Chapitre 2: Le théorème de projection et ses applications

*

21 décembre 2007

1 Introduction

En géométrie élémentaire, si P est un plan et x un point qui n'appartient pas à P, il existe un unique point $y \in P$ qui est le plus proche de x au sens de la distance euclidienne. Ce point y est en fait la projection orthogonale de x sur le plan P. On va généraliser de manière abstraite cette propriété aux espaces de Hilbert.

2 Projection sur un convexe fermé

2.1 Le résultat général

Soit V un espace normé sur $\mathbb{R}$ ou $\mathbb{C}$. On rappelle les définitions suivantes

Définition 2.1 *: 1) On appelle segment d'extrémités $a, b \in V$, le sous ensemble de V défini par*

$$[a,b] = \{\lambda a + (1-\lambda)b \in V;\ 0 \leq \lambda \leq 1\} \tag{1}$$

2) un sous ensemble $A \subset V$ est convexe si pour tous $a, b \in A$, $[a,b] \subset A$.

Exemple : Tout sous espace vectoriel de V est un ensemble convexe.

On considère à partir de maintenant un espace de Hilbert H sur $\mathbb{C}$ avec sa norme $||x|| = \sqrt{< x, x >}$.

Théorème 2.2 *(de projection) : Soit A un sous ensemble convexe fermé (et non vide) de H. Alors pour tout $x \in H$, il existe un unique $y \in A$ tel que*

$$\inf_{a \in A} ||x - a|| = ||x - y||. \tag{2}$$

Autrement dit il existe un unique point $y \in A$ qui est à une distance de x la plus petite possible[1]. Ce point y s'appelle la projection de x sur A.

*Notes du cours sur les espaces de Hilbert de M. L. Gallardo, Licence 3-ième année, Université de Tours, année 2007-2008. Les démonstrations sont données dans le cours oral.

[1] on notera qu'une borne inf n'est pas toujours atteinte. Ici le théorème dit que l'inf est atteint et en un unique point.

2.2 Projection sur un sous espace fermé

Le cas particulier le plus important du théorème précédent est la projection sur un sous epace vectoriel **fermé** F de H. Soit alors $x \in H$.

Corollaire 2.3 *: 1) soit $y \in F$ tel que $||x-y|| = \inf_{z\in F} ||x-z||$. Alors $x-y$ est orthogonal à F (i.e. orthogonal à tous les vecteurs $z \in F$).*
2) Réciproquement si $y \in F$ est tel que $x-y \perp F$, alors $||x-y|| = \inf_{z\in F} ||x-z||$ i.e. y est la projection de x sur F.

Définition 2.4 *(et* **Notation***) : Dans le cas du résultat précédent, on notera $y = P_F(x)$ et on dira que y est la* **projection orthogonale** *de x sur F.*

Corollaire 2.5 *: La projection orthogonale $P_F : H \to F$ est une contraction linéaire (i.e. une application linéaire telle que pour tout $x \in H$, $||P_F(x)|| \leq ||x||$).*

3 Applications du théorème de projection

3.1 Supplémentaires orthogonaux

Soit H un espace de Hilbert et $A \subset H$, $A \neq \emptyset$ un sous ensemble quelconque.

Définition 3.1 *: On appelle orthogonal de A, l'ensemble*

$$A^\perp = \{x \in H;\ \forall y \in A,\ < x, y >= 0\}\,. \tag{3}$$

Proposition 3.2 *: $A^\perp$ est toujours un sous espace vectoriel fermé de H.*

Exercice : Soit M un sous espace vectoriel de H. Montrer que $x \in M^\perp$ si et seulement si

$$\forall y \in M,\ ||x-y|| \geq ||x||. \tag{4}$$

(indication : pour la condition suffisante, pour $y_0 \in M$ et $F = \mathbb{C}y_0$ la droite engendrée par y_0, on pourra remarquer que $P_F(x) = 0$ puis utiliser la propriété caractéristique de P_F).

Théorème 3.3 *: Soit F un sous espace vectoriel* **fermé** *de H. Alors*

$$H = F \oplus F^\perp \quad \textit{(somme directe orthogonale)}. \tag{5}$$

Autrement dit :

$$\forall x \in H,\ \exists! y \in F,\ \exists! z \in F^\perp,\ x = y + z \tag{6}$$

(ou le quantificateur $\exists!$ signifie «il existe un unique»).

Remarque et Exercice : L'hypothèse que F est fermé dans H est fondamentale. Par exemple si $H = l^2$, le sous espace vectoriel E des $x = (x_i)_{i\in\mathbb{N}^*}$ tels que seules un nombre fini des coordonnées x_i sont non nulles[2], n'est pas fermé dans H et on a $E^\perp = \{0\}$. Dans ce cas $E \oplus E^\perp = E$ n'est pas égal à H.

[2] i.e. $x_i = 0$ pour tout i assez grand.

3.2 Décompositions orthogonales

Soit D un ensemble fini ou dénombrable. Dans toute la suite on conviendra que $D = \{1, 2, \dots, d\}$ si D est fini et $D = \mathbb{N}^*$ si D est dénombrable. Soit alors H un espace de Hilbert et $(e_n)_{n\in D}$ une famille de vecteurs de H.

Définition 3.4 *: On dit que $(e_n)_{n\in D}$ est un système orthonormal dans H si :*

$$\forall n \neq m \in D, \; < e_n, e_m >= 0. \tag{7}$$

$$\forall n \in D, \; ||e_n|| = \sqrt{< e_n, e_n >} = 1 \tag{8}$$

Proposition 3.5 *: Si $(e_n)_{n\in D}$ est un système orthonormal, alors c'est une famille (algébriquement) libre de vecteurs de H.*

Corollaire 3.6 *: Si $(e_n)_{n\in D}$ est un système orthonormal dans H, alors $cardD \leq dimH$. En particulier dans un espace de Hilbert de dimension finie les systèmes orthonormaux sont finis.*

Théorème 3.7 *(projection sur un s.e.v. de dimension finie) : Soit $e_1, \dots, e_n$ un système orthonormal fini et $V = V[e_1, \dots, e_n]$ le sous espace vectoriel de H engendré par les e_i. Alors*

$$\forall x \in H, \; P_V(x) = \sum_{i=1}^{n} < x, e_i > e_i. \tag{9}$$

Corollaire 3.8 *: Si $x \in V[e_1, \dots, e_n]$, alors on a*

$$x = \sum_{i=1}^{n} < x, e_i > e_i \quad et \quad ||x||^2 = \sum_{i=1}^{n} | < x, e_i > |^2. \tag{10}$$

Corollaire 3.9 *(inégalité de Bessel) : Soit $(e_n)_{n\in D}$ est un système orthonormal dans H, alors*

$$\forall x \in H, \; \sum_{n\in D} | < x, e_i > |^2 \leq ||x||^2. \tag{11}$$

(où $\sum_{n\in D} | < x, e_i > |^2$ est une somme finie si D est fini et une série convergente si $D = \mathbb{N}^*$).

3.3 Séries de Fourier associées à un système orthonormal

Soit $(e_n)_{n\in D}$ est un système orthonormal dans H qu'on suppose fixé pour les définitions et les résultats qui suivent.

Définition 3.10 *: Pour tout $x \in H$, on appelle :*
1) coefficient de Fourier d'ordre $n \in D$ (où n-ième coefficient de Fourier) le nombre $< x, e_n >$ $(\in \mathbb{C})$,
2) série de Fourier de x la série[3] $\sum_{n\in D} < x, e_n > e_n$.

[3] c'est une série de vecteurs de H si $D = \mathbb{N}^*$ et une somme finie si $cardD < +\infty$.

Théorème 3.11 : *Pour tout $x \in H$, la série de Fourier de x est convergente dans H et sa somme est telle que*

$$(12) \qquad \left\| \sum_{n \in D} < x, e_n > e_n \right\|^2 = \sum_{n \in D} | < x, e_n > |^2.$$

Exercice : Soit $(\lambda_n)_{n \in D}$ une suite de nombres complexes. Montrer que la série $\sum_{n \in D} \lambda_n e_n$ est une série de vecteurs convergente dans H si et seulement si la série numérique $\sum_{n \in D} |\lambda|^2$ est convergente. Dans ce cas si $x = \sum_{n \in D} \lambda_n e_n$ est la somme de la série, quels sont les coefficients de Fourier de x et que vaut $||x||^2$?

3.4 Bases hilbertiennes

On a vu au paragraphe précédent que la série de Fourier d'un vecteur $x \in H$ est toujours convergente, mais quelle est la valeur de sa somme ? Nous allons aborder cette question dans ce paragraphe puis y répondre complétement dans le paragraphe 3.5. On suppose toujours qu'un système orthonormal $(e_n)_{n \in D}$ de H est donné.

Définition 3.12 : *On dit que $(e_n)_{n \in D}$ est un système total dans H si*

$$(13) \qquad \{e_n;\ n \in D\}^{\perp} = \{0\}.$$

Autrement dit si $x \in H$ est tel que $< x, e_n >= 0$ pour tout $n \in D$, alors $x = 0$.

Proposition 3.13 : *Si $\overline{V[e_n;\ n \in D]}$ est le sous espace fermé*[4] *engendré par les e_n $(n \in D)$, alors $(e_n)_{n \in D}$ est un système total si et seulement si*

$$(14) \qquad \left(\overline{V[e_n;\ n \in D]}\right)^{\perp} = \{0\}.$$

Théorème 3.14 : *Si $(e_n)_{n \in D}$ est un système orthonormal total dans H, alors*

$$(15) \qquad \forall x \in H,\ x = \sum_{n \in D} < x, e_n > e_n.$$

Autrement dit tout vecteur de H est somme de sa série de Fourier. De plus pour tous $x, y \in H$, on a l'égalité de Bessel-Parseval

$$(16) \qquad < x, y >= \sum_{n \in D} < x, e_n > \overline{< y, e_n >},$$

en particulier pour $x = y$, on obtient

$$(17) \qquad ||x||^2 = \sum_{n \in D} | < x, e_n > |^2.$$

Définition 3.15 : *Un système orthonormal total $(e_n)_{n \in D}$ de H est appelé aussi* **base hilbertienne** *de H.*

[4] On rappelle que $V := V[e_n;\ n \in D]$ est l'ensemble de toutes les combinaisons linéaires (finies) $\sum \lambda_k e_{n_k}$ et que $\overline{V[e_n;\ n \in D]}$ est l'adhérence de V dans H.

Remarque : Une base hibertienne $(e_n)_{n\in D}$ de H n'est une base algébrique de H que si $card D < +\infty$ i.e. si H est de dimension finie.

Exercice : Montrer que si $(e_n)_{n\in\mathbb{N}^*}$ est une base hilbertienne de H, le vecteur $\sum_{n=1}^{\infty} \frac{1}{n} e_n \in H$ ne peut pas s'écrire comme une combinaison linéaire (finie) des vecteurs e_n. En déduire que $(e_n)_{n\in\mathbb{N}^*}$ n'est pas une base de H au sens algébrique. (On remarquera que d'après l'exercice du paragraphe 3.3, la série $\sum_{n=1}^{\infty} \frac{1}{n} e_n$ est bien convergente dans H et elle représente donc un vecteur de H).

Exemple : Si $H = l^2$ (voir le chapitre 1), pour tout $n \in \mathbb{N}^*$ considérons l'élément $e^{(n)} = (e_k^{(n)})_{k\in\mathbb{N}^*} \in l^2$ tel que $e_k^{(n)} = 0$ si $k \neq n$ et $e_n^{(n)} = 1$ i.e. $e^{(n)}$ a toutes ses coordonnées nulles sauf la n-ième qui vaut 1. Le système $(e^{(n)})_{n\in\mathbb{N}^*}$ est une base hilbertienne de l^2 qu'on appelle parfois base hilbertienne canonique.

3.5 Cas général d'un système orthonormal quelconque

Soit $(e_n)_{n\in D}$ un système orthonormal dans H.

Remarque : on notera que $(e_n)_{n\in D}$ est toujours une base hilbertienne du sous espace vectoriel fermé de H qu'il engendre.

Théorème 3.16 *: Soit $V = \overline{V[e_n;\ n \in D]}$ le sous espace vectoriel fermé de H engendré par les e_n $(n \in D)$ et P_V l'opérateur de projection orthogonale sur V. Alors pour tout $x \in H$, on a*

$$(18)\qquad P_V(x) = \sum_{n\in D} < x, e_n > e_n \quad et \quad ||P_V(x)||^2 = \sum_{n\in D} |< x, e_n >|^2.$$

On notera que lorsque le système $(e_n)_{n\in D}$ est total, on a $V = H$, $P_V(x) = x$ et on retrouve le résultat du théorème 3.14. comme cas particulier.

Question pratique : Soit V un sous espace vectoriel fermé de H, comment déterminer $P_V(x)$?
Si $dim V = d < +\infty$, il suffit de déterminer une base orthonormale puis d'appliquer le théorème 3.7.
Si $dim V = +\infty$, il suffit de disposer d'un système orthonormal total dans V et d'appliquer le théorème 3.16. Ceci n'est possible que dans les espaces de Hilbert séparables comme on le verra dans le paragraphe 4.

Le procédé d'orthonormalisation de Gram-Schmidt : Soit $(b_n)_{n\in D}$ (où D est fini ou dénombrable), une famille libre de vecteurs de H et soit $V = V[b_n;\ n \in D]$ le sous espace vectoriel engendré par les b_n. Pour tout $k \in \mathbb{N}^*$, soit $V_k = V[b_1, \ldots, b_k]$ le sous espace engendré par les k premiers vecteurs. On considère alors la suite des vecteurs $(c_n)_{n\in D}$ obtenus de la manière suivante :

$$(19)\qquad c_1 = b_1, c_2 = b_2 - P_{V_1}(b_2), \ldots, c_k = b_k - P_{V_{k-1}}(b_k), \ldots (k \in \mathbb{N}^*).$$

Théorème 3.17 *: Les vecteurs $e_n = \frac{c_k}{||c_k||}$, $n \in \mathbb{N}^*$ forment un système orthonormal qui engendre V i.e. tel que*

$$(20)\qquad V = V[b_n;\ n \in D] = V[e_n;\ n \in D]$$

Corollaire 3.18 *: Soit $(b_n)_{n\in D}$ (où D est fini ou dénombrable), une famille libre de vecteurs de H. Alors la famille orthonormale $(e_n)_{n\in D}$ obtenue par le procédé d'orthonormalisation de Gram-Schmidt du théorème précédent, est une base hilbertienne de $\overline{V\,[b_n;\ n\in D]}$.*

3.6 Structure des espaces de Hilbert ayant une base hilbertienne

Dans cette partie nous abordons deux questions théoriques sur d'une part la structure algébrique et d'autre part la structure topologique d'un espace de Hilbert ayant une base hibertienne.

3.7 Le théorème d'isomorphisme

Définition 3.19 *: Deux espaces de Hilbert H et K sont dits isomorphes s'il existe une application linéaire $U : H \to K$ telle que :*
1) U est bijective.
2) U est unitaire i.e. conserve le produit scalaire :

$$\forall x, y \in H,\ < U(x), U(y) >_K = < x, y >_H, \tag{21}$$

ou, ce qui est équivalent, U conserve la norme :

$$\forall x \in H,\ ||U(x)||_K = ||x||_H. \tag{22}$$

Exercice : Démontrer l'équivalence affirmée dans la définition précédente i.e. une application linéaire $U : H \to K$ conserve le produit scalaire si et seulement si elle conserve la norme.

Théorème 3.20 *(théorème d'isomorphisme) : Soit H un espace de Hilbert ayant une base hilbertienne $(e_n)_{n\in D}$. Alors :*
-ou bien $cardD = d < +\infty$ et H est isomorphe à $\mathbb{C}^d$.
-ou bien $D = \mathbb{N}^*$ et H est isomorphe à l^2.

3.8 Espaces de Hilbert séparables

Définition 3.21 *: Un espace normé $(E, ||.||)$ est dit séparable s'il possède un sous ensemble dénombrable dense, c'est à dire s'il existe $\mathfrak{D} = \{b_n;\ n \in \mathbb{N}^*\} \subset E$ tel que*

$$\forall x \in E, \forall \epsilon > 0,\ \exists b \in \mathfrak{D} :\ ||x - b|| \le \epsilon. \tag{23}$$

Théorème 3.22 *: Un espace de Hilbert H possède une base hilbertienne si et seulement s'il est séparable.*

Pour la preuve de ce résultat, on utilise le lemme suivant :

Lemme 3.23 *: Si $\mathfrak{D} = \{b_n;\ n \in \mathbb{N}^*\}$ est un sous ensemble dénombrable dense d'un espace de Hilbert H, alors il existe une sous famille libre $\{b_{n_k};\ k \in D\}$ (finie ou dénombrable) de vecteurs de $\mathfrak{D}$ dont le sous espace fermé engendré est égal à H i.e. $\overline{V\,[b_{n_k};\ k \in D]} = H$.*

4 Exemple d'orthonormalisation et de base hilbertienne

Sur l'espace vectoriel $C([-1,1])$ des fonctions $f:[0,1]\to\mathbb{C}$ continues, muni du produit scalaire

$$(24)\qquad <f,g>=\int_{-1}^{1} f(t)\bar{g(t)}dt,$$

on considère la famille libre des fonctions monôme :

$$(25)\qquad b_n: t\mapsto t^n \quad (n\in\mathbb{N}).$$

Le sous-espace $V_k=V[b_0,b_1,\dots,b_k]$ est constitué des fonctions polynomiales de degré $\le k$ et $V=V[b_n;\ n\in\mathbb{N}]$ est le sous-espace de toutes les fonctions polynomiales. Le problème est d'orthonormaliser la famille $(b_n)_{n\in\mathbb{N}}$.
Avec les notations de (19), on calcule :
1) $c_0=b_0:t\mapsto 1$, $||c_0||^2=2$ et donc $e_0:t\mapsto\frac{1}{\sqrt{2}}$.
2) $c_1=b_1-P_{V_0}(b_1)$. Mais $<b_0,b_1>=0$ donc $P_{V_0}(b_1)=0$ et $c_1=b_1$, $||c_1||^2=2/3$, d'où $e_1:t\mapsto t\sqrt{2/3}$.
3) $c_2=b_2-P_{V_1}(b_2)$, $P_{V_1}(b_2)=<b_2,e_0>e_0+<b_2,e_1>e_1=\frac{\sqrt{2}}{3}e_0$ donc $c_2(t)=t^2-\frac{1}{3}$ et $e_2(t)=\sqrt{\frac{45}{8}}\left(t^2-\frac{1}{3}\right)$.
A ce stade on constate que

$$(26)\qquad e_1(t)=\frac{\sqrt{1+\frac{1}{2}}}{2^1.1!}\frac{d}{dt}(t^2-1),\quad e_2(t)=\frac{\sqrt{2+\frac{1}{2}}}{2^2.2!}\frac{d^2}{dt^2}(t^2-1)^2.$$

On fait l'hypothèse de récurrence que pour tout $k\le n$:

$$(27)\qquad e_k(t)=\frac{\sqrt{k+\frac{1}{2}}}{2^k.k!}\frac{d^k}{dt^k}(t^2-1)^k.$$

On vérifie assez facilement cette hypothèse en montrant que $e_{n+1}\perp V_n=V[e_0,e_1,\dots,e_n]$ et que $||e_{n+1}||=1$. Pour cela on utilise les propriétés suivantes des polynômes $P_n(t)=\frac{d^n}{dt^n}(t^2-1)^n$:
1) P_n est de degré n.
2) $\int_{-1}^{1}\frac{d^n}{dt^n}(t^2-1)^n\frac{d^m}{dt^m}(t^2-1)^m dt=(-1)^r\int_{-1}^{1}\frac{d^{n-r}}{dt^{n-r}}(t^2-1)^n\frac{d^{m+r}}{dt^{m+r}}(t^2-1)^m dt$.
3) $\int_{-1}^{1}P_n(t)P_m(t)dt=0$ si $n\neq m$.
4) $\int_{-1}^{1}P_n^2(t)dt=\frac{2}{2n+1}$.
Les polynômes e_n donnés par (27) s'appellent les polynômes de Legendre (normalisés par la condition $||e_n||_2=1$). Ils constituent une famille orthonormale de $C([-1,1])$ telle que $V=V[b_n;\ n\in\mathbb{N}]=V[e_n;\ n\in\mathbb{N}]$. Mais l'espace $C([-1,1])$ est seulement un espace préhilbertien[5]. On sait d'après le cours d'intégration que $C([-1,1])$ est dense dans l'espace $L^2([-1,+1])$ des classes de fonctions de carré intégrable sur $[-1,1]$ pour la mesure de Lebesgue[6].

Théorème 4.1 *: Les polynômes de Legendre forment une base hilbertienne de l'espace de Hilbert $L^2([-1,+1])$.*

---

[5]voir le chapitre 1.

[6]En fait $L^2([-1,+1])$ est le complété de $C([-1,1])$ pour la norme $||.||_2$ associée au produit scalaire (24).

Pour la démonstration il suffit de montrer que l'espace $V[e_n;\ n \in \mathbb{N}]$ est dense dans $C([-1,1])$ pour la norme $||.||_2$. Mais comme la norme $||.||_2$ est moins fine que la norme $||.||_\infty$ de la convergence uniforme puisque pour $f \in C([-1,1])$, on a

$$||f||_2^2 = \int_{-1}^{1} |f(t)|^2 dt \leq \int_{-1}^{1} \sup_{t\in[0,1]} |f(t)|^2 dt = 2||f||_\infty^2, \tag{28}$$

il suffit de savoir que les polynômes sont denses dans $C([-1,1])$ pour la convergence uniforme. Mais ce résultat est bien connu ; c'est le théorème de Weierstrass :

Théorème 4.2 *: Pour toute fonction f continue sur un intervalle compact $[a,b]$, il existe une suite (P_n) de fonctions polynôme qui converge uniformément vers f sur $[a,b]$.*

Pour la démonstration de ce résultat, on peut supposer que $[a,b]=[0,1]$ (faire le changement de variable $t = a + t'(b-a)$, avec $t' \in [0,1]$) et on montre alors le résultat plus explicite suivant :

Théorème 4.3 *(Bernstein) : Si $f \in C([0,1])$, la suite P_n des polynômes tels que*

$$P_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}, \tag{29}$$

converge uniformément vers f sur $[0,1]$ quand $n \to +\infty$.

Le seul point délicat dans la preuve de Bernstein est l'estimation suivante

Lemme 4.4 *: Soit $x \in [0,1]$ et pour tout $\alpha > 0$, soit I_α l'ensemble des entiers $k \in \{0,1,\ldots,n\}$ tels que $|k-nx| > n\alpha$. Alors*

$$\sum_{k\in I_\alpha} C_n^k x^k (1-x)^{n-k} \leq \frac{x(1-x)}{\alpha^2 n}. \tag{30}$$

Ce résultat est facile si on connait la loi binomiale. En effet soit X une variable aléatoire de loi binomiale $\mathcal{B}(n,x)$. Alors puisque $\mathbb{E}(X) = nx$, la quantité du membre de gauche dans (30) est justement la probabilité $\mathbb{P}(|X-\mathbb{E}(X)| > \alpha n)$. Or l'inégalité de Bienaymé-Tchebychev donne justement

$$\mathbb{P}(|X-\mathbb{E}(X)| > \alpha n) \leq \frac{VarX}{(\alpha n)^2} = \frac{nx(1-x)}{\alpha^2 n^2}. \tag{31}$$

C'est l'inégalité du lemme.