exercices.	56