



N° d'ordre :

N° de séries :

Mémoire de fin d'études
Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : EDP et applications.

Thème

**Autour de la projection et la méthode des
projections alternées**

Présenté par :

Bouedjourn Fouzia

Djafri Bessma

Devant le jury :

Président : T.Haddad Prof. Université de Jijel

Encadreur : I.Kecis M.C.B Université de Jijel

Examineur : I.Touil M.C.B Université de Jijel

Examineur : I.Soualhia M.C.B Université de Jijel

Table des matières

Introduction	ii
1 Préliminaires	1
1.1 Notations	1
1.2 Définitions et premières propriétés	2
1.3 Quelques rappels d'analyse convexe	5
1.4 Topologie faible	9
2 Autour de la projection et ses caractérisations	12
2.1 Projection : définitions et propriétés générales	12
2.2 Caractérisations de la Projection sur différents types d'ensembles	19
2.2.1 Projection sur un convexe fermé	21
2.2.2 Projection sur un sous espace vectoriel fermé	24
2.2.3 Projection sur un cône convexe fermé	27
3 Méthode des projections alternées MPA	32
3.1 Cas de deux sous espaces vectoriels	32
3.1.1 Taux de convergence de MPA	35
3.2 Cas de plusieurs ensembles convexes fermés : algorithme de Dykstra	39
3.3 Application : Systèmes linéaires	49

Table des figures

1.1	boule unité $\ \cdot\ _2$	4
1.2	boule unité $\ \cdot\ _1$	4
1.3	boule unité $\ \cdot\ _\infty$	4
1.4	6
2.1	Graphes de $P_C(\cdot)$ avec $C = [-1, 1] \cup [2, 3]$	14
2.2	Projection par rapport à la norme $\ \cdot\ _1$	14
2.3	Projection par rapport à la norme $\ \cdot\ _2$	15
2.4	Projection par rapport à la norme $\ \cdot\ _\infty$	15
2.5	19
2.6	Projection sur un ensemble non convexe : $\langle x - y, z - y \rangle$ peut être positif	20
2.7	interprétation géométrique de la caractérisation (2.8)	22
2.8	Caractérisation de la projection sur un cône convexe	27
3.1	Algorithme des projections alternées MAP	33
3.2	37
3.3	Algorithme de MAP ne converge pas vers le point de projection	40
3.4	Algorithme de Dykstra avec deux ensembles convexes ($r = 2$)	42

Introduction

Un problème de l'admissibilité ou faisabilité convexe, the convex feasibility problem en anglais, s'écrit sous la forme suivante

$$\text{Trouver si possible au moins un } \bar{x} \in \bigcap_{i=1}^N C_i \quad (1)$$

Où $C_i, i = 1, \dots, N$ sont des sous ensembles fermés d'un espace de Hilbert H . Les problèmes de type (1) ont des domaines d'applications variés en mathématique, citons par exemple

- Inégalités variationnelles [6], Équations multivoques d'évolution [16].
- Meilleure approximation et problèmes de minimisation (optimisation) [10, 18].
- Équations aux dérivées partielles, les solutions d'inégalités convexes [12].

Dans le cas classique, les ensembles C_i sont des sous-espaces vectoriels fermés ou des ensembles convexes fermés. Dans ce cadre, différentes méthodes itératives ont été développées pour la résolution du problème (1). La nature itérative de ces méthodes les rend indispensables pour résoudre des problèmes à grande échelle et hautement non linéaires. Ce sont des cas où les méthodes d'optimisation directes sont intraitables.

Le prototype des algorithmes pour résoudre le problème (1) est la méthode des projections alternées, MPA en abrégé. Cette méthode a été introduite pour la première fois par Von Neumann [19] en 1933 pour trouver la projection d'un point donné sur l'intersection de deux sous-espaces vectoriels fermés d'un espace de Hilbert. La philosophie de cette méthode se base sur le fait que le calcul des projections sur chaque ensemble individuellement est plus facile que directement sur l'intersection. Depuis son introduction, cette méthode a reçu une attention considérable et a trouvé des applications dans de nombreux domaines des mathématiques et de la physique ainsi que dans d'autres domaines de la science et de l'ingénierie. Cet algorithme repose sur la construction d'une suite de projections successives sur les ensembles considérés, à partir d'un point initial $x \in H$ comme suit

$$x_0 := x, \quad x_n = \left(P_{C_N} P_{C_{N-1}} \cdots P_{C_1} \right)^n (x_0) \quad (2)$$

Avec deux sous-espaces vectoriels fermés C_1 et C_2 , Von Neumann a démontré que la suite $(x_n)_n$ générée par relation (2) converge vers le point $P_{C_1 \cap C_2}(x_0)$. Le passage de 2 à N sous espaces vectoriels a été fait par Halperin dans [14] avec les mêmes résultats sur la convergence.

Le changement des hypothèses en remplaçant les sous espaces vectoriels C_i par des ensembles convexes fermés change complètement les propriétés de la convergence. D'un coté, on perd la convergence forte (dans la dimension infinie) de la suite de Von Neumann vers le point de l'intersection des C_i (Théorème de Bregman [4]), et d'un autre coté, on perd la convergence (même dans la dimension finie) vers le point de projection $P_{\bigcap_{i=1}^N C_i}(x)$.

Heureusement, Dykstra [9] a trouvé une modification intelligente du schéma de Von Neumann pour lequel la convergence vers le point de l'intersection le plus proche au point initial est garantie. La méthode a été développée à l'origine pour projeter un point donné dans un espace de dimension finie sur un cône convexe fermé.

On peut décrire le matériel présenté dans ce travail en trois chapitres, on commence par un chapitre introductif qui rappelle et présente les résultats fondamentaux et les concepts de base que l'on va utiliser dans les autres chapitres, notamment les notions de l'analyse convexe et de la topologie faible.

Le deuxième chapitre est consacré à la notion fondamentale de projection. On va s'intéresser à étudier la fonction de projection tout en donnant les propriétés majeures de cette application avec différents types d'ensembles (convexes, s.e.v, cônes convexes), notamment ses caractérisations qui donnent des définitions alternatives plus simples et élégantes que l'on utilisera dans le dernier chapitre.

Le dernier chapitre est dévolu à l'étude de la méthode des projection alternées en considérant deux types d'ensembles : l'intersection des sous espaces vectoriels et celle des ensembles convexes fermés. Dans la première partie, on présentera l'algorithme de Von Neumann ainsi que ses résultats de convergence. Cette présentation sera suivie par une étude sur la vitesse ou le taux de convergence qui sera décrit en terme d'angle formé par les sous espaces considérés. On traitera dans la deuxième partie, l'algorithme de Dykstra en donnant sa définition, en suite, on va démontrer le résultat de convergence de cette méthode. On terminera par une application qui utilise la méthode des projection alternées pour la résolution d'un système linéaire.

Préliminaires

Ce chapitre introductif a pour but de rappeler les résultats fondamentaux et les concepts de base que l'on va utiliser dans les autres chapitres. Selon les problèmes qui se présentent au cours de ce travail, on est amené à utiliser quelques résultats et outils auxiliaires que l'on accepte (souvent) sans démonstration, pour plus d'informations on se réfère à [5, 11, 17]. Dans tout ce qui suit, sauf mention contraire, X désigne un espace vectoriel sur \mathbb{R} et X^* représente son dual topologique formé de toutes les formes linéaires $x^* \in X : X \rightarrow \mathbb{R}$

1.1 Notations

- s.e.v : sous espace vectoriel .
- c-à-d : c'est à dire.
- resp. : respectivement.
- \mathbb{R} l'ensemble des nombres réels.
- $\overline{\mathbb{R}} = [-\infty, +\infty]$.
- min, max, inf, sup minimum, maximum, infimum et supremum respectivement.
- H un espace de Hilbert.
- X^* dual topologique d'un espace normé X .
- $\overline{\mathbb{B}}_X$ la boule unité fermée de X .
- $\langle \cdot, \cdot \rangle$ produit scalaire sur H .
- $L^p(0, T; H) := \{f : [0, T] \rightarrow H \text{ mesurable} : \int_0^T \|f(t)\|^p < +\infty\}$.
- $L^\infty(0, T; H) = \{f : [0, T] \rightarrow H \text{ mesurable} : \exists C \geq 0, \|f(t)\| \leq C \text{ p.p } t \in [0, T]\}$.
- $\text{int}(A), \overline{A}$: l'intérieur et l'adhérence d'un ensemble A respectivement.
- C° cône polaire
- $N_C(x)$ cône normal à C au point x
- $\ker f := \{x \in X : f(x) = 0\}$: le noyau de f .

- $x_n \downarrow x$: une suite $(x_n)_n$ décroissante et converge vers un élément x .
- \rightharpoonup ou \xrightarrow{w} la convergence faible.
- \rightarrow ou $\xrightarrow{\|\cdot\|}$ la convergence forte.
- $P_C(x)$ la projection de x sur C
- $\sigma(X, X^*)$ la topologie faible sur X .
- $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées sur \mathbb{R} .
- $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices carrées symétriques.
- $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrices carrées symétriques semi définies positive.
- $\mathcal{S}_n^-(\mathbb{R})$ l'ensemble des matrices carrées symétriques semi définies négative

1.2 Définitions et premières propriétés

Définition 1.1. On appelle produit scalaire sur X , toute forme bilinéaire $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$ possédant les propriétés supplémentaires suivantes

$$\begin{aligned} \langle x, y \rangle &\geq 0 \text{ (positivité)} \\ \langle x, x \rangle &= 0 \Leftrightarrow x = 0 \\ \langle x, y \rangle &= \langle y, x \rangle \text{ (symétrie)} \end{aligned}$$

Exemple 1.2.

1. Soit $n \in \mathbb{N}$, on note $l_2(n)$ l'ensemble de fonctions

$$\begin{aligned} x(\cdot) : \{1, 2, \dots, n\} &\rightarrow \mathbb{R} \\ i &\mapsto x(i) \end{aligned} \tag{1.1}$$

alors l'application

$$\begin{aligned} \langle \cdot, \cdot \rangle : l_2(n) &\rightarrow \mathbb{R} \\ x, y &\mapsto \langle x, y \rangle = \sum_{i=1}^n x(i)y(i) \end{aligned}$$

est un produit scalaire dont la norme induite est

$$\|x\| = \left(\sum_{i=1}^n |x(i)|^2 \right)^{\frac{1}{2}} \tag{1.2}$$

2. L'espace $L^2(\Omega)$ de fonction carrée intégrables est un espace de Hilbert dont le produit scalaire est

$$\langle f, g \rangle_{L^2} = \int_{\Omega} f(x)g(x)dx$$

et la norme associée est

$$\|f\|_{L^2} = \left(\int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}$$

Le produit scalaire permet de mesurer l'angle formé par deux vecteurs x et y via l'équivalence suivante

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \|y\|} \Leftrightarrow \theta = \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

Cette dernière expression est bien définie grâce à l'inégalité de Cauchy-Schwartz qui donne

$$|\langle x, y \rangle| \leq \|x\| \|y\| \Rightarrow |\cos(\theta)| = \frac{|\langle x, y \rangle|}{\|x\| \|y\|} \leq 1.$$

L'application $\|\cdot\| : x \in X \mapsto \|x\| = \sqrt{\langle x, x \rangle}$ est une norme sur X appelée la norme hilbertien ou la norme associée ou induite par le produit scalaire $\langle \cdot, \cdot \rangle$. L'espace X muni du produit scalaire $\langle \cdot, \cdot \rangle$ s'appelle un espace pré-hilbertien. Si de plus, X est complet pour la norme induite par $\langle \cdot, \cdot \rangle$ alors dans ce cas, $(X, \langle \cdot, \cdot \rangle)$ est dit un espace de Hilbert que l'on note généralement H .

Sur \mathbb{R}^n , les normes les plus couramment utilisées sont

$$\begin{aligned} \|x\|_1 &:= \sum_{i=1}^n |x_i| \\ \|x\|_2 &:= \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}, \text{ (norme euclidienne)} \\ \|x\|_\infty &:= \max_{1 \leq i \leq n} |x_i|, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n \end{aligned}$$

Il est à noter que les propriétés de deux espaces définis sur le même ensemble et munis de deux normes différentes seront bien différentes. Il convient alors de distinguer soigneusement les espaces normés en ajoutant la norme utilisée ($(\mathbb{R}^n, \|\cdot\|_1)$, $(\mathbb{R}^n, \|\cdot\|_2)$, ...etc).

Définition 1.3. *Étant donné un point $x_0 \in \mathbb{R}^n$ et un réel $r > 0$, les sous ensembles de X notés*

$$\begin{aligned} \mathbb{B}[x_0, r] &:= \{x \in X : \|x - x_0\| \leq r\} \\ \mathbb{B}(x_0, r) &:= \{x \in X : \|x - x_0\| < r\} \\ \mathbb{S}[x_0, r] &:= \{x \in X : \|x - x_0\| = r\} \end{aligned}$$

sont appelés, respectivement

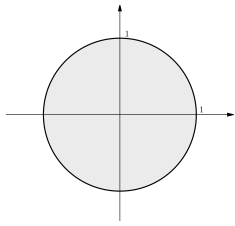
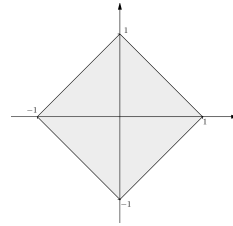
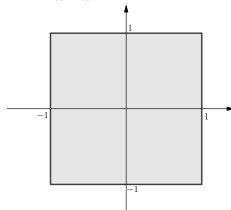
boule fermée de centre x_0 et de rayon r

boule ouverte de centre x_0 et de rayon r

sphère de centre x_0 et de rayon r

En particulier, si $r = 1$ et $x_0 = 0$ alors $\overline{\mathbb{B}}_X$ (resp. \mathbb{B}_X) représente la boule unité fermée (resp. Ouverte). Il est clair qu'une sphère peut être vide

Les représentations graphiques des boules unités associées aux trois normes définies ci-dessus sont données par les Figures 1.3, 1.1 et 1.3.

FIGURE 1.1 – boule unité $\|\cdot\|_2$ FIGURE 1.2 – boule unité $\|\cdot\|_1$ FIGURE 1.3 – boule unité $\|\cdot\|_\infty$

Définition 1.4 (Identité du parallélogramme). Soit $\|\cdot\|$ une norme sur X alors, $\|\cdot\|$ vérifie l'identité du parallélogramme si

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Cette dernière égalité est vérifiée dans chaque espace de Hilbert dont la norme est celle associée au produit scalaire.

Définition 1.5.

- (a) Deux vecteurs $x, y \in H$ sont dits orthogonaux et on écrit $x \perp y$ si $\langle x, y \rangle = 0$. Cette relation est bien sûr symétrique.
- (b) Deux ensembles $A, B \subset H$ sont orthogonaux et on écrit $A \perp B$ si $\langle x, y \rangle = 0$ pour tous $x \in A$ et $y \in B$.
- (c) Soit $A \subset H$, l'orthogonale de A que l'on note A^\perp est l'ensemble défini par

$$A^\perp = \{x \in H : \langle x, y \rangle = 0, \forall y \in A\} \quad (1.3)$$

Le vecteur nul est orthogonal à tous les vecteurs de H , ceci entraîne que si $\langle x, y \rangle = 0$ pour tout $y \in H$ alors $x = 0$.

Lemme 1.6. Soit F un sous-espace vectoriel (s.e.v) fermé de H alors $(F^\perp)^\perp = F$. Si F n'est pas nécessairement fermé on a $(F^\perp)^\perp = \overline{F}$.

Proposition 1.7. Si $F \subset H$ est un sous espace vectoriel alors F^\perp l'est aussi, de plus il est fermé.

Preuve. Tout d'abord, il est clair que $0 \in F^\perp$ ce qui implique la non-vacuité de F^\perp .

- (•) Soient $x, y \in F^\perp$ et $\alpha, \beta \in \mathbb{R}$ alors pour tout $z \in F$

$$\langle \alpha x + \beta y, z \rangle = \alpha \underbrace{\langle x, z \rangle}_{=0} + \beta \underbrace{\langle y, z \rangle}_{=0} = 0 \Rightarrow \alpha x + \beta y \in F^\perp$$

- (•) F^\perp Soit $(x_n)_n \subset F^\perp$ une suite converge vers un élément $x \in H$, alors pour tout $y \in H$

$$\begin{aligned} \langle x_n, y \rangle = 0, \forall n \geq 0 &\Rightarrow \lim_{n \rightarrow \infty} \langle x_n, y \rangle = 0 \Rightarrow \langle \lim_{n \rightarrow \infty} x_n, y \rangle = 0 \\ &\Rightarrow \langle x, y \rangle = 0, \forall y \in H \\ &\Leftrightarrow x \in F^\perp \end{aligned}$$

ce qui exprime la fermeture de F^\perp . □

1.3 Quelques rappels d'analyse convexe

Définition 1.8.

1. On appelle *segment joignant* ou *reliant* deux points $x, y \in X$ l'ensemble

$$[x, y] := \{\lambda x + (1 - \lambda)y, \lambda \in [0, 1]\}$$

2. Un sous ensemble $C \subset X$ est dit *convexe* si pour tous $x, y \in C$ le segment reliant x à y reste dans C c-à-d,

$$\lambda x + (1 - \lambda)y \in C \text{ pour tout } \lambda \in [0, 1].$$

3. Une fonction $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ est dite *convexe* si, pour tous $x, y \in \text{dom } f$ et pour tout $\lambda \in [0, 1]$ l'inégalité suivante, dite "*l'inégalité de convexité*" est vérifiée

$$\forall \lambda \in [0, 1] : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

où $\text{dom}(f)$ est le domaine de f défini par

$$\text{dom}(f) = \{x \in X, f(x) < +\infty\}.$$

la fonction f est dite *strictement convexe* si l'inégalité est stricte. La définition de la convexité signifie que le segment reliant les deux points $(x, f(x))$ et $(y, f(y))$ se trouve toujours au dessus de la courbe de f .

4. On appelle *épigraphe* de $f : X \rightarrow \mathbb{R}$ le sous-ensemble de $X \times \mathbb{R}$ noté et défini comme suit

$$\text{epi}(f) := \{(x, t) \in X \times \mathbb{R} / f(x) \leq t\}$$

Exemple 1.9.

1. Si $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ est convexe alors les ensembles de sous niveau $[f \leq r] := \{x \in X : f(x) \leq r\}, r \in \mathbb{R}$ sont convexes, mais ceci ne caractérise pas les fonctions convexes (penser à la fonction $f : x \in \mathbb{R} \mapsto \sqrt{|x|}$ représentée dans la Figure 1.4).

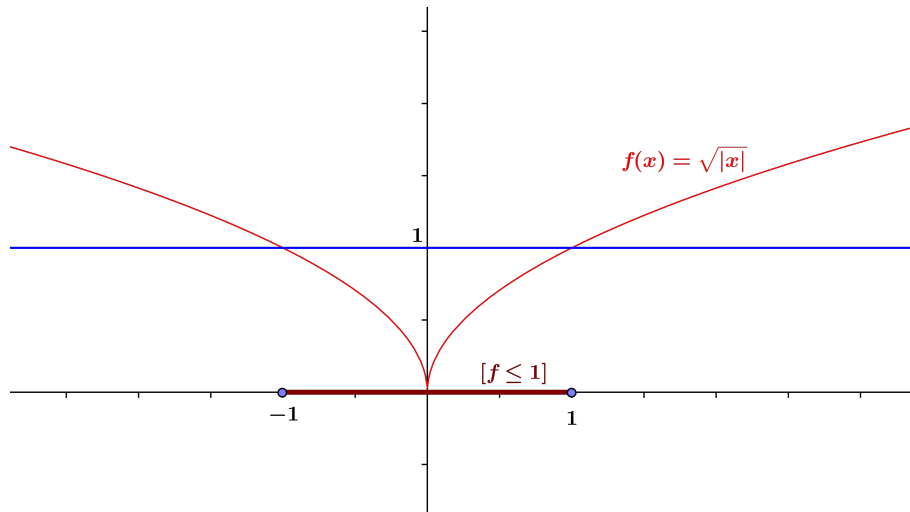


FIGURE 1.4 –

2. Les boules $\mathbb{B}[x_0, r]$ et $\mathbb{B}(x_0, r)$ sont des ensembles convexe pour tous $x_0 \in X, r \in \mathbb{R}_+$.
3. Tous sous espace vectoriel de X est convexe.
4. La sphère \mathbb{S} n'est pas convexe. en effet, soient $a := (-1, 0), b := (0, 1) \in \mathbb{S}$ alors

$$\frac{1}{2}(-1, 0) + \frac{1}{2}(0, 1) = (0, 0) \notin \mathbb{S}.$$

5. La fonction $f : x \mapsto x^2$ est convexe, en effet, soient $x, y \in \mathbb{R}$ et $\lambda \in [0, 1]$ alors

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) - f(\lambda x + (1 - \lambda)y) &= \lambda x^2 + (1 - \lambda)y^2 - \lambda^2 x^2 \\ &\quad - 2\lambda(1 - \lambda)xy - (1 - \lambda)^2 y^2 \\ &= \lambda(1 - \lambda)(x^2 - 2xy + y^2) \\ &= \lambda(1 - \lambda)(x - y)^2 \geq 0 \end{aligned}$$

Théorème 1.10. *L'épigraphe nous permet de lier la convexité des fonctions à celle des ensembles par l'équivalence suivante*

f est convexe si et seulement si $\text{epi}(f)$ est convexe dans $X \times \mathbb{R}$.

Preuve. \Rightarrow) Supposons que f soit convexe alors, pour tous $(x_1, r_1), (x_2, r_2) \in \text{epi}(f)$ et $\lambda \in [0, 1]$

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &\leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda r_1 + (1 - \lambda)r_2 \\ \Rightarrow (\lambda x_1 + (1 - \lambda)x_2, \lambda r_1 + (1 - \lambda)r_2) &\in \text{epi}(f) \\ \Leftrightarrow \lambda(x_1, r_1) + (1 - \lambda)(x_2, r_2) &\in \text{epi}(f) \\ \Leftrightarrow \text{epi}(f) \text{ est convexe.} \end{aligned}$$

\Leftarrow) On suppose que $\text{epi}(f)$ soit un ensemble convexe de $X \times \mathbb{R}$, alors pour tous $(x_1, r_1), (x_2, r_2) \in \text{epi}(f)$ et $\lambda \in [0, 1]$

$$\begin{aligned} \lambda(x_1, r_1) + (1 - \lambda)(x_2, r_2) \in \text{epi}(f) &\Rightarrow (\lambda x_1 + (1 - \lambda)x_2, \lambda r_1 + (1 - \lambda)r_2) \in \text{epi}(f) \\ &\Rightarrow f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda r_1 + (1 - \lambda)r_2 \end{aligned}$$

en particulier, pour $(x_i, r_i) = (x_i, f(x_i)) \in \text{epi}(f), i = 1, 2$, on déduit la convexité de f . \square

Proposition 1.11. *L'intersection quelconque d'ensembles convexes est convexe.*

Preuve. Soit $(C_i)_{i \in I}$ une famille quelconque de convexes, posons $C = \bigcap_{i \in I} C_i$. Si $C = \emptyset$ on a rien à prouver car l'ensemble vide est convexe sinon, soient $x, y \in C$, alors $x, y \in C_i$, pour tout $i \in I$, comme C_i est convexe alors pour tout $i \in I$, alors

$$\forall \lambda \in [0, 1] : \lambda x + (1 - \lambda)y \in C_i, \forall i \in I \Rightarrow \lambda x + (1 - \lambda)y \in C.$$

\square

Pour plus d'information sur la théorie des fonctions convexes vous pouvez consulter les références[1, 15]

Remarque 1.12. *L'union d'ensembles convexes n'est pas nécessairement convexe en effet, Si on prend $C = [0, 1] \cup [2, 3]$ alors pour $x_1 = 1, x_2 = 2 \in C$ et $\lambda = \frac{1}{2}$ on obtient*

$$\lambda x_1 + (1 - \lambda)x_2 = \frac{1}{2} + 1 = \frac{3}{2} \notin C.$$

Définition 1.13.

- Un sous ensemble $C \subset X$ est dit un cône si $\lambda C \subset C$ pour tout $\lambda \geq 0$; autrement dit, C est invariant par l'homothétie de centre 0 et de rapport $\lambda \in \mathbb{R}_+$. Si de plus C est convexe alors C est dit cône convexe, il est clair qu'un cône contient toujours l'origine ($\lambda = 0 : 0C = \{0\} \subset C$).
- Le cône polaire (ou cône polaire négatif, ou cône dual) de $C \subset X$ est l'ensemble

$$C^\circ := \{y \in X^*, \langle y, x \rangle \leq 0 \text{ pour tout } x \in C\},$$

d'autres notations sont également utilisées pour le cône polaire de $C : C^-, C^\ominus \dots$ etc.

Exemple 1.14.

1. Cône de Pareto de \mathbb{R}^n (orthant positif)

$$C := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n, x_i \geq 0 \text{ pour tout } i = 1, \dots, n\} = \mathbb{R}_+^n$$

Alors

$$C^\circ = -C = \{-x = (-x_1, \dots, -x_n) \in \mathbb{R}^n\} = \mathbb{R}_-^n$$

2. Cône des matrices symétriques semi définies positives (SDP) :

Sur $\mathcal{S}_n(\mathbb{R}) := \{A \in \mathcal{M}_n(\mathbb{R}) : A \text{ symétrique}\}$, on définit le produit scalaire

$$\begin{aligned} \langle\langle \cdot, \cdot \rangle\rangle: \mathcal{S}_n^2(\mathbb{R}) &\longrightarrow \mathbb{R} \\ A, B &\mapsto \langle\langle A, B \rangle\rangle = \text{trace}(AB) \end{aligned}$$

si on pose

$$C = \mathcal{S}_n^+(\mathbb{R}) := \{A \in \mathcal{S}_n(\mathbb{R}) : A \text{ semi-définie positive}\}$$

alors

$$C^\circ = \mathcal{S}_n^-(\mathbb{R}) := \{B \in \mathcal{S}_n(\mathbb{R}) : B \text{ semi-définie négative}\}.$$

Dans l'analyse convexe, les cônes les plus connus sont le cône tangente et le cône normal. Soit $x \in C$, alors

$$T_C(x) := \overline{\mathbb{R}_+ \{C - x\}}$$

est le cône tangente à C au point x , et

$$N_C(x) := (T_C(x))^\circ \subset X^*$$

est le cône normal à C au point x .

- La définition de $T_C(x)$, nous permet d'écrire

$$N_C(x) = \{x^* \in X^*, \langle x^*, y - x \rangle \leq 0, \text{ pour tout } y \in C\} \quad (1.4)$$

Il est facile de vérifier que si $x \in \text{int}(C)$ avec $\text{int}(C) \neq \emptyset$ alors $N_C(x) = \{0\}$. En effet, soit $x \in \text{int}(C)$, alors il existe $\gamma > 0$ tel que $\mathbb{B}(x, \gamma) \subset C$. Soient $v \in N_C(x)$ et $t > 0$ suffisamment petit, tel que $x + tv \in \mathbb{B}(x, \gamma) \subset C$ alors

$$\begin{aligned} \langle v, x + tv - x \rangle \leq 0 &\Rightarrow \langle v, tv \rangle \leq 0 \Rightarrow t\|v\|^2 \leq 0 \Rightarrow v = 0 \\ &\Rightarrow N_C(v) \subset \{0\} \Rightarrow N_C(v) = \{0\} \end{aligned}$$

Exemple 1.15.

Sur \mathbb{R} on considère le sous ensemble $C = [0, 1]$ alors

$$T_{[0,1]}(1) =] - \infty, 0], \quad N_{[0,1]}(1) = [0, +\infty[.$$

En effet

$$\begin{aligned} C - \{1\} &= \{x - 1, x \in [0, 1]\} = [-1, 0] \\ &\Rightarrow \mathbb{R}_+(C - \{1\}) = \{\lambda x, x \in C - \{1\}\} \\ &= \{\lambda x, x \in [-1, 0]\} =] - \infty, 0] \\ &\Rightarrow T_{[0,1]}(1) = \overline{\mathbb{R}_+(C - \{1\})} = \overline{] - \infty, 0]} =] - \infty, 0]. \end{aligned}$$

$$N_{[0,1]}(1) = (T_{[0,1]}(1))^\circ = \{v \in \mathbb{R} : v(y - 1) \leq 0, \forall y \in [0, 1]\} = [0, +\infty[$$

1.4 Topologie faible

On appelle topologie faible sur X que l'on note $\sigma(X, X^*)$ la topologie la moins fine rendant toutes les applications $f \in X^* : X \rightarrow \mathbb{R}$ continues. On peut recenser les ouverts qui doivent appartenir à la topologie faible $\sigma(X, X^*)$ de la manière suivante : si $f \in X^*$ et $U \in \mathbb{R}$ est un ouvert, il faut que $f^{-1}(U)$ soit un ouvert de $\sigma(X, X^*)$. Mais comme les intervalles représentent une base de voisinage de \mathbb{R} , on voit que ceci revient à dire que pour tout intervalle I et tout $f \in X^* : f^{-1}(I) \in \sigma(X, X^*)$.

Proposition 1.16. *La topologie $\sigma(X, X^*)$ est la topologie la moins fine contenant tous les ensembles $f^{-1}(I)$ pour tout $f \in X^*$ et tout intervalle I de \mathbb{R} .*

Définition 1.17. (*Proposition*)

Soit $(x_n)_n$ une suite dans un espace normé X . On dit que la suite $(x_n)_n$ converge faiblement vers un $x \in X$ et on écrit $x_n \rightharpoonup x$ ou bien $x = w - \lim_{n \rightarrow +\infty} x_n$ ou bien $x_n \xrightarrow{w} x$ si et seulement si

$$\lim_{n \rightarrow +\infty} \langle f, x_n \rangle = \langle f, x \rangle \text{ pour tout } f \in X^*.$$

Notre tâche maintenant consiste à relier la notion de la convergence faible avec un nouveau concept de limite inférieure. Pour cela nous allons introduire la notions de la limite inférieure (resp. supérieure) pour une suite ordinaire de $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$.

Définition 1.18. *Soit $(x_n)_n$ une suite de $\overline{\mathbb{R}}$. On définit les limites inférieure et supérieure de $(x_n)_n$ comme suit*

$$\begin{aligned} \liminf_{n \rightarrow \infty} x_n &= \sup_{n \in \mathbb{N}} \left(\inf_{k \geq n} x_k \right) \\ \limsup_{n \rightarrow \infty} x_n &= \inf_{n \in \mathbb{N}} \left(\sup_{k \geq n} x_k \right) \end{aligned}$$

Ces deux semi-limites vérifient les inégalités suivantes

$$\liminf x_n \leq \lim x_n \leq \limsup x_n$$

Exemple 1.19.

Considérons la suite $(x_n)_{n \geq 0}$ définie comme suit

$$x_n = (-1)^n \left(1 + \frac{1}{n+1} \right) = \left\{ 2, \frac{-3}{2}, \frac{4}{3}, \frac{-5}{4}, \frac{6}{5}, \frac{-7}{6}, \dots \right\},$$

il est clair que $\inf x_n = \min x_n = \frac{-3}{2}$ et $\sup x_n = \max x_n = 2$. Posons

$$y_n := \inf_{k \geq n} x_k, \quad z_n := \sup_{k \geq n} x_k$$

alors

$$\begin{aligned} y_0 &= \frac{-3}{2}, y_1 = \frac{-3}{2}, y_2 = \frac{-5}{4}, \dots, y_n = -\left(1 + \frac{1}{2n}\right) \\ z_0 &= 2, z_1 = \frac{4}{3}, z_2 = \frac{4}{3}, \dots, z_n = 1 + \frac{1}{2n+1}. \end{aligned}$$

La suite $(y_n)_n$ (resp. $(z_n)_n$) est croissante (resp. décroissante) alors

$$\begin{aligned}\limsup_n x_n &= \inf_n \sup_{k \geq n} x_k = \lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2n+1}\right) = 1 \\ \liminf_n x_n &= \sup_n \inf_{k \geq n} x_k = \lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} -\left(1 + \frac{1}{2n}\right) = -1.\end{aligned}$$

Lemme 1.20. La limite inférieure d'une suite $(x_n)_n$ peut être caractérisée par l'égalité suivante

$$\liminf_n x_n = \min\left\{\lim_j x_{n_j}, (x_{n_j})_j \text{ est une sous suite convergente de } (x_n)_n\right\}.$$

autrement dit, $\liminf_n x_n$ est le minimum de l'ensemble des limites de sous-suites convergentes de $(x_n)_n$.

Proposition 1.21. Soient $(x_n)_n$ et $(f_n)_n$ deux suites d'un espace normé X et son dual X^* respectivement, les assertions suivantes ont lieu

- (a) La convergence forte implique la convergence faible, c-à-d, si $x_n \rightarrow x$ alors $x_n \rightharpoonup x$.
La réciproque est fausse.
- (b) Si $x_n \xrightarrow{w} x$ alors la suite $(\|x_n\|)_n$ est bornée et

$$\|x\| \leq \liminf_{n \rightarrow +\infty} \|x_n\|. \quad (1.5)$$

- (c) Si $x_n \rightharpoonup x$ et f_n converge fortement vers f dans X^* alors

$$\langle f_n, x_n \rangle \rightarrow \langle f, x \rangle.$$

Lemme 1.22 (Kadec-Klee). Soit $(x_n)_n$ une suite de H telle que

$$x_n \xrightarrow{w} x \text{ et } \|x_n\| \rightarrow \|x\|$$

alors (x_n) converge fortement vers x .

Théorème 1.23.

1. Tout ensemble faiblement fermé est fermé
2. Un ensemble convexe est fermé si et seulement s'il est faiblement fermé

On sait bien que dans un espace normé général (de dimension infinie) on perd la compacité forte de la boule unité fermé, ce qui n'est pas le cas avec la compacité faible.

Théorème 1.24. (Banach-Alaoglu)

Soit H un espace de Hilbert, alors La boule unité fermé $\overline{\mathbb{B}}_H = \{x \in H : \|x\| \leq 1\}$ est faiblement compacte, c-à-d, pour toute suite bornée $(x_n)_n$ on peut extraire une sous suite de $(x_n)_n$ faiblement convergente.

Proposition 1.25. *Dans un espace de dimension finie, les topologies fortes et faibles coïncident.*

Définition 1.26.

Soient $f : X \rightarrow \overline{\mathbb{R}}$ et $x_0 \in X$, on dit que f est

(a) propre si et seulement si

$$f(x) \neq -\infty, \forall x \in X \quad \text{et} \quad \exists y_0 \in X : f(y_0) \neq +\infty.$$

(b) semi-continue inférieurement (s.c.i en abrégé) si pour tout $r \in]-\infty, f(x_0)]$, il existe un voisinage V de x_0 tel que $f(V) \subset]r, +\infty]$.

La semi-continuité inférieure peut se caractériser par l'une des deux propriétés suivantes : $\text{epi}(f)$ est fermé dans $X \times \mathbb{R}$ ou les ensembles de niveau $[f \leq r]$ sont fermés dans X pour tout $r \in \mathbb{R}$.

(c) Soient C un sous ensemble de H et $f : H \rightarrow \mathbb{R}$, une suite $(x_n)_n \subset C$ est dite une suite minimisante de f si

$$\lim_{n \rightarrow \infty} f(x_n) = \inf_{x \in C} f(x)$$

Lemme 1.27. *Soient X un espace de Banach réflexif et $A \subset X$ un ensemble convexe et fermé. Soit $\varphi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe s.c.i telle que*

$$\lim_{x \in A, \|x\| \rightarrow +\infty} \varphi(x) = +\infty$$

alors, $\varphi(\cdot)$ atteint son minimum sur A , c-à-d,

$$\exists x_0 \in A : \varphi(x_0) = \min_{x \in A} \varphi(x).$$

Autour de la projection et ses caractérisations

Ce second chapitre est consacré à la notion fondamentale de **projection**. On va s'intéresser à étudier la fonction de projection $P_C(\cdot)$ tout en donnant les propriétés fondamentales de cette application avec différents types d'ensembles (convexes, s.e.v, cônes convexes), notamment ses caractérisations qui donnent des définitions alternatives plus simples et élégantes que l'on utilisera dans le dernier chapitre.

2.1 Projection : définitions et propriétés générales

Dans cette section on va donner la définition de la projection sur un ensemble quelconque $C \in X$ tout en citant quelques exemples explicatifs. On va voir aussi, que la projection est une notion qui dépend de la norme utilisée ce qui implique que, pour un seul point $x \in H$, il va exister plusieurs projections. Le résultat fondamental qui donne les conditions et de l'existence et l'unicité de projection sera établi également dans cette partie.

Définition 2.1. *Soit C un sous ensemble non vide d'un espace normé X , et soit $x \in X$. Un élément $y \in C$ est une projection de x sur C , ou le point de C le plus proche à x si*

$$\|x - y\| = \inf_{z \in C} \|x - z\| = d(x, C). \quad (2.1)$$

On note $P_C(x)$ l'ensemble de tous les points $y \in C$ vérifiant (2.1), c-à-d

$$P_C(x) := \{y \in C, \|x - y\| = d(x, C)\}.$$

Un élément $y \in P_C(x)$ se caractérise également par l'inclusion suivante

$$y \in \mathbb{B}[x, d(x, C)] \cap C = \mathbb{S}(x, d(x, C)) \cap C. \quad (2.2)$$

Si $x \in C$ alors $P_C(x) = \{x\}$, autrement dit, C est l'ensemble des points fixes de $P_C(\cdot)$.

Exemple 2.2.

1. Si $C = \overline{\mathbb{B}}_X$, alors

$$\forall x \in X : P_C(x) = \begin{cases} x & \text{si } x \in C \\ \frac{x}{\|x\|} & \text{sinon} \end{cases} = \frac{x}{\max\{\|x\|, 1\}}.$$

En effet, si $x \notin C$, alors il suffit de prouver que $d(x, C) = \|x - \frac{x}{\|x\|}\|$. On a

$$\|x - \frac{x}{\|x\|}\| = \|\frac{\|x\| - 1}{\|x\|}x\| = \|x\| - 1,$$

comme $\frac{x}{\|x\|} \in \overline{\mathbb{B}}_X$ alors

$$\|x - \frac{x}{\|x\|}\| \geq d(x, C) = \inf_{z \in C} \|x - z\|$$

d'autre part, pour tout $z \in C$

$$\|x\| - 1 \leq \|x\| - \|z\| \leq \|x - z\| \Rightarrow \|x - \frac{x}{\|x\|}\| = \|x\| - 1 \leq \inf_{z \in C} \|x - z\| = d(x, C)$$

donc $d(x, C) = \|x - \frac{x}{\|x\|}\|$.

2. Si $C = [-1, 1] \cup [2, 3]$, alors (voir la Figure 2.1)

$$P_C(x) = \begin{cases} -1 & \text{si } x \leq -1 \\ x & \text{si } x \in C \\ 1 & \text{si } x \in [1, \frac{3}{2}[\\ 2 & \text{si } x \in]\frac{3}{2}, 2] \\ \{1, 2\} & \text{si } x = \frac{3}{2} \\ 3 & \text{si } x \geq 3 \end{cases}$$

Remarque 2.3. La définition de la projection dépend de la norme choisie, ça signifie que pour un seul point $x \in X$, il va exister plusieurs points de projection de x sur l'ensemble $C \subset X$. Cette propriété peut être vue clairement dans l'exemple suivant.

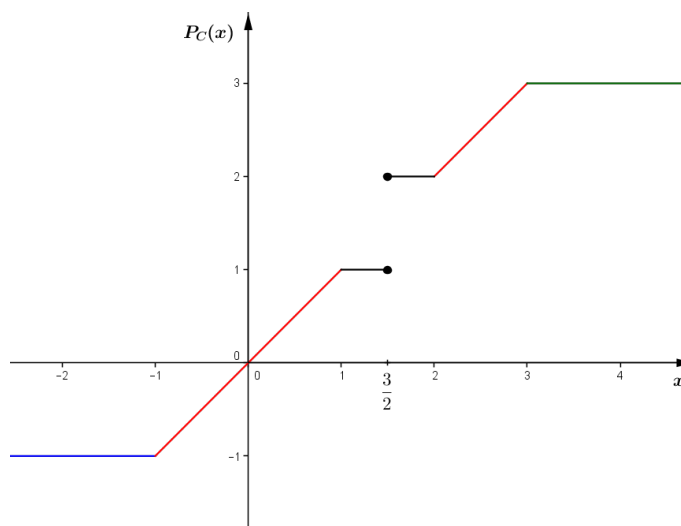
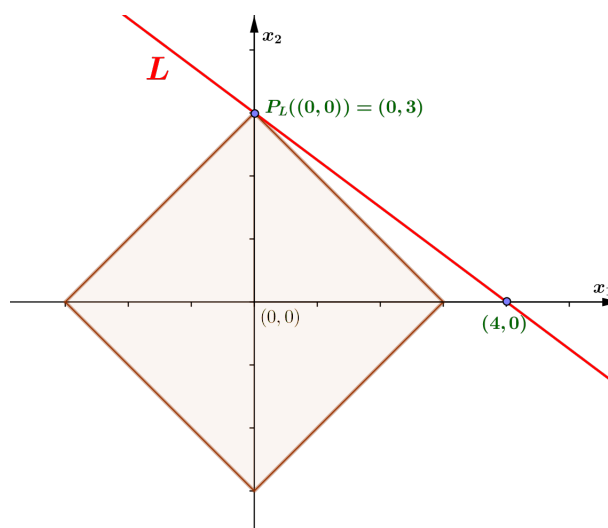
Exemple 2.4. Dans le plan \mathbb{R}^2 , on considère la droite L dont l'équation est $3x_1 + 4x_2 = 12$. On va calculer le point de L le plus proche à l'origine $(0, 0)$ en utilisant les trois normes définies ci-dessus, de façon plus précise

chercher $y \in L : y = P_L(0, 0)$ avec $\|y\|_\lambda = d((0, 0), L) = \inf_{z \in L} \|z\|_\lambda, \lambda \in \{1, 2, \infty\}$

1. Norme $\|\cdot\|_1$:

D'après l'égalité (2.2), la projection de $(0, 0)$ sur L est le premier point d'intersection de L avec la boule de centre $(0, 0)$ associée à la norme $\|\cdot\|_1$ en partant de l'origine vers l'extérieur. Selon la Figure 2.2, la droite L et cette boule se touchent, pour la première fois au point $(0, 3)$ ce qui signifie que

$$P_L((0, 0)) = (0, 3).$$

FIGURE 2.1 – Graphe de $P_C(\cdot)$ avec $C = [-1, 1] \cup [2, 3]$ FIGURE 2.2 – Projection par rapport à la norme $\|\cdot\|_1$

2. Norme $\|\cdot\|_2$:

L'équation de L s'écrit également sous la forme

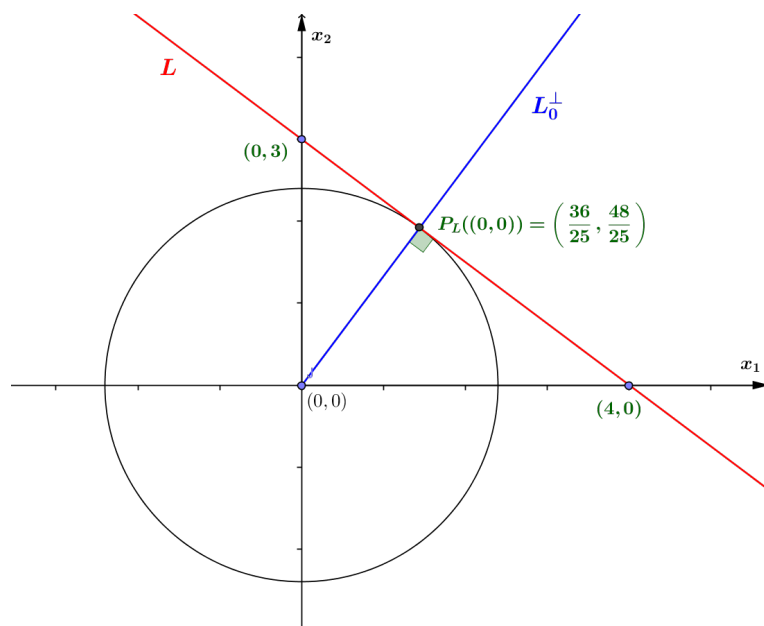
$$x = (x_1, x_2) = (0, 3) + \mu((4, 0) - (0, 3)) = (0, 3) + \mu(4, -3), \mu \in \mathbb{R}$$

Soit L_0^\perp la droite passant par l'origine et orthogonale à L , alors l'équation de L_0^\perp a la forme

$$L_0^\perp : x = (x_1, x_2) = 0 + \lambda(3, 4) = \lambda(3, 4).$$

il est clair que le point $P_L(0, 0)$ se donne comme le point d'intersection de L avec L_0^\perp (voir (2.12)), il suffit alors de résoudre l'équation suivante dont les inconnus sont μ et λ

$$(0, 3) + \mu(4, -3) = \lambda(3, 4)$$

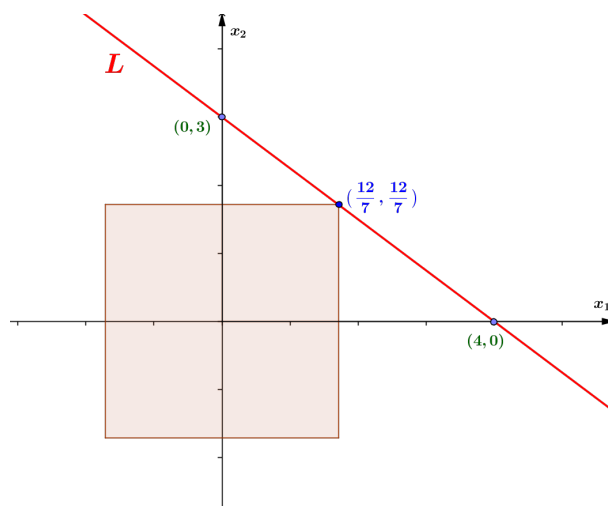
FIGURE 2.3 – Projection par rapport à la norme $\|\cdot\|_2$

ce qui donne $P_L((0,0)) = \left(\frac{36}{25}, \frac{48}{25}\right)$.

Dans ce cas, on a utilisé la notion d'orthogonalité par ce que l'on a un produit scalaire à partir duquel on a défini la norme $\|\cdot\|_2$, ce qui nous permet de définir l'orthogonal d'un ensemble.

$$\|x\|_2 = \langle x, x \rangle^{1/2} = (x^t x)^{1/2}, \forall x \in \mathbb{R}^n.$$

3. Norme $\|\cdot\|_\infty$:

FIGURE 2.4 – Projection par rapport à la norme $\|\cdot\|_\infty$

Le point de L le plus proche à l'origine que l'on note (\bar{x}_1, \bar{x}_2) , est le point d'intersection de la plus petite boule associée à $\|\cdot\|_\infty$ centrée en $(0,0)$ avec L . D'après

la Figure 2.4, (\bar{x}_1, \bar{x}_2) est le point d'intersection de la première bissectrice $M = \{x \in \mathbb{R}^2, x = (\lambda, \lambda), \lambda \in \mathbb{R}_+\}$ avec L , pour cela, il suffit alors de résoudre l'équation $(0, 3) + \mu(4, -3) = (\lambda, \lambda)$ ce qui donne $P_L((0, 0)) = (\frac{12}{7}, \frac{12}{7})$.

Proposition 2.5. Soit C un sous ensemble fermé de X et soit $x \in X$ alors

$$P_C(y) = P_C(x) \text{ pour tout } y \in [x, P_C(x)].$$

Preuve. Supposons qu'il existe $y \in [x, P_C(x)]$ tel que $P_C(y) \neq P_C(x)$ alors

$$\exists z \in C : \|y - z\| < \|y - P_C(x)\|$$

on a

$$y \in [x, P_C(x)] \Rightarrow \exists \lambda \in]0, 1[: y = \lambda x + (1 - \lambda)P_C(x)$$

donc

$$\begin{aligned} \|x - z\| &\leq \|x - y\| + \|y - z\| < \|x - y\| + \|y - P_C(x)\| \\ &= \|x - \lambda x - (1 - \lambda)P_C(x)\| + \|\lambda x + (1 - \lambda)P_C(x) - P_C(x)\| \\ &= \|x - P_C(x)\| \end{aligned}$$

ce qui contredit la définition de $P_C(x)$. □

Lemme 2.6. Soit C un sous ensemble non vide de X , alors pour tous $x, y \in X$ et $\lambda \in \mathbb{R}$

$$P_{C+y}(x + y) = P_C(x) + y, \quad P_{\lambda C}(\lambda x) = \lambda P_C(x) \quad (2.3)$$

Preuve. Tout d'abord

$$\begin{aligned} d(x + y, C + y) &= \inf_{z \in C+y} \|x + y - z\| = \inf_{z=z'+y, z' \in C} \|x + y - z\| = \inf_{z' \in C} \|x + y - (z' + y)\| \\ &= \inf_{z' \in C} \|x - z'\| = d(x, C) \end{aligned}$$

alors

$$\begin{aligned} z \in P_{C+y}(x + y) &\Leftrightarrow z \in C + y \text{ et } \|x + y - z\| = d(x + y, C + y) \\ &\Leftrightarrow z - y \in C \text{ et } \|x - (z - y)\| = d(x, C) \\ &\Leftrightarrow z - y \in P_C(x) \Leftrightarrow z \in P_C(x) + y. \end{aligned}$$

d'autre part,

$$d(\lambda x, \lambda C) = \inf_{y \in \lambda C} \|\lambda x - \lambda y\| = |\lambda| \inf_{y \in C} \|x - y\| = |\lambda| d(x, C)$$

alors, pour tout $\lambda \in \mathbb{R} \setminus \{0\}$

$$\begin{aligned} y \in P_{\lambda C}(\lambda x) &\Leftrightarrow y \in \lambda C \quad \text{et} \quad \|\lambda x - y\| = d(\lambda x, \lambda C) \\ &\Leftrightarrow \frac{1}{\lambda} y \in C \text{ et } |\lambda| \left\| x - \frac{1}{\lambda} y \right\| = |\lambda| d(x, C) \\ &\Leftrightarrow \frac{y}{\lambda} \in P_C(x). \end{aligned}$$

□

Le théorème suivant donne les conditions de l'existence et l'unicité de la projection dans un espace de Hilbert. On rappelle que ce théorème n'est plus valable dans un espace normé général.

Théorème 2.7 (Théorème de la projection sur un convexe fermé). *Soit C un ensemble convexe fermé de H , alors pour tout $x \in H$, il existe une et une seule projection $y = P_C(x)$, c-à-d,*

$$\forall x \in H, \exists! y \in C : d(x, C) = \|x - y\|.$$

Preuve.

I. Existence

1^{ère} **méthode** : Soient $x \in H$ et $\varphi(\cdot)$ la fonction définie par

$$\begin{aligned} \varphi : C &\rightarrow \mathbb{R} \cup \{+\infty\} \\ z &\mapsto \|x - z\| \end{aligned}$$

la fonction φ est convexe et continue (donc s.c.i), de plus $\lim_{z \in C, \|z\| \rightarrow +\infty} \varphi(z) = +\infty$. D'autre part, H est un Hilbert donc il est réflexif ce qui nous permet d'appliquer le Lemme 1.27 sur la fonction φ , donc la fonction φ atteint son minimum sur C , c-à-d, il existe $y \in C$ tel que $\|x - y\| = \min_{z \in C} \|x - z\|$

2^{ème} **méthode** :

Soit (d_n) une suite strictement décroissante de limite d (suite minimisante) avec

$$d = \inf_{z \in C} \|x - z\|$$

c-à-d : pour chaque $n \in \mathbb{N}^*$, il existe $y_n \in C$,

$$d_n = \|x - y_n\| \downarrow d = \inf_{z \in C} \|x - z\| \quad (2.4)$$

on va montrer que $(y_n)_n$ est une suite de Cauchy, pour cela, on utilise l'identité de parallélogramme

$$\begin{aligned} \|y_n - y_m\|^2 &= \|(x - y_n) - (x - y_m)\|^2 \\ &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|x - y_n + x - y_m\|^2 \\ &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\|x - \frac{y_n + y_m}{2}\|^2 \end{aligned}$$

or $\frac{1}{2}(y_n + y_m) \in C$ (car C est convexe), donc $\|x - \frac{y_n + y_m}{2}\| \geq d$. On obtient

$$\|y_n - y_m\|^2 \leq 2(d_n^2 + d_m^2) - 4d^2$$

Il est clair que le terme $2(d_n^2 + d_m^2) - 4d^2$ tend vers 0 quand n et m tendent vers l'infini grâce à (2.4) ce qui signifie que $(y_n) \subset C$ est une suite de Cauchy dans H qu'est complet,

elle converge alors vers une limite $y \in \overline{C} = C$.

on a

$$d \leq \|x - y_n\| = d_n$$

Par passage à la limite

$$d \leq \lim_n \|x - y_n\| = \|x - \lim_n y_n\| = \|x - y\| = \lim_n d_n = d$$

d'où

$$\|x - y\| = d = d(x, C).$$

II. Unicité

Supposons qu'il existe deux points de projection $y, y' \in H$ tels que $y \neq y'$ avec $d(x, C) = \|x - y\| = \|x - y'\|$. La convexité de C implique $\frac{y+y'}{2} \in C$ alors, d'après l'identité du parallélogramme on obtient

$$\begin{aligned} \left\| \frac{y+y'}{2} - x \right\|^2 &= \left\| \frac{1}{2}(y-x) + \frac{1}{2}(y'-x) \right\|^2 \\ &= 2 \times \frac{1}{4} \|y-x\|^2 + 2 \times \frac{1}{4} \|y'-x\|^2 - \frac{1}{4} \|y-y'\|^2 \\ &= \frac{1}{2} (\|y-x\|^2 + \|y'-x\|^2) - \frac{1}{4} \|y-y'\|^2 \\ &= d^2 - \frac{1}{4} \|y-y'\|^2 < d^2 \end{aligned}$$

ce qui contredit la définition de d donc $y = y'$. □

Définition 2.8. *Un ensemble fermé $C \subset H$ est dit de Chebychev si chaque point $x \in H$ admet un et un seul point à distance minimale dans C c-à-d, $P_C(x)$ est un singleton non vide (réduit à un seul élément) pour tout $x \in H$.*

Il est clair, d'après le Théorème 2.7 que chaque ensemble convexe (fermé) $C \subset H$ est un ensemble de Chebychev. La question que l'on peut poser ensuite est : La réciproque est-elle vraie ? Autrement dit, un ensemble de Chebychev est-il nécessairement convexe ?

Avant d'aller plus loin, il est bon d'observer que le problème posé ne peut pas être étudié que dans le cadre Hilbertien où toutes les hypothèses de la Définition 2.8 comptent ce qu'est justifié par les deux exemples suivants :

- Si $\|\cdot\|$ n'est pas une norme hilbertienne, la question n'a pas lieu d'être posé même si H est de dimension finie. Par exemple, si \mathbb{R}^2 est muni de la norme $\|(x_1, x_2)\|_\infty := \max(|x_1|, |x_2|)$ et si

$$C := \{(x_1, x_2) \in \mathbb{R}^2 : (x_2 \geq x_1 \geq 0) \text{ ou } (x_2 \geq 2x_1 \text{ et } x_1 \leq 0)\}.$$

la représentation graphique de C (voir Figure 2.5) montre qu'il est un ensemble de Chebychev, c-à-d tout point de \mathbb{R}^2 admet un et un seul point à distance minimale dans C par rapport à la norme $\|\cdot\|_\infty$, alors que C est loin d'être convexe.

- Si $\|\cdot\|$ est la norme euclidienne usuelle de \mathbb{R}^2 et si

$$C := \{(x_1, x_2) \in \mathbb{R}^2 : \|(x_1, x_2)\| \geq 1\}$$

alors tout point sauf l'origine admet un et un seul point à distance minimale dans C , tandis que C est voilement non convexe (ici la condition "pour tout point $x \in H$ " dans la Définition 2.8a n'est pas prise en considération).

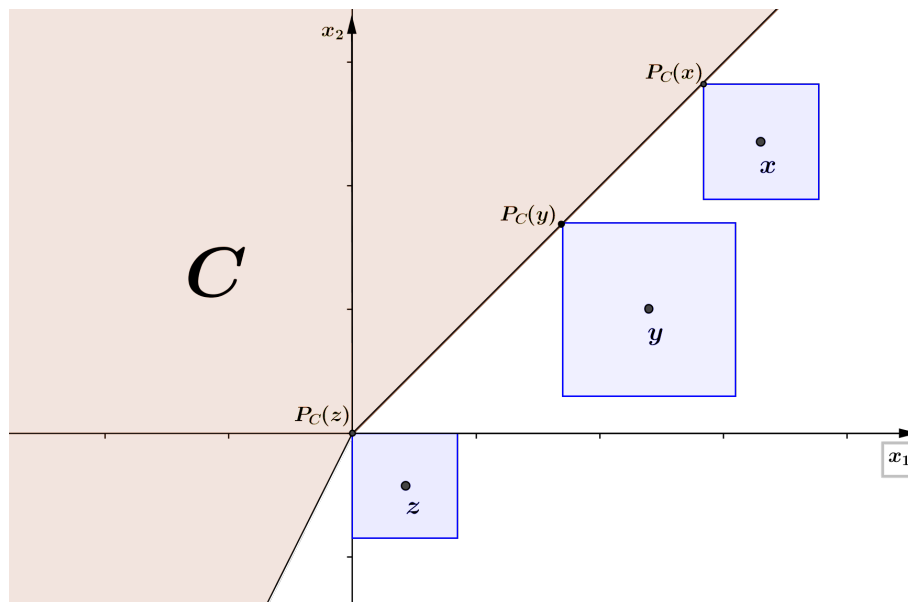


FIGURE 2.5 –

La question peut être posée, donc seulement dans un espace de Hilbert, dans cette situation, on distingue deux cas

- ◆ Si H est de dimension finie, la réponse est connue comme étant oui et ce depuis plus de soixante ans, c'est le fameux théorème de L.N.H. Bunt (dans sa thèse en 1934).
- ◆ Si H est de dimension infinie, la question a été clairement posée par V. Klee vers 1961 et cet auteur conjecturait même que la réponse était non, et bien qu'il y ait eu des progrès considérables depuis, la question reste ouverte à ce jour.

2.2 Caractérisations de la Projection sur différents types d'ensembles

On commence cette section par la caractérisation de la projection sur un ensemble quelconque de H , et à partir de cette propriété, on va prolonger l'étude aux différents types d'ensembles considérés ci-dessous. Ce résultat est très important dans la théorie d'approximation où on peut remplacer le problème de minimisation (chercher le point à

distance minimale) par un autre problème équivalent se donne via une inégalité variationnelle.

Théorème 2.9. *Soient C une partie quelconque non vide de H et $x \notin C$. Les assertions suivantes sont équivalentes*

(a) $y \in P_C(x)$ c-à-d,

$$y \in C \text{ et } \|x - y\| = d(x, C);$$

(b)

$$y \in C \quad \text{et} \quad \forall z \in C, \langle x - y, z - y \rangle \leq \frac{1}{2} \|z - y\|^2; \quad (2.5)$$

(c)

$$y \in C \quad \text{et} \quad y \in P_C[y + t(x - y)], \forall t \in]0, 1] \quad (2.6)$$

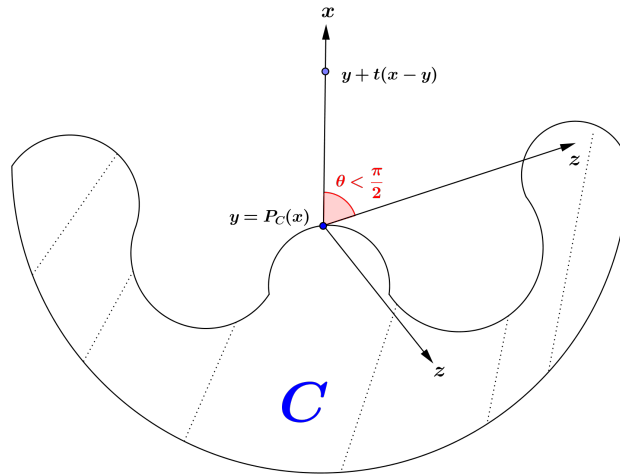


FIGURE 2.6 – Projection sur un ensemble non convexe : $\langle x - y, z - y \rangle$ peut être positif

Preuve.

$$\begin{aligned} (a) &\Leftrightarrow y \in C \text{ et } \|x - y\| \leq \|x - z\|, \forall z \in C \\ &\Leftrightarrow y \in C \text{ et } \|x - y\|^2 \leq \|(x - y) + (y - z)\|^2 \\ &\Leftrightarrow y \in C \text{ et } \|x - y\|^2 \leq \|x - y\|^2 + \|y - z\|^2 + 2\langle x - y, z - y \rangle \\ &\Leftrightarrow y \in C \text{ et } \langle x - y, z - y \rangle \leq \frac{1}{2} \|y - z\|^2, \forall z \in C. \end{aligned}$$

d'autre part

$$\begin{aligned} (b) &\Leftrightarrow y \in C \text{ et } 2\langle x - y, z - y \rangle \leq \frac{1}{t} \|z - y\|^2, \forall t \in]0, 1] \\ &\Leftrightarrow y \in C \text{ et } 2\langle [y + t(x - y)] - y, z - y \rangle \leq \|z - y\|^2 \forall t \in]0, 1] \end{aligned} \quad (2.7)$$

l'équivalence entre (a) et (b) donne

$$(2.7) \Leftrightarrow \langle z' - y, z - y \rangle \leq \frac{1}{2} \|z - y\|^2, \forall z \in C \text{ avec } z' = y + t(x - y) \\ \Leftrightarrow y \in P_C(z') \text{ i.e } y \in P_C[y + t(x - y)].$$

□

2.2.1 Projection sur un convexe fermé

L'inégalité (2.5) caractérise la projection sur un ensemble quelconque C d'un espace de Hilbert. La question qui vient à l'esprit naturellement est : comment cette caractérisation devient-elle lorsque l'ensemble C est convexe ? Le théorème suivant répond à cette question

Théorème 2.10. *Soit C un ensemble convexe fermé non vide de H , la projection sur C se caractérise par l'équivalence suivante*

$$y = P_C(x) \Leftrightarrow y \in C \text{ et } \forall z \in C, \langle z - y, x - y \rangle \leq 0 \quad (2.8)$$

Preuve. \Rightarrow) En tenant compte du Théorème 2.9 on obtient

$$y = P_C(x) \Leftrightarrow y \in C \text{ et } \langle x - y, z - y \rangle \leq \frac{1}{2} \|z - y\|^2, \forall z \in C$$

Soient $t \in]0, 1[$ et $z_t = y + t(z - y) \in C$ alors

$$\langle x - y, y + t(z - y) - y \rangle \leq \frac{1}{2} \|y + t(z - y) - y\|^2, \forall z \in C \quad \forall t \in]0, 1[\\ \Rightarrow t \langle x - y, z - y \rangle \leq \frac{t^2}{2} \|z - y\|^2, \forall z \in C \quad \forall t \in]0, 1[$$

par passage à la limite $t \downarrow 0$ on obtient l'inégalité donnée dans (2.8)

\Leftarrow) Soit $z \in C$, alors

$$y \in C \text{ et } \langle x - y, z - y \rangle \leq 0 \leq \frac{1}{2} \|z - y\|^2 \Leftrightarrow y = P_C(x)$$

□

Interprétation géométrique

La caractérisation de la projection sur un convexe fermé (voir Figure 2.7) s'exprime par le fait que l'angle θ formé par les vecteurs $x - y$ et $z - y, z \in C$ est toujours obtus, c-à-d, $\theta \geq \frac{\pi}{2}$, ce qui entraîne que $\cos \theta = \langle x - y, z - y \rangle \leq 0, \forall z \in C$, par contre si C n'est pas forcément convexe (voir Figure 2.6) cet angle peut être aigu et par suite $\cos \theta = \langle x - y, z - y \rangle$ peut être positif (Théorème 2.9 (b)).

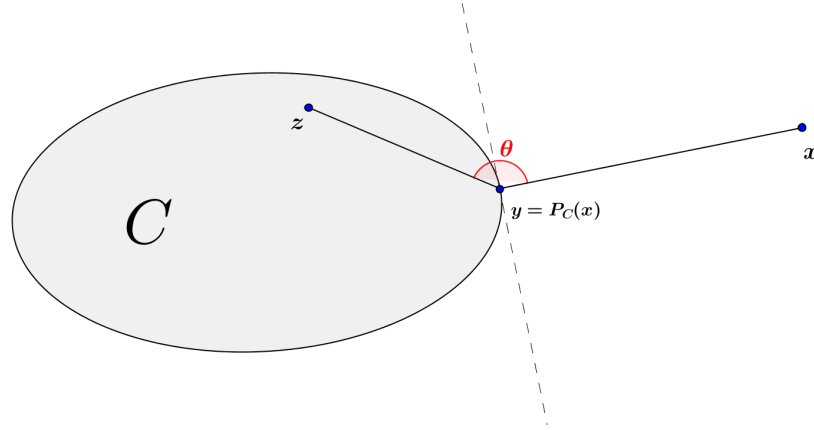


FIGURE 2.7 – interprétation géométrique de la caractérisation (2.8)

Théorème 2.11. Soit C un ensemble convexe fermé non vide de H . les propriétés suivantes ont lieu

1. pour tous $x, y \in H$, on a

$$P_{y+C}(x) = y + P_C(x - y) \quad (2.9)$$

2. pour tous $x \in H$ et pour tout $\lambda \in \mathbb{R}_+$

$$P_C(P_C(x) + \lambda(x - P_C(x))) = P_C(x).$$

Cette égalité montre que $P_C(y)$ est l'unique projection de chaque point y appartenant, non seulement au segment $[P_C(x), x]$ (voir (2.6)), mais à toute la demi-droite commençant en $P_C(x)$ et passant par x .

3. pour tout $x \in H$

$$y = P_C(x) \Leftrightarrow x - y \in N_C(y).$$

Preuve.

1. La caractérisation de la projection sur un convexe fermé dit que pour tout $z \in C$

$$\begin{aligned} \langle (y + z) - (y + P_C(x - y)), x - (y + P_C(x - y)) \rangle &= \langle z - P_C(x - y), (x - y) - P_C(x - y) \rangle \\ &\leq 0, \end{aligned}$$

ce qui traduit l'égalité $P_{y+C}(x) = y + P_C(x - y)$.

2. Soient $\lambda \in \mathbb{R}_+$ et $y \in C$, d'après le Théorème 2.10

$$\begin{aligned} \langle y - P_C(x), P_C(x) + \lambda(x - P_C(x)) - P_C(x) \rangle &= \lambda \langle y - P_C(x), x - P_C(x) \rangle \leq 0 \\ \Rightarrow P_C(P_C(x) + \lambda(x - P_C(x))) &= P_C(x) \end{aligned}$$

3. D'après la Définition 1.13

$$\begin{aligned} x - y \in N_C(y) &\Leftrightarrow \langle x - y, z - y \rangle \leq 0, \forall z \in C \\ &\stackrel{(2.8)}{\Leftrightarrow} y = P_C(x). \end{aligned}$$

□

Proposition 2.12. *Soit $C \subset H$ un ensemble convexe fermé, alors*

(a)

$$\langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle \geq \|P_C(x_1) - P_C(x_2)\|^2, \forall x_1, x_2 \in H$$

ce qui implique la monotonie de $P_C(\cdot)$

(b) Pour tous $x, y \in H$

$$\|x_1 - x_2\|^2 \geq \|P_C(x_1) - P_C(x_2)\|^2 + \|x_1 - P_C(x_1) - [x_2 - P_C(x_2)]\|^2$$

(c) L'application $P_C(\cdot)$ est lipschitzienne de rapport 1 c-à-d

$$\|P_C(x_1) - P_C(x_2)\| \leq \|x_1 - x_2\|, \forall x_1, x_2 \in H$$

par conséquent, $P_C(\cdot)$ est uniformément continue.

Preuve. (a) la caractérisation de la projection sur un convexe fermé nous donne

$$\langle x_1 - P_C(x_1), z_1 - P_C(x_1) \rangle \leq 0, \forall z_1 \in C \quad (2.10)$$

$$\langle x_2 - P_C(x_2), z_2 - P_C(x_2) \rangle \leq 0, \forall z_2 \in C \quad (2.11)$$

On prend $z_1 = P_C(x_2) \in C, z_2 = P_C(x_1) \in C$ respectivement dans (2.10) et (2.11), on obtient

$$\langle x_1 - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle \leq 0 \text{ et } \langle x_2 - P_C(x_2), P_C(x_1) - P_C(x_2) \rangle \leq 0$$

l'addition membre à membre donne

$$\begin{aligned} \langle x_1 - x_2 + P_C(x_2) - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle &\leq 0 \\ \Rightarrow \langle P_C(x_2) - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle &\leq \langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle \\ \Rightarrow \|P_C(x_2) - P_C(x_1)\|^2 &\leq \langle P_C(x_2) - P_C(x_2), x_2 - x_1 \rangle \end{aligned}$$

(b) En utilisant (a) on trouve

$$\begin{aligned} \|x_1 - x_2\|^2 &= \|[x_1 - P_C(x_1)] + [P_C(x_1) - P_C(x_2)] + [P_C(x_2) - x_2]\|^2 \\ &= \|P_C(x_1) - P_C(x_2)\|^2 + \|x_1 - P_C(x_1) - [x_2 - P_C(x_2)]\|^2 \\ &\quad + 2\langle P_C(x_1) - P_C(x_2), x_1 - x_2 - [P_C(x_1) - P_C(x_2)] \rangle \\ &= \|P_C(x_1) - P_C(x_2)\|^2 + \|x_1 - P_C(x_1) - [x_2 - P_C(x_2)]\|^2 \\ &\quad + 2\langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle - 2\|P_C(x_1) - P_C(x_2)\|^2 \\ &\stackrel{(a)}{\geq} \|P_C(x_1) - P_C(x_2)\|^2 + \|x_1 - P_C(x_1) - [x_2 - P_C(x_2)]\|^2. \end{aligned}$$

(c) Il suffit d'utiliser (b) ou d'appliquer l'inégalité de Cauchy-Schwartz à l'inégalité (a)

$$\begin{aligned} \|P_C(x_2) - P_C(x_1)\|^2 &\leq \langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle \leq \|P_C(x_1) - P_C(x_2)\| \|x_1 - x_2\| \\ &\Rightarrow \|P_C(x_2) - P_C(x_1)\| \leq \|x_1 - x_2\| \end{aligned}$$

□

2.2.2 Projection sur un sous espace vectoriel fermé

En se basant sur la caractérisation de la projection sur un convexe fermé établie dans le Théorème 2.7, on va étendre ce résultat au cas des sous espaces vectoriels.

Théorème 2.13. *Soient F un sous espace vectoriel fermé non vide de H et $x \in H$, le point de projection de x sur F se caractérise via l'équivalence suivante*

$$y = P_F(x) \Leftrightarrow y \in F \text{ et } \langle x - y, z \rangle = 0, \forall z \in F. \quad (2.12)$$

Autrement dit, y est la projection de x sur F si et seulement si le vecteur $x - y$ est orthogonal à F : $(x - y) \perp F$, c'est pour cette raison, $P_F(x)$ s'appelle parfois la projection **orthogonale** de x sur F .

Preuve.

\Rightarrow) Soient $x \in H$ et $y = P_F(x)$, d'après la caractérisation (2.8) (tout s.e.v est un convexe) on a

$$y \in F, \text{ et } \langle x - y, z - y \rangle \leq 0, \forall z \in F$$

on pose $z' = z - y$, alors $z' \in F$ et par suite

$$\langle x - y, z' \rangle \leq 0, \forall z' \in F$$

comme F est un s.v.e, alors la dernière inégalité reste vraie pour $-z \in F$ et ça donne

$$\langle x - y, z' \rangle \geq 0, \forall z' \in F$$

ce qui est vrai seulement si $\langle x - y, z' \rangle = 0, \forall z' \in F$.

\Leftarrow) D'après la caractérisation (2.8) il suffit de prouver que $\langle x - y, z - y \rangle \leq 0, \forall z \in F$, on a

$$\begin{aligned} y \in F \text{ et } \langle x - y, z \rangle = 0, \forall z \in F &\Rightarrow y \in F \text{ et } \langle x - y, y \rangle = 0 \\ &\Rightarrow y \in F \text{ et } \langle x - y, z - y \rangle = 0, \forall z \in F \\ &\stackrel{(2.8)}{\implies} y = P_F(x). \end{aligned}$$

□

La proposition suivante collecte les propriétés fondamentales de la projection sur un s.e.v

Proposition 2.14. *Soit F un sous espace vectoriel fermé de H .*

(a) *Chaque élément $x \in H$ se décompose sous la forme*

$$x = P_F(x) + P_{F^\perp}(x) \quad \text{c-à-d, } Id = P_F + P_{F^\perp},$$

de plus cette décomposition est unique dans le sens : si $x = y + z$ avec $y \in F, z \in F^\perp$ alors $y = P_F(x)$ et $z = P_{F^\perp}(x)$, on écrit également : $H = F \oplus F^\perp$. Cela signifie que déterminer $P_F(x)$ et déterminer $P_{F^\perp}(x)$ sont deux problèmes équivalents : quand on a l'une on a l'autre.

(b) $P_F(\cdot)$ *est linéaire continue avec $\|P_F\| = 1$.*

(c) $\|x\|^2 = \|P_F(x)\|^2 + \|P_{F^\perp}(x)\|^2$, *pour tout $x \in H$.*

(d)

$$\|P_F(x)\| \leq \|x\| \quad \text{et} \quad \langle P_F(x), x \rangle = \|P_F(x)\|^2, \forall x \in H$$

(e) $P_F(\cdot)$ *est auto-adjoint, c-à-d, $\langle P_F(x), y \rangle = \langle x, P_F(y) \rangle$, pour tout $x \in H$.*

(f) $P_F^2 = P_F$ *et $\ker P_F = P_F^{-1}(\{0\}) = F^\perp$.*

Preuve.

Avant de commencer la preuve, il est important de rappeler que l'ensemble F^\perp est un sous espace vectoriel fermé grâce à la fermeture de F , de plus $F^{\perp\perp} = F$ ce qui nous permet d'appliquer, sur F^\perp les mêmes propriétés appliquées sur F .

(a) Soit $x \in H$, alors $P_F(x)$ existe grâce au Théorème 2.7, on pose $y = x - P_F(x)$, alors, on doit montrer que $y = P_{F^\perp}(x)$. Il est clair que $y \in F^\perp$ (selon (2.12)), de plus

$$\begin{aligned} P_F(x) \in F &\Rightarrow \langle x - y, z \rangle = \langle x - (x - P_F(x)), z \rangle = \langle P_F(x), z \rangle = 0, \forall z \in F^\perp \\ &\Rightarrow y = x - P_F(x) \perp F^\perp \stackrel{(2.12)}{\implies} x - P_F(x) = P_{F^\perp}(x) \Rightarrow x = P_F(x) + P_{F^\perp}(x). \end{aligned}$$

il ne reste que démontrer l'unicité de la décomposition, pour cela, on considère une autre décomposition $x = y + z$ avec $y \in F, z \in F^\perp$. Pour tout $c \in F$ on a

$$\langle x - y, c \rangle = \langle z, c \rangle = 0$$

ce qui traduit l'égalité $y = P_F(x)$. De façon similaire on trouve que $z = P_{F^\perp}(x)$.

(b) (•) linéarité de $P_F(\cdot)$: soient $x, y \in H$ alors on a

$$\begin{aligned} x - P_F(x) \perp F &\Rightarrow \langle x - P_F(x), y \rangle = 0, \forall y \in F \\ z - P_F(z) \in F^\perp &\Rightarrow \langle z - P_F(z), y \rangle = 0, \forall y \in F \\ \Leftrightarrow \langle (x + z) - [P_F(x) + P_F(z)], y \rangle &= 0, \forall y \in F \\ \stackrel{(2.12)}{\implies} P_F(x + z) &= P_F(x) + P_F(z) \end{aligned}$$

ainsi

$$\begin{aligned} x - P_F(x) \perp F &\Rightarrow \langle x - P_F(x), y \rangle = 0, \forall y \in F \\ &\Rightarrow \langle \lambda x - \lambda P_F(x), y \rangle = 0, \forall y \in F, \lambda \in \mathbb{R} \\ &\stackrel{(2.12)}{\implies} P_F(\lambda x) = \lambda P_F(x), \forall \lambda \in \mathbb{R} \end{aligned}$$

- (•) continuité de P_F : est une conséquence directe de la Lipschitzité de $P_F(\cdot)$ établie dans la Proposition 2.12 (c).
- (•) $\|P_F\| = 1$: on a, d'après (d), $\|P_F(x)\| \leq \|x\|$ pour tout $x \in H$ donc $\|P_F\| \leq 1$, d'autre part

$$\begin{aligned} \text{pour tout } x \in F : P_F(x) = x &\Rightarrow \|x\| = \|P_F(x)\| \leq \|P_F\| \|x\| \\ &\Rightarrow \|P_F\| \geq 1. \end{aligned}$$

(c) Tout d'abord

$$\begin{aligned} 0 &\stackrel{(2.12)}{=} \langle x - P_F(x), P_F(x) \rangle \stackrel{(a)}{=} \langle x - P_F(x), x - P_{F^\perp}(x) \rangle \\ &= \|x\|^2 - \langle x, P_{F^\perp}(x) \rangle - \langle P_F(x), x \rangle + \langle P_F(x), P_{F^\perp}(x) \rangle \\ &= \|x\|^2 - \langle x, P_{F^\perp}(x) \rangle - \langle x, P_F(x) \rangle \end{aligned}$$

alors

$$\begin{aligned} \|x\|^2 &= \langle x, P_{F^\perp}(x) \rangle + \langle x, P_F(x) \rangle \\ &= \langle x + P_{F^\perp}(x) - P_{F^\perp}(x), P_{F^\perp}(x) \rangle + \langle x + P_F(x) - P_F(x), P_F(x) \rangle \\ &= \langle x - P_{F^\perp}(x), P_{F^\perp}(x) \rangle + \langle P_{F^\perp}(x), P_{F^\perp}(x) \rangle + \langle x - P_F(x), P_F(x) \rangle + \langle P_F(x), P_F(x) \rangle \\ &= \|P_F(x)\|^2 + \|P_{F^\perp}\|^2. \end{aligned}$$

(d)

$$\begin{aligned} \|x\|^2 &= \|P_F(x)\|^2 + \|P_{F^\perp}(x)\|^2 \Rightarrow \|x\| \geq \|P_F(x)\| \\ \langle P_F(x), x \rangle &\stackrel{(a)}{=} \langle P_F(x), P_F(x) + P_{F^\perp}(x) \rangle = \|P_F(x)\|^2 + \underbrace{\langle P_F(x), P_{F^\perp}(x) \rangle}_{=0} = \|P_F(x)\|^2 \end{aligned}$$

(e) Soient $x, y \in H$ alors

$$\begin{aligned} 0 &\stackrel{(2.12)}{=} \langle x - P_F(x), P_F(y) \rangle = \langle x, P_F(y) \rangle - \langle P_F(x), P_F(y) \rangle \\ &\Rightarrow \langle x, P_F(y) \rangle = \langle P_F(x), P_F(y) \rangle \end{aligned} \tag{2.13}$$

de même

$$\begin{aligned} 0 &\stackrel{(2.12)}{=} \langle y - P_F(y), P_F(x) \rangle = \langle y, P_F(x) \rangle - \langle P_F(x), P_F(y) \rangle \\ &\Rightarrow \langle y, P_F(x) \rangle = \langle P_F(x), P_F(y) \rangle \end{aligned} \tag{2.14}$$

en combinant (2.13) et (2.14) on trouve que $\langle y, P_F(x) \rangle = \langle x, P_F(y) \rangle$.

(f) $P_F^2 = P_F$ trivial.

$$x \in \ker P_F \Leftrightarrow P_F(x) = 0 \stackrel{(2.12)}{\iff} x - 0 = x \perp F \Leftrightarrow x \in F^\perp$$

□

2.2.3 Projection sur un cône convexe fermé

C'est en quelque sorte la situation intermédiaire entre celle rappelée dans la partie 2.2.1 et celle traitée au 2.2.2. Mais, du fait que C sera pris un cône convexe fermé, on va aller beaucoup plus loin que dans la sous-section 2.2.1 et se rapprocher de ce qu'on obtenait dans le contexte de la sous-section 2.2.2.

Théorème 2.15. *Soit C un cône convexe fermé de H , le point $P_C(x)$ se caractérise par les équivalences suivantes*

$$y = P_C(x) \Leftrightarrow y \in C, x - y \in C^\circ \text{ et } x - y \perp y \quad (2.15)$$

$$\Leftrightarrow \begin{cases} \langle x - y, y \rangle = 0 \\ \forall z \in C : \langle x - y, z \rangle \leq 0 \end{cases} \quad (2.16)$$

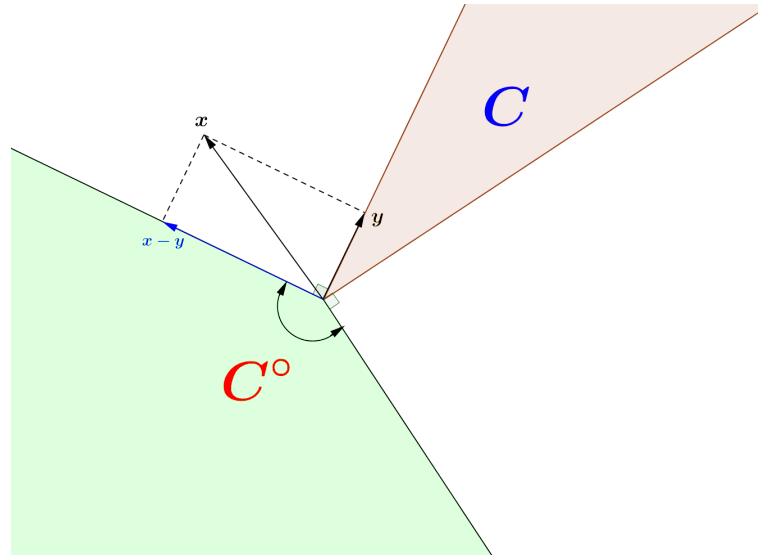


FIGURE 2.8 – Caractérisation de la projection sur un cône convexe

Preuve.

Supposons que $y = P_C(x)$ alors $y \in C$ et $\langle x - y, z - y \rangle \leq 0, \forall z \in C$ comme C est un cône alors d'après (2.8) $\lambda y \in C$ pour tout $\lambda \geq 0$ ce qui implique $\langle x - y, \lambda y - y \rangle \leq 0$ (en prenant $z = \lambda y \in C$) alors $(\lambda - 1)\langle x - y, y \rangle \leq 0$. En particulier, pour $\lambda = 0$ on obtient $\langle x - y, y \rangle \geq 0$ et pour $\lambda = 2$ on trouve $\langle x - y, y \rangle \leq 0$ pour tout $z \in C$ ce qui est vraie seulement si $\langle x - y, y \rangle = 0$, c'est à dire $x - y \in C^\circ$.

Réciproquement, supposons que y vérifie la partie droite de l'équivalence (2.15). On considère la fonction $f : H \rightarrow \mathbb{R}$ qui à chaque $z \in H$ associe $f(z) = \|x - z\|^2$ alors

$$\begin{aligned} f(z) &= \|x - y + y - z\|^2 = \|x - y\|^2 + \|y - z\|^2 + 2\langle x - y, y - z \rangle \\ &\geq \|y - z\|^2 + 2\langle x - y, y \rangle - 2\langle x - y, z \rangle \\ &\geq f(y) \end{aligned}$$

par conséquent $f(z) \geq f(y)$ pour tout $z \in C$, ce qui exprime bien que y est le point de C le plus proche à x c-à-d $y = P_C(x)$. \square

Théorème 2.16 (Décomposition de Moreau). *Pour tout $x \in H$, la décomposition suivante a lieu*

$$x = P_C(x) + P_{C^\circ}(x) \quad \text{et} \quad \langle P_C(x), P_{C^\circ}(x) \rangle = 0$$

de plus, cette décomposition est unique, c-à-d, s'il existe $u \in C, v \in C^\circ$ tels que $x = u + v$ et $\langle P_C(x), P_{C^\circ}(x) \rangle = 0$ alors $u = P_C(x)$ et $v = P_{C^\circ}(x)$.

Preuve.

Soient $x \in H$ et $c_0 = x - P_C(x)$ alors (2.15) donne

$$c_0 \in C^\circ \quad \text{et} \quad \langle c_0, P_C(x) \rangle = \langle c_0, x - c_0 \rangle = 0 \quad (2.17)$$

soit $y \in C^\circ$ alors

$$\langle x - c_0, y \rangle = \underbrace{\langle P_C(x), y \rangle}_{\in C} \leq 0 \Rightarrow x - c_0 \in (C^\circ)^\circ$$

posons $K := C^\circ$, il est clair que K un cône convexe qui vérifie

$$c_0 \in K, \quad x - c_0 \in K^\circ, \quad \langle x - c_0, c_0 \rangle = 0$$

ce qu'est équivalent selon (2.15) à $c_0 = P_K(x) = P_{C^\circ}(x)$ et par conséquent

$$x = P_C(x) + c_0 = P_C(x) + P_{C^\circ}(x) \quad \text{et} \quad \langle P_C(x), P_{C^\circ}(x) \rangle = \langle P_C(x), c_0 \rangle \stackrel{(2.17)}{=} 0$$

il nous reste que de démontrer l'unicité de la décomposition, pour cela, on considère une autre décomposition $x = y + z$ avec $y \in C, z \in C^\circ$ et $\langle y, z \rangle = 0$. Pour tout $c \in C$ on a

$$\langle x - y, c \rangle = \langle z, c \rangle \leq 0 \quad \text{et} \quad \langle x - y, y \rangle = \langle z, y \rangle = 0$$

ce qui traduit l'égalité $y = P_C(x)$. De façon similaire on trouve que $z = P_{C^\circ}(x)$. \square

Exemple 2.17.

1. On considère l'orthant positif $K = \mathbb{R}_+^n$ et son polaire $K^\circ = \mathbb{R}_-^n$, alors pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ on a

$$P_K(x) = (x_1^+, \dots, x_n^+) \quad \text{et} \quad P_{K^\circ}(x) = (x_1^-, \dots, x_n^-)$$

avec

$$x_i^+ = \begin{cases} x_i & \text{si } x_i \geq 0 \\ 0 & \text{si } x_i < 0 \end{cases}, \quad x_i^- = \begin{cases} x_i & \text{si } x_i \leq 0 \\ 0 & \text{si } x_i > 0 \end{cases} \Rightarrow x_i^+ x_i^- = 0$$

la décomposition de Moreau est donc

$$x = (x_1, \dots, x_n) = P_K(x) + P_{K^\circ}(x) = (x_1^+ + x_1^-, \dots, x_n^+ + x_n^-)$$

avec

$$\langle P_K(x), P_{K^\circ}(x) \rangle = \sum_{i=1}^n x_i^+ x_i^- = 0$$

2. Projection sur le cône $K = \mathcal{S}_n^+(\mathbb{R})$ de $\mathcal{S}_n(\mathbb{R})$

Soit $A \in \mathcal{S}_n(\mathbb{R})$ alors, d'après le théorème spectral (décomposition de Schur) on trouve

$$U^T A U = U^{-1} A U = \text{diag}(\lambda_1, \dots, \lambda_n)$$

avec $U \in \mathcal{M}_n(\mathbb{R})$ est une matrice orthogonale ($U U^t = U^t U = I_n$) et $\lambda_i, i = 1, \dots, n$ sont les valeurs propres de A . Si on pose

$$A_1 = U \text{diag}(\lambda_1^+, \dots, \lambda_n^+) U^t, \quad A_2 = U \text{diag}(\lambda_1^-, \dots, \lambda_n^-) U^t$$

alors, $A_1 \in \mathcal{S}_n^+(\mathbb{R})$ et $A_2 \in \mathcal{S}_n^-(\mathbb{R})$. La décomposition de Moreau associée au produit scalaire $\langle \cdot, \cdot \rangle$ est donc

$$\forall A \in \mathcal{S}_n(\mathbb{R}) : A = A_1 + A_2$$

avec

$$\begin{aligned} \langle A_1, A_2 \rangle &= \text{trace}(A_1 A_2) = \text{trace}\left(U \text{diag}(\lambda_1^+, \dots, \lambda_n^+) U^T U \text{diag}(\lambda_1^-, \dots, \lambda_n^-) U^T\right) \\ &= \text{trace}\left(U \text{diag}(\lambda_1^+ \lambda_1^-, \dots, \lambda_n^+ \lambda_n^-) U^T\right) = 0 \end{aligned}$$

Une autre manière de dire les choses :

A_1 est la matrice semie-définie positive la plus proche de A

A_2 est la matrice semie-définie négative la plus proche de A

(au sens de la norme matricielle associée à $\langle \cdot, \cdot \rangle$.)

Remarque 2.18. La décomposition de Moreau généralise la décomposition classique (fondamentale) établie lorsque C est un sous-espace vectoriel fermé F de H

$$x = P_F(x) + P_{F^\perp}(x)$$

Il y a néanmoins quelques différences essentielles :

- P_C n'est pas une application linéaire (voir Proposition 2.19).
- En projetant x sur C et sur C° , on n'était pas sûr d'obtenir des éléments orthogonaux alors que pour un sous-espace vectoriel F , tout élément de F est orthogonal à tout élément de F^\perp .
- La décomposition de $x \in H$ en $x = x_1 + x_2$, où $x_1 \in C$ et $x_2 \in C^\circ$, n'est pas unique

Les propriétés majeures de la projection sur un cône convexe peuvent être résumées dans la proposition suivante

Proposition 2.19. Soient C un cône convexe fermé et $x, z \in H$ alors

- (a) $P_C^2 = P_C$ (P_C est idempotent);
 (b) Pour tout réel positif $\lambda \in \mathbb{R}_+$: $P_C(\lambda x) = \lambda P_C(x)$;
 (c)

$$P_C(x + z) = P_C(x) + P_C(z) \Leftrightarrow \langle P_C(x), z \rangle = \langle P_C(x), P_C(z) \rangle = \langle x, P_C(z) \rangle$$

- (d) $\|x\|^2 = \|P_C(x)\|^2 + \|P_{C^\circ}(x)\|^2, \forall x \in H$
 (e) $C^\circ = \{x \in H, P_C(x) = 0\}$ et $C = \{x \in H, P_{C^\circ}(x) = 0\}$
 (f) $\|P_C(x)\| \leq \|x\|, \forall x \in H$

Preuve.

- (a) On sait que si $x \in C$ alors $x = P_C(x)$ donc $P_C(P_C(x)) = P_C^2(x) = P_C(x)$.
 (b) D'après la caractérisation de la projection sur un convexe fermé (2.8)

$$\begin{aligned} y = P_C(\lambda x) &\Leftrightarrow \langle z - y, \lambda x - y \rangle \leq 0, \forall z \in C \\ &\Leftrightarrow \langle \frac{z}{\lambda} - \frac{1}{\lambda}y, x - \frac{1}{\lambda}y \rangle \leq 0, \forall \frac{z}{\lambda} \in C \\ &\Leftrightarrow \langle z' - \frac{1}{\lambda}y, x - \frac{1}{\lambda}y \rangle \leq 0, \forall z' \in C \\ &\Leftrightarrow \frac{1}{\lambda}y = P_C(x) \\ &\Leftrightarrow y = \lambda P_C(x). \end{aligned}$$

- (c) Tout d'abord, il vient de (2.16) que pour tous $x, z \in H$

$$\begin{aligned} \langle P_C(z), x - P_C(x) \rangle + \langle P_C(x), z - P_C(z) \rangle &= \langle P_C(z), x - P_C(x) \rangle + \underbrace{\langle P_C(x), x - P_C(x) \rangle}_{=0} \\ &+ \langle P_C(x), z - P_C(z) \rangle + \underbrace{\langle P_C(z), z - P_C(z) \rangle}_{=0} \\ &= \langle P_C(x) + P_C(z), x + z - (P_C(x) + P_C(z)) \rangle \end{aligned} \quad (2.18)$$

Supposons que $P_C(x + y) = P_C(x) + P_C(y)$ alors

$$\begin{aligned} (2.18) &\Rightarrow \langle P_C(z), x - P_C(x) \rangle + \langle P_C(x), z - P_C(z) \rangle \\ &= \langle P_C(x + z), x + z - P_C(x + z) \rangle \stackrel{(2.16)}{=} 0 \end{aligned} \quad (2.19)$$

d'autre part, l'équivalence (2.16) donne

$$\langle P_C(z), x - P_C(x) \rangle \leq 0 \text{ et } \langle P_C(x), z - P_C(z) \rangle \leq 0$$

en combinaison ceci avec (2.19) on obtient

$$\begin{aligned} \langle P_C(z), x - P_C(x) \rangle &= \langle P_C(x), z - P_C(z) \rangle = 0 \\ \Leftrightarrow \langle P_C(x), z \rangle &= \langle P_C(x), P_C(z) \rangle = \langle x, P_C(z) \rangle \end{aligned} \quad (2.20)$$

Réciproquement, on suppose que les dernières égalités aient lieu, alors

$$x + z = \underbrace{P_C(x) + P_C(z)}_{\in C} + \underbrace{x - P_C(x)}_{\in C^\circ \text{ selon (2.15)}} + \underbrace{z - P_C(z)}_{\in C^\circ \text{ selon (2.15)}}$$

$$x + z = u + v : u \in C \text{ et } v \in C^\circ$$

ce qui signifie que $x + z$ se décompose sous la forme d'une somme de deux éléments de C et C° respectivement, de plus

$$\begin{aligned} \langle P_C(x) + P_C(z), x - P_C(x) + z - P_C(z) \rangle &= \underbrace{\langle P_C(x), x - P_C(x) \rangle}_{=0 \text{ selon (2.16)}} + \underbrace{\langle P_C(x), z - P_C(z) \rangle}_{=0 \text{ selon (2.20)}} + \\ &\quad \underbrace{\langle P_C(z), x - P_C(x) \rangle}_{=0 \text{ selon (2.20)}} + \underbrace{\langle P_C(z), z - P_C(z) \rangle}_{=0 \text{ selon (2.16)}} = 0 \end{aligned}$$

l'unicité de la décomposition de Moreau établie dans le Théorème 2.16 entraîne que $u = P_C(x)$ et $v = P_C(z)$ et par suite $x + z = P_C(x) + P_C(z)$.

(d) Soit $x \in H$, d'après la décomposition de Moreau, $x = P_C(x) + P_{C^\circ}(x)$ avec $\langle P_C(x), P_{C^\circ}(x) \rangle = 0$

Alors

$$\begin{aligned} \|P_C(x)\|^2 + \|P_{C^\circ}(x)\|^2 &= \|x - P_C(x)\|^2 + \|x - P_{C^\circ}(x)\|^2 \\ &= \|x - P_C(x)\|^2 + \|x - P_{C^\circ}(x)\|^2 + 2\underbrace{\langle x - P_C(x), x - P_{C^\circ}(x) \rangle}_{=0} \\ &= \|x - P_C(x) + x - P_{C^\circ}(x)\|^2 = \|x\|^2 \end{aligned}$$

(e) \subset

$$\begin{aligned} x \in C^\circ &\Leftrightarrow x = P_{C^\circ}(x) \\ &\Rightarrow x = P_{C^\circ}(x) + 0 \text{ et } \langle P_{C^\circ}(x), 0 \rangle = 0 \\ &\quad \xrightarrow[\text{Théorème 2.16}]{\text{unicité}} 0 = P_C(x) \end{aligned}$$

\supset

$$x \in H \xrightarrow{\text{Théorème 2.16}} x = \underbrace{P_C(x)}_{=0} + P_{C^\circ}(x) = P_{C^\circ}(x)$$

(f) C'est une conséquence directe de (d) en supprimant le terme $\|P_{C^\circ}(x)\|^2$.

□

Méthode des projections alternées MPA

Des nombreux et importants problèmes en mathématiques (optimisation, inégalités variationnelles, ... etc) peuvent être ramenés à chercher le point d'intersection d'un nombre fini d'ensembles $C_i, i = \overline{1, N}$ supposés généralement convexes. Ce genre de problème s'appelle **un problème d'admissibilité ou faisabilité convexe** et a la forme mathématique suivante

$$\text{Trouver un point } \bar{x} \in H \text{ tel que : } \bar{x} \in C := \bigcap_{i=1}^N C_i \quad (3.1)$$

Dans ce chapitre, nous allons considérer deux types d'ensembles. Dans la première partie on va supposer que les ensembles C_i sont des sous espaces vectoriels fermés, tandis que dans la deuxième, on s'intéresse à étudier le problème (3.1) en supposant que les ensembles C_i soient convexes fermés.

Dans la théorie d'optimisation (approximation) on s'intéresse généralement à chercher des points (des solutions) les plus proches possible ou à distance minimale par rapport à un ensemble donné (problème de minimisation), il est donc important d'améliorer les résultats obtenus par la résolution du problème (3.1) en considérant le problème suivant

$$\text{Trouver un point } \bar{x} \in C := \bigcap_{i=1}^N C_i \text{ tel que } \bar{x} = P_C(x). \quad (3.2)$$

3.1 Cas de deux sous espaces vectoriels

Notre objectif dans cette section consiste à étudier le problème (3.2) en considérant dans un premier temps, l'intersection de deux sous espaces vectoriels que l'on note V_1 et V_2 . Le prototype des algorithmes pour résoudre ce type de problèmes est la méthode des projections alternées, MPA en abrégé, que l'on va étudier dans ce chapitre. Cet algorithme a été introduit par J. Von Neumann [19], elle se base sur la construction d'une suite de

projections successives sur les ensembles considérés en espérant d'avoir la convergence vers le point $P_C(x)$.

Définition 3.1 (suite des projections alternées).

Soient V_1, V_2 deux sous-espace vectoriels fermés de H et $x \in H$, une suite des projections alternées est une suite de points $(x_n)_n$ de H vérifie la condition suivante

$$\begin{cases} x_0 = x \\ x_{2n+1} = P_{V_1}(x_{2n}) \quad \text{et} \quad x_{2n+2} = P_{V_2}(x_{2n+1}), n = 0, 1, 2, \dots \end{cases} \quad (3.3)$$

Il est clair que $x_{2n} \in V_2$ et $x_{2n+1} \in V_1$, ce qui implique $P_{V_1}(x_{2n+1}) = x_{2n+1}$ et $P_{V_2}(x_{2n}) = x_{2n}$. De plus, la suite $(x_n)_n$ peut être définie comme suit $x_0 = x$ et $x_{n+1} = (P_{V_2} \circ P_{V_1})(x_n)$, ou avec la relation suivante

$$x_0 = x \quad x_n = (P_{V_2} P_{V_1})^n x_0.$$

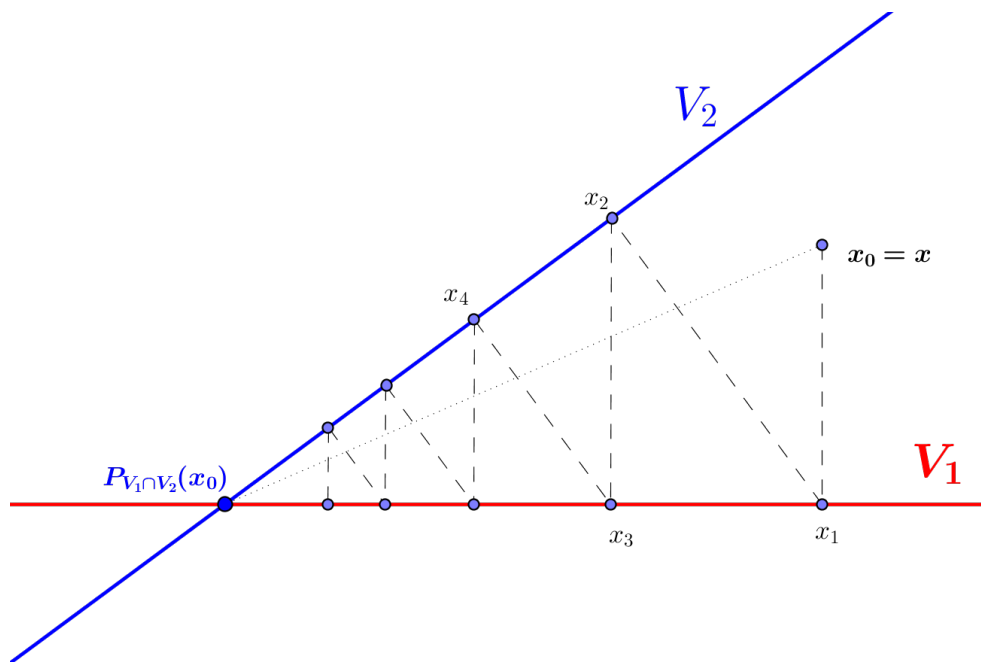


FIGURE 3.1 – Algorithme des projections alternées MAP

D'après la dernière définition, l'idée de cet algorithme est de répéter le processus suivant : projeter un point initial x sur V_1 , puis projeter le point obtenu sur V_2 , puis de nouveau sur V_1 , et ainsi de suite, comme l'illustre la Figure 3.1.

Théorème 3.2 (Théorème de J.Von Neumann [19]).

La suite $(x_n)_n$ des projections alternées définie dans (3.3) converge vers le point de projection de x sur $V_1 \cap V_2$, c-à-d,

$$\lim_{n \rightarrow \infty} x_n = P_{V_1 \cap V_2}(x).$$

Dorénavant, pour simplifier, on note $P_i = P_{V_i}, i = 1, 2$.

Esquisse de la démonstration Voici quelques points qui résument la preuve

- La suite $(\|x_n\|)_n$ est décroissante
- La sous suite (x_{2n}) est une suite de Cauchy dans H
- Toute la suite $(x_n)_n$ converge vers un élément $\bar{x} \in V_1 \cap V_2$.
- Le point \bar{x} obtenu est bien la projection de x sur $V_1 \cap V_2$.

Preuve. Comme $\|P_2\| \leq 1$ (voir la Proposition 2.14.(b)) alors

$$\|x_{2n}\| = \|P_2(x_{2n-1})\| \leq \|P_2\| \|x_{2n-1}\| \leq \|x_{2n-1}\| = \|P_1(x_{2n-2})\| \leq \|x_{2n-2}\| = \|x_{2(n-1)}\|$$

ceci traduit la décroissance de la suite des normes $\|x_n\|$ et par conséquent elle converge (parce qu'elle inférieurement bornée) vers une limite notée λ . En utilisant les faits que que les P_i sont auto-adjoints, $x_{2n} \in V_2$ et $x_{2n-1} \in V_1$ on trouve que pour tous $n, m \in \mathbb{N}$

$$\begin{aligned} \langle x_{2n}, x_{2m} \rangle &= \langle P_2(x_{2n-1}), x_{2m} \rangle = \langle x_{2n-1}, P_2(x_{2m}) \rangle = \langle x_{2n-1}, x_{2m} \rangle \\ &= \langle P_1(x_{2n-1}), x_{2m} \rangle = \langle x_{2n-1}, P_1(x_{2m}) \rangle = \langle x_{2n-1}, x_{2m+1} \rangle \end{aligned}$$

de façon similaire, on obtient

$$\langle x_{2n+1}, x_{2m+1} \rangle = \langle x_{2n}, x_{2m+2} \rangle, n, m \in \mathbb{N}$$

en Combinant ces deux dernières égalités, on déduit que pour tous $n, m \in \mathbb{N}$ avec $n \geq m$

$$\langle x_{2n}, x_{2m} \rangle = \langle x_{2n-1}, x_{2m+1} \rangle = \langle x_{2n-2}, x_{2m+2} \rangle = \dots = \langle x_{2n-k}, x_{2m+k} \rangle \text{ pour tout } 0 \leq k \leq 2n$$

en particulier, pour $k = n - m$ on obtient

$$\langle x_{2n}, x_{2m} \rangle = \langle x_{n+m}, x_{n+m} \rangle = \|x_{n+m}\|^2$$

donc, pour $m \leq n$

$$\begin{aligned} \|x_{2n} - x_{2m}\|^2 &= \|x_{2n}\|^2 - 2\langle x_{2n}, x_{2m} \rangle + \|x_{2m}\|^2 \\ &= \|x_{2n}\|^2 - 2\|x_{n+m}\|^2 + \|x_{2m}\|^2 \xrightarrow{m \rightarrow \infty} \lambda^2 - 2\lambda^2 + \lambda^2 = 0 \end{aligned}$$

ce qui exprime que x_{2n} est une suite de Cauchy dans V_2 qu'est fermé, alors elle converge vers une limite que l'on note $y \in V_2$. D'autre part

$$\begin{aligned} \langle x_{2n}, x_{2n-1} \rangle &= \langle P_2(x_{2n}), x_{2n-1} \rangle = \langle x_{2n}, P_2(x_{2n-1}) \rangle = \langle x_{2n}, x_{2n} \rangle = \|x_{2n}\|^2 \\ \Rightarrow \|x_{2n} - x_{2n-1}\|^2 &= \|x_{2n}\|^2 - 2\langle x_{2n}, x_{2n-1} \rangle + \|x_{2n-1}\|^2 \\ &= -\|x_{2n}\|^2 + \|x_{2n-1}\|^2 \xrightarrow{n \rightarrow \infty} -\lambda^2 + \lambda^2 = 0 \end{aligned}$$

donc

$$\|x_{2n-1} - y\| \leq \|x_{2n} - x_{2n-1}\| + \|x_{2n} - y\| \xrightarrow[n \rightarrow \infty]{} 0 \Rightarrow \lim_n x_{2n-1} = y$$

mais $(x_{2n-1})_n \subset V_1$ et V_1 est fermé, alors $y \in V_1$. On a démontré alors que

$$\lim_{n \rightarrow \infty} x_{2n} = \lim_{n \rightarrow \infty} x_{2n-1} = y \in V_1 \cap V_2 \Rightarrow \lim_{n \rightarrow \infty} x_n = y \in V_1 \cap V_2.$$

Pour terminer la preuve, il nous reste que de prouver que $y = P_{V_1 \cap V_2}(x)$. Pour cela, il suffit d'utiliser le Théorème 2.13 en montrant que $(x - y) \perp V_1 \cap V_2$. Soit $z \in V_1 \cap V_2$, alors

$$\begin{aligned} \langle x_{2n}, z \rangle &= \langle x_{2n} - x_{2n-1} + x_{2n-1}, z \rangle = \langle P_1(x_{2n-1}) - x_{2n-1} + x_{2n-1}, z \rangle \\ &= \underbrace{\langle P_1(x_{2n-1}) - x_{2n-1}, z \rangle}_{=0 \text{ selon (2.12)}} + \langle x_{2n-1}, z \rangle = \langle x_{2n-1}, z \rangle \end{aligned}$$

de même

$$\langle x_{2n-1}, z \rangle = \langle x_{2n-2}, z \rangle$$

par récurrence on trouve que pour tout $n \in \mathbb{N}$

$$\langle x_{2n}, z \rangle = \langle x_{2n-1}, z \rangle = \langle x_{2n-2}, z \rangle = \dots = \langle x_0, z \rangle = \langle x, z \rangle \Rightarrow \lim_n \langle x_{2n}, z \rangle = \langle x, z \rangle$$

alors

$$\begin{aligned} \langle x - y, z \rangle &= \langle x, z \rangle - \langle y, z \rangle = \langle x, z \rangle - \lim_n \langle x_{2n}, z \rangle = \langle x, z \rangle - \langle x, z \rangle = 0 \\ &\Rightarrow (x - y) \perp V_1 \cap V_2. \end{aligned}$$

□

Le passage de 2 à N sous espaces vectoriels n'est pas évident, néanmoins il y a une astuce qui permet d'étendre le théorème de Von Neumann aux plusieurs s.e.v. Ceci a été fait par Halperin (voir [14]) et le résultat est donné par le théorème suivant

Théorème 3.3 (Halperin[14]). *Soient $V_1, V_2 \dots V_N$, N sous-espaces vectoriels fermés de H et $x \in H$ alors*

$$\lim_{n \rightarrow \infty} (P_{V_N} P_{V_{N-1}} \dots P_{V_2} P_{V_1})^n(x) = P_{\bigcap_{i=1}^N V_i}(x)$$

3.1.1 Taux de convergence de MPA

Nous montrons que le taux de convergence va dépendre de l'angle entre les deux sous espaces. Ce dernier concept (angle) a une importance particulière et mérite d'être bien étudié. On sait que si x et y sont deux vecteurs alors l'angle entre eux se donne via l'équivalence suivante

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \|y\|} \Leftrightarrow \theta = \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

la question est donc : comment peut-t-on définir l'angle entre deux sous espaces vectoriels ?

La définition suivante, introduite pour la première fois par **Friedrichs** [13] en 1937, est la plus utilisée (acceptée) dans la littérature de MPA.

Définition 3.4. Soient M et N deux sous-espaces fermés de H . L'angle $\theta(M, N)$ entre M et N est un angle de $[0, \frac{\pi}{2}]$ dont le cosinus que l'on note $c(M, N) := \cos \theta(M, N)$ est défini par

$$c(M, N) := \sup\{|\langle x, y \rangle|, x \in M \cap (M \cap N)^\perp, \|x\| \leq 1, y \in N \cap (M \cap N)^\perp, \|y\| \leq 1\}.$$

Exemple 3.5. Soient M et N deux sous espaces de $l_2(2)$ définis par

$$M = \{x(\cdot) \in l_2(2), x(1) = x(2)\} \text{ et } N = \{x(\cdot) \in l_2(2), x(2) = 0\}$$

et munis de la norme définie dans (1.2) par

$$\|x\| = \left(|x(1)|^2 + |x(2)|^2 \right)^{\frac{1}{2}}.$$

Il est clair que M n'est que la première bissectrice et N est l'axe des abscisses. On va démontrer que

$$\theta(M, N) = \frac{\pi}{4}.$$

En effet, d'après la Figure 3.2, on voit que $M \cap N = \{0\}$ donc

$$(M \cap N)^\perp = \{0\}^\perp = l_2(2) \Rightarrow M \cap (M \cap N)^\perp = M \text{ et } N \cap (M \cap N)^\perp = N$$

soient $x \in M \cap (M \cap N)^\perp = M : \|x\| \leq 1$ et $y \in N \cap (M \cap N)^\perp = N : \|y\| \leq 1$ alors

$$\begin{aligned} \|x\| &= \left(|x(1)|^2 + |x(2)|^2 \right)^{\frac{1}{2}} = \left(|x(1)|^2 + |x(1)|^2 \right)^{\frac{1}{2}} = \sqrt{2}|x(1)| \leq 1 \\ &\Rightarrow |x(1)| = |x(2)| \leq \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \end{aligned}$$

et

$$\|y\| = \left(|y(1)|^2 + \underbrace{|y(2)|^2}_{=0} \right)^{\frac{1}{2}} = |y(1)| \leq 1$$

par conséquent

$$|\langle x, y \rangle| = |x(1)y(1)| \leq \frac{\sqrt{2}}{2}$$

ce qui entraîne que

$$\begin{aligned} c(M, N) &= \sup\{|\langle x, y \rangle|, x \in M \cap (M \cap N)^\perp, \|x\| \leq 1, y \in N \cap (M \cap N)^\perp, \|y\| \leq 1\} \\ &= \sup\{|x(1)y(1)|, x \in M, \|x\| \leq 1, y \in N, \|y\| \leq 1\} \leq \frac{\sqrt{2}}{2} \end{aligned}$$

ce dernier supremum est atteint lorsque $x(1) = x(2) = \frac{\sqrt{2}}{2}$ et $y(1) = 1$, ce qui signifie que

$$c(M, N) = \frac{\sqrt{2}}{2} \text{ avec } \theta(M, N) \in [0, \frac{\pi}{2}]$$

et par suite $\theta(M, N) = \frac{\pi}{4}$.

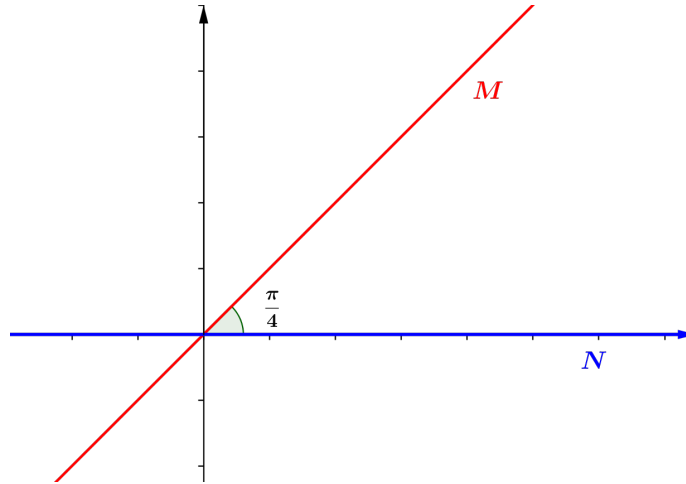


FIGURE 3.2 –

le lemme suivant résume les propriétés basiques de l'angle engendré par deux sous espaces.

Lemme 3.6. *Soient M et N deux sous-espaces fermés de H , alors*

- (a) $0 \leq c(M, N) \leq 1$ et $c(M, N) = c(N, M)$
- (b) $c(M, N) = C(M \cap (M \cap N)^\perp, N \cap (M \cap N)^\perp)$
- (c) $c(M, N) := \sup\{|\langle x, y \rangle|, x \in M, \|x\| \leq 1, y \in N \cap (M \cap N)^\perp, \|y\| \leq 1\}$
- (d) $c(M, N) := \sup\{|\langle x, y \rangle|, x \in M \cap (M \cap N)^\perp, \|x\| \leq 1, y \in N, \|y\| \leq 1\}$
- (e)

$$|\langle x, y \rangle| \leq c(M, N) \|x\| \|y\|$$

à chaque fois que $x \in M, y \in N$ et au moins l'un de x ou y est dans $(M \cap N)^\perp$. La dernière inégalité représente une version plus forte de celle de Cauchy-Schwartz pour certains paires de vecteurs $x, y \in H$. Pour la preuve de ce lemme on se réfère à [8]

Nous allons montrer dans le théorème suivant que la vitesse de convergence de MPA dépend de l'angle entre les deux sous espaces en question, de façon plus précise, la suite (x_n) générées par la MPA s'approche de l'intersection $V_1 \cap V_2$ à vitesse géométrique.

Théorème 3.7 (Taux de convergence). *Soient M_1 et M_2 deux sous-espaces fermés de H et $c := c(M_1, M_2)$ alors, pour tout $\forall x \in H$*

$$\|(P_{M_2} P_{M_1})^n(x) - P_{M_1 \cap M_2}(x)\| \leq c^{2n-1} \|x - P_{M_1 \cap M_2}(x)\| \leq c^{2n-1} \|x\| \quad (3.4)$$

cela signifie que la vitesse de convergence de la MPA est directement proportionnelle à l'angle entre les deux sous espaces, c-à-d, plus que l'angle est grand (resp. petit), plus que la convergence est rapide (resp. lente).

Preuve. Soient $M = M_1 \cap M_2, P_i = P_{M_i}, i = 1, 2$ et $P = P_2 P_1$. On montre dans un premier temps que

$$P_i(M^\perp) \subset M^\perp, i = 1, 2 \quad (3.5)$$

ce qui donne

$$P(M^\perp) \subset M^\perp \quad (3.6)$$

pour cela, soient $x \in M^\perp$ et $y \in M$, on utilise les faits que $P_i, i = 1, 2$ sont auto-adjoints et idempotents on trouve

$$\langle P_i(x), y \rangle = \langle x, P_i(y) \rangle = \langle x, P_2(P_1(y)) \rangle \underset{y \in M_1}{=} \langle x, P_2(y) \rangle \underset{y \in M_2}{=} \langle x, y \rangle = 0$$

Ce qui traduit l'inclusion $P_i(M^\perp) \subset M^\perp$ et par suite $P(M^\perp) \subset M^\perp$, en effet

$$P(M^\perp) = P_2 \underbrace{[P_1(M^\perp)]}_{\subset M^\perp} \subset P_2(M^\perp) \subset M^\perp$$

on montre à présent que

$$\|P(x)\| \leq c \|P_1(x)\| \leq c \|x\|, \text{ pour tout } x \in M^\perp \quad (3.7)$$

Soit $x \in M^\perp$ alors

$$\langle P(x) - P_1(x), P(x) \rangle = \langle P_{M_2}(P_1(x)) - P_1(x), \underbrace{P(x)}_{\in M_2} \rangle \underset{(2.12)}{=} 0$$

en combinant cette dernière égalité avec les inclusions (3.5), (3.6) et la propriété (e) du Lemme 3.6 on trouve

$$\begin{aligned} \|P(x)\|^2 &= \langle P(x), P(x) \rangle = \langle P(x) - P_1(x) + P_1(x), P(x) \rangle \\ &= \langle P_1(x), P(x) \rangle \leq c \|P_1(x)\| \|P(x)\| \leq c \|x\| \|P(x)\| \text{ car } \|P_i\| \leq 1 \end{aligned}$$

en divisant les deux membres de l'inégalité par $\|P_C(x)\|$ on obtient (3.7).

Maintenant, soit $x \in P(M^\perp)$, alors $x \in M_2 \cap M^\perp$ par (3.6). D'autre part, on a $P_1(x) \in M_1 \cap M^\perp$ grâce à (3.5), on rapplique à nouveau le Lemme 3.6 (e) on obtient

$$\|P_1(x)\|^2 = \langle P_1(x), P_1(x) \rangle = \langle x, P_1(x) \rangle \leq c \|x\| \|P_1(x)\| \Rightarrow \|P_1(x)\| \leq c \|x\|$$

l'inégalité (3.7) devient

$$\|P(x)\| \leq c \|P_1(x)\| \leq c^2 \|x\| \text{ pour tout } x \in P(M^\perp) \quad (3.8)$$

il est facile de vérifier en utilisant la récurrence (induction) que

$$\|P(x)\| \leq c^{2n} \|x\| \text{ pour tout } x \in P(M^\perp) \quad (3.9)$$

en effet, supposons que cette dernière inégalité ait lieu pour $n - 1$, alors pour tout $x \in P(M^\perp) \subset M^\perp$ (donc $P(x) \in P(M^\perp)$)

$$\|P^n(x)\| = \|P^{n-1}(\underbrace{P(x)}_{\in P(M^\perp)})\| \leq c^{2(n-1)}\|P(x)\| \stackrel{(3.8)}{\leq} c^{2(n-1)}c^2\|x\| = c^{2n}\|x\|$$

en tenant compte de (3.6), (3.7) et (3.9) on trouve

$$\begin{aligned} \|(P_{M_2}P_{M_1})^n(x) - P_{M_1 \cap M_2}(x)\| &= \|(P_2P_1)^n(x) - P_M(x)\| = \|P^n(x) - P^n P_M(x)\| \\ &= \|P^n[x - P_M(x)]\| = \|P^{n-1}[P(\underbrace{x - P_M(x)}_{\in M^\perp})]\| \\ &\leq c^{2(n-1)}\|P(x - P_M(x))\| \\ &\leq c^{2n-1}\|x - P_M(x)\| \leq c^{2n-1}\|x\| \end{aligned}$$

ce qui termine la preuve. □

Remarque 3.8. *Le Théorème 3.7 montre facilement que si $c(M_1, M_2) = 0$, c-à-d, $\theta(M_1, M_2) = \frac{\pi}{2}$ alors la suite des projections alternées converge après deux itération ($n = 2$) et on a $P_{M_2}P_{M_1} = P_{M_1 \cap M_2}$.*

3.2 Cas de plusieurs ensembles convexes fermés : algorithme de Dykstra

On est tenté d'étendre l'algorithme des projections alternées de Von Neumann au cas de deux convexes fermés C_1 et C_2 (pas nécessairement des sous-espaces ou des hyperplans), et de penser que la suite ainsi construite converge vers la projection du point donnée x sur $C_1 \cap C_2$. le contre exemple suivant montre que ce résultat n'est plus valable dans cette situation même si C_1 et C_2 sont deux demi-espaces fermés. Considérons les deux ensembles suivants

$$\begin{aligned} C_1 &= \{(y, z) \in \mathbb{R}^2 : 0 \leq y \leq z\} \\ C_2 &= \{(y, z) \in \mathbb{R}^2 : z = 1\} \end{aligned}$$

Les ensembles C_1 et C_2 sont deux demi-espaces fermés de \mathbb{R}^2 , ils sont alors convexes. Supposons que le point initiale est $x = x_0 = (1, -1)$ alors $P_{C_1}(x_0) = (0, 0)$. D'après la Figure 3.3, il est clair que la suite $(x_n)_n$ générée par la MPA converge vers le point $x_2 = (0, 1) \in C_1 \cap C_2$ après deux itérations, tandis que le point de projection de x sur $C_1 \cap C_2$ est $(1, 1)$, autrement dit

$$x_n = (P_{C_2}P_{C_1})^n(x) = x_2, \forall n \geq 2, \quad \text{et} \quad \lim x_n \neq P_{C_1 \cap C_2}(x)$$

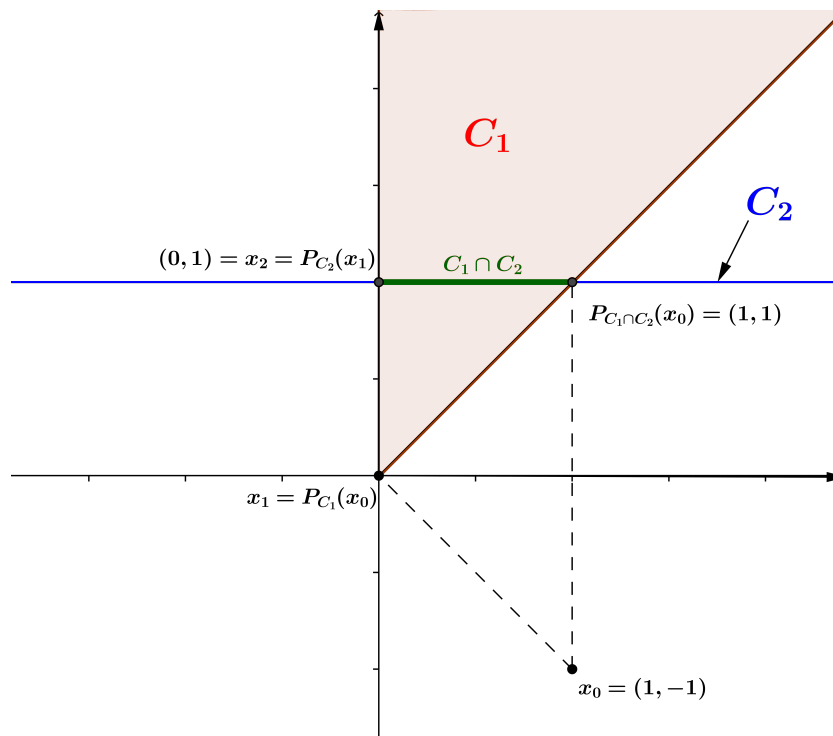


FIGURE 3.3 – Algorithme de MAP ne converge pas vers le point de projection

Il est à noter que le passage vers les ensembles convexes fermés dans un espace de Hilbert général (n'est pas nécessairement de dimension finie) conduit à perdre la convergence forte de la suite de Von Neumann (MPA) vers le point de l'intersection des C_i . Néanmoins, Bregman [4] a réussi à donner une extension "faible" du théorème de Von Neumann. Ce résultat est donné par le théorème suivant

Théorème 3.9 (Bregman [4]). *Soit $x_0 \in H$ et soient $C_i, i = \overline{1, r}$, r ensembles convexes fermés d'intersection non vide, alors*

$$\left(P_{C_r} P_{C_{r-1}} \cdots P_{C_1} \right)^n (x_0) \xrightarrow{w} x^* \in \bigcap_{i=1}^r C_i$$

autrement dit, la suite des projections alternées converge faiblement vers un point de l'intersection qui n'est pas nécessairement le point $P_{\bigcap_{i=1}^r C_i}(x)$.

Il est clair que cette dernière convergence devient forte lorsque la dimension de l'espace H est fini (\mathbb{R}^n par exemple) grâce à la coïncidence, dans ce cas, entre la topologie faible et la topologie forte.

Heureusement, Dykstra [9] a trouvé une modification intelligente du schéma de Von Neumann pour lequel la convergence vers le point de l'intersection le plus proche au point initial est garantie. La méthode a été développée à l'origine pour projeter un point donné

dans un espace de dimension finie sur un cône convexe fermé K ([9]) qui s'écrit sous la forme

$$K = \bigcap_{i=1}^r K_i, r \in \mathbb{N}, K_i \text{ cône convexe fermé}$$

Notre objectif dans cette partie consiste à décrire un algorithme permettant de résoudre le problème suivant

$$\min_{y \in K} \|x - y\| \text{ avec } K = \bigcap_{i=1}^r K_i \quad (3.10)$$

où $K_i, i = \overline{1, r}$ sont des ensembles **convexes fermés** (**pas nécessairement des sous espaces**).

La solution de problème (3.10) se donne via l'algorithme de Dykstra [9] qui utilise la projection sur les ensembles K_i au lieu de l'ensemble K . Cet algorithme construit deux suites, des itérations $(x_n)_n$ et des **correcteurs** $(e_n)_n$ par les relations de récurrence suivantes

$$\begin{aligned} x_0 &:= x, \quad e_{-(r-1)} = e_{-(r-2)} = \dots = e_{-1} = e_0 = 0 \\ x_n &:= P_{K_{[n]}}(x_{n-1} + e_{n-r}), \\ e_n &:= x_{n-1} + e_{n-r} - x_n = x_{n-1} + e_{n-r} - P_{K_{[n]}}(x_{n-1} + e_{n-r}) \end{aligned} \quad (3.11)$$

où

$$[n] := \{1, 2, \dots, r\} \cap \{n - kr, k = 0, 1, 2, \dots\}$$

autrement dit

$$[n] = \begin{cases} n \bmod r & \text{si } n \neq kr \\ r & \text{si } n = kr \end{cases}$$

où $n \bmod r$ représente le reste de la division euclidienne de n sur r , par exemple

$$[1] = 1, [2] = 2, \dots, [r-1] = r-1, [r] = r, [r+1] = 1, [r+2] = 2, \dots, [2r] = r, \dots$$

La figure 3.4 donne une illustration de l'algorithme de Dykstra.

Théorème 3.10 (Boyle et Dykstra [3]).

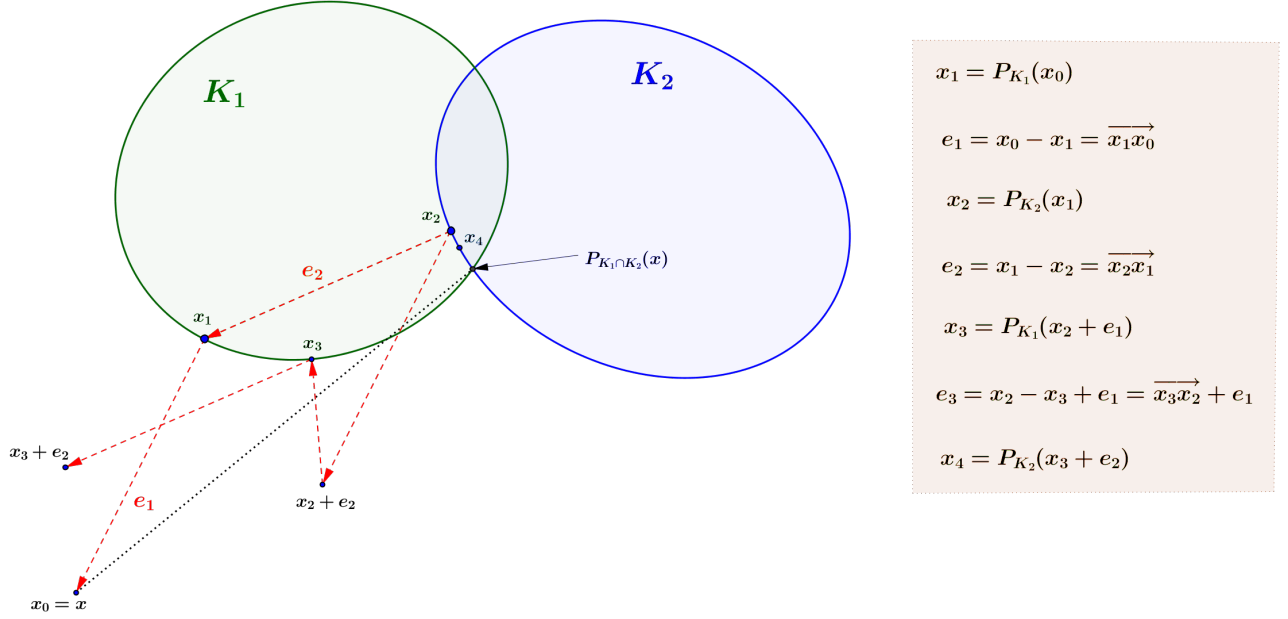
La suite $(x_n)_n$ générée par l'algorithme de Dykstra converge vers $P_K(x)$, c-à-d

$$\lim_n x_n = P_{\bigcap_{i=1}^r K_i}(x).$$

Il est commode de décomposer la preuve de ce théorème à plusieurs parties que nous désignons comme des lemmes. Pour la simplicité de la notation, on pose $P_i = P_{K_i}, i = 1, 2, \dots, r$.

Lemme 3.11. *Pour tout n*

$$\langle x_n - z, e_n \rangle \geq 0 \quad \forall z \in K_{[n]}$$

FIGURE 3.4 – Algorithme de Dykstra avec deux ensembles convexes ($r = 2$)

Preuve. en utilisant l'inégalité (2.8) avec $x = x_{n-1} + e_{n-r}$, $y = P_{[n]}(x_{n-1} + e_{n-r})$ et $K = K_{[n]}$, on obtient, pour tout $z \in K_{[n]}$

$$\langle x_n - z, e_n \rangle = \langle P_{[n]}(x_{n-1} + e_{n-r}) - z, x_{n-1} + e_{n-r} - P_{[n]}(x_{n-1} + e_{n-r}) \rangle \geq 0$$

□

Lemme 3.12. Pour tout $n \in \mathbb{N}$

$$x - x_n = e_{n-(r-1)} + e_{n-(r-2)} + \dots + e_{n-1} + e_n. \quad (3.12)$$

Preuve. on utilise la récurrence sur n , pour $n = 0$, $x - x_0 = x - x = 0$ et $e_{-(r-1)} + e_{-(r-2)} + \dots + e_{-1} + e_0 = 0$ par (3.11). Supposons que le résultat est vrai pour $n \geq 0$ alors

$$\begin{aligned} x - x_{n+1} &= (x - x_n) + (x_n - x_{n+1}) \\ &= (e_{n-(r-1)} + e_{n-(r-2)} \dots + e_{n-1} + e_n) + (e_{n+1} - e_{n+1-r}) \\ &= e_{n-(r-2)} + e_{n-(r-3)} + \dots + e_n + e_{n+1} \\ &= e_{n+1-(r-1)} + e_{n+1-(r-2)} + \dots + e_n + e_{n+1} \end{aligned}$$

Ce qui justifie la validité de (3.12) pour $n + 1$.

□

Lemme 3.13. Pour tous $n \in \mathbb{N}$, $0 \leq m \leq n$ et $y \in K$,

$$\|x_m - y\|^2 = \|x_n - y\|^2 + \sum_{k=m+1}^n \|x_k - x_{k-1}\|^2 + 2 \sum_{k=m+1}^n \langle e_{k-r}, x_{k-r} - x_k \rangle$$

$$+2 \sum_{k=n-(r-1)}^n \langle e_k, x_k - y \rangle - 2 \sum_{k=m-(r-1)}^m \langle e_k, x_k - y \rangle. \quad (3.13)$$

Preuve. Soit $\{y_m, y_{m+1}, \dots, y_{n+1}\} \subset H$ un ensemble de vecteurs de H alors

$$\begin{aligned} \|y_m - y_{n+1}\|^2 &= \|(y_m - y_{m+1}) + (y_{m+1} - y_{m+2}) + \dots + (y_n - y_{n+1})\|^2 \\ &= \sum_{k=m+1}^{n+1} \|y_{k-1} - y_k\|^2 + 2 \sum_{m+1 \leq i < j \leq n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \\ &= \|y_n - y_{n+1}\|^2 + \sum_{k=m+1}^n \|y_{k-1} - y_k\|^2 \\ &\quad + 2 \sum_{i=m+1}^n \left(\sum_{j=i+1}^{n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \right) \end{aligned} \quad (3.14)$$

mais

$$\begin{aligned} \sum_{i=m+1}^n \left(\sum_{j=i+1}^{n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \right) &= \sum_{i=m+1}^n \left\langle y_{i-1} - y_i, \sum_{j=i+1}^{n+1} (y_{j-1} - y_j) \right\rangle \\ &= \sum_{i=m+1}^n \langle y_{i-1} - y_i, y_i - y_{n+1} \rangle \end{aligned} \quad (3.15)$$

en substituant y_{n+1} par y et y_i par x_i pour tout $i \leq n$ dans (3.15) on trouve

$$\begin{aligned} \sum_{i=m+1}^n \langle y_{i-1} - y_i, y_i - y_{n+1} \rangle &= \sum_{i=m+1}^n \langle x_{i-1} - x_i, x_i - y \rangle \\ &= \sum_{i=m+1}^n \langle e_i - e_{i-r}, x_i - y \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1}^n \left[\langle e_{i-r}, x_i - x_{i-r} \rangle + \langle e_{i-r}, x_{i-r} - y \rangle \right] \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1-r}^{n-r} \langle e_i, x_i - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle \\ &= \sum_{i=n-r+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m-r+1}^m \langle e_i, x_i - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle. \end{aligned}$$

en combinant cette dernière égalité avec celles dans (3.14) et (3.15) tout en gardant les substitutions considérées ci-dessus on obtient

$$\begin{aligned} \|x_m - y\|^2 &= \|x_n - y\|^2 + \sum_{k=m+1}^n \|x_k - x_{k-1}\|^2 + 2 \sum_{k=n-(r-1)}^n \langle e_k, x_k - y \rangle \\ &\quad + 2 \sum_{k=m+1}^n \langle e_{k-r}, x_{k-r} - x_k \rangle - 2 \sum_{k=m-(r-1)}^m \langle e_k, x_k - y \rangle \end{aligned}$$

□

Lemme 3.14. *La suite (x_n) générée par l'algorithme de Dykstra est une suite bornée, de plus*

$$\sum_1^{\infty} \|x_{k-1} - x_k\| < +\infty, \quad (3.16)$$

en particulier

$$\lim_{n \rightarrow \infty} \|x_{n-1} - x_n\| = 0 \quad (3.17)$$

Preuve. Posons $m = 0$ dans l'égalité (3.13) on obtient

$$\begin{aligned} \|x_0 - y\|^2 &= \|x_n - y\|^2 + \sum_{k=1}^n \|x_k - x_{k-1}\|^2 + 2 \sum_{k=1}^n \underbrace{\langle e_{k-r}, x_{k-r} - x_k \rangle}_{\geq 0 \text{ selon Lemme 3.11}} \\ &+ 2 \sum_{k=n-(r-1)}^n \underbrace{\langle e_k, x_k - y \rangle}_{\geq 0 \text{ selon Lemme 3.11}} \\ &\geq \|x_n - y\|^2 + \sum_{k=1}^n \|x_k - x_{k-1}\|^2 \text{ pour tout } n \in \mathbb{N} \end{aligned}$$

ce qui implique que la suite $(x_n)_n$ est bornée et justifie (3.16). La convergence dans (3.17) est une conséquence directe de (3.16). \square

Lemme 3.15. *Pour tout $n \in \mathbb{N}$*

$$\|e_n\| \leq \sum_{k=1}^n \|x_{k-1} - x_k\| \quad (3.18)$$

Preuve. On utilise la récurrence sur n , Pour $n = 1$

$$\|e_1\| = \|x_0 - x_1 - e_{1-r}\| = \|x_0 - x_1\| \text{ car } e_i = 0, \forall i \leq 0$$

supposons maintenant que (3.18) est vraie pour tout $m \leq n$, alors

$$\begin{aligned} \|e_{n+1}\| &= \|x_n - x_{n+1} + e_{n+1-r}\| \leq \|x_n - x_{n+1}\| + \|e_{n+1-r}\| \\ &\leq \|x_n - x_{n+1}\| + \sum_{k=1}^{n+1-r} \|x_{k-1} - x_k\| \leq \sum_{k=1}^{n+1} \|x_{k-1} - x_k\| \text{ car } n+1-r \leq n \end{aligned}$$

Ce qui traduit la validité de (3.18) pour $n+1$. \square

Lemme 3.16.

$$\liminf_n \sum_{k=n-(r-1)}^n |\langle x_k - x_n, e_k \rangle| = 0. \quad (3.19)$$

Preuve. L'inégalité de Schwartz et le Lemme 3.15, donnent

$$\begin{aligned}
\sum_{k=n-(r-1)}^n |\langle x_k - x_n, e_k \rangle| &\leq \sum_{k=n-(r-1)}^n \|e_k\| \|x_k - x_n\| \\
&\leq \sum_{k=n-(r-1)}^n \left(\sum_{j=1}^k \|x_{j-1} - x_j\| \right) \left(\sum_{i=k+1}^n \|x_{i-1} - x_i\| \right) \\
&\leq r \sum_{j=1}^n \|x_{j-1} - x_j\| \left(\sum_{i=n-(r-2)}^n \|x_{i-1} - x_i\| \right)
\end{aligned}$$

en posant $a_i = \|x_{i-1} - x_i\|$, alors il suffit de montrer que

$$\liminf_n \sum_{j=1}^n a_j \left(\sum_{i=n-(r-2)}^n a_i \right) = 0 \quad (3.20)$$

d'après le lemme 3.14, on a

$$A := \sum_{i=1}^{\infty} a_i^2 = \sum_{i=1}^{\infty} \|x_{i-1} - x_i\|^2 < \infty$$

On utilisant l'inégalité de Schwartz, on obtient

$$\sum_{j=1}^n a_j \leq \sqrt{n} \left(\sum_{j=1}^n a_j^2 \right)^{\frac{1}{2}} \leq \sqrt{n} A^{\frac{1}{2}}$$

alors, pour avoir (3.20), il suffit de prouver

$$\liminf_n \sqrt{n} \sum_{i=n-(r-2)}^n a_i = 0. \quad (3.21)$$

Pour cela, on pose

$$\rho := \liminf_n \sqrt{n} \sum_{i=n-(r-2)}^n a_i$$

On suppose que (3.21) soit fautive, alors $\rho > 0$ doit être strictement positif. On suppose que $\rho < \infty$, (la preuve pour $\rho = \infty$ est similaire) alors

$$\begin{aligned}
\sqrt{n} \sum_{i=n-(r-2)}^n a_i &> \frac{1}{2} \rho \text{ pour } n \text{ suffisamment grand} \\
\Rightarrow \sum_{i=n-(r-2)}^n a_i &> \frac{\rho}{(2\sqrt{n})}
\end{aligned}$$

en utilisant l'inégalité de Schwartz

$$\frac{\rho^2}{4n} < \left(\sum_{i=n-(r-2)}^n a_i \right)^2 \leq (r-1) \sum_{i=n-(r-2)}^n a_i^2$$

pour certain nombre entier N on obtient, alors

$$\frac{\rho^2}{4n} < (r-1) \sum_{i=n-(r-2)}^n a_i^2 \text{ pour tout } n \geq N.$$

Par conséquent

$$\begin{aligned} \infty &= \sum_{n=N}^{\infty} \frac{\rho^2}{4n} \leq (r-1) \sum_{n=N}^{\infty} \sum_{i=n-(r-2)}^n a_i^2 \\ &\leq (r-1) \sum_{n=N}^{\infty} [a_{n-(r-2)}^2 + a_{n-(r-3)}^2 + \dots + a_n^2] \leq (r-1)^2 \sum_1^{\infty} a_i^2 < \infty, \end{aligned}$$

Ce qui est absurde. Donc (3.20) est vraie. \square

Lemme 3.17. *Il existe une sous suite x_{n_j} de x_n telle que*

$$\limsup_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0 \text{ pour tout } y \in K \quad (3.22)$$

et

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| = 0. \quad (3.23)$$

Preuve. En utilisant le Lemme 3.12, on trouve que pour tout $y \in K$

$$\begin{aligned} \langle y - x_n, x - x_n \rangle &= \langle y - x_n, e_{n-(r-1)} + e_{n-(r-2)} + \dots + e_n \rangle \\ &= \sum_{k=n-(r-1)}^n \langle y - x_n, e_k \rangle \\ &= \sum_{k=n-(r-1)}^n \langle y - x_k, e_k \rangle + \sum_{k=n-(r-1)}^n \langle x_k - x_n, e_k \rangle \end{aligned} \quad (3.24)$$

d'après le Lemme 3.11, la premier somme est négative ce qui entraîne

$$\langle y - x_n, x - x_n \rangle \leq \sum_{k=n-(r-1)}^n \langle x_k - x_n, e_k \rangle. \quad (3.25)$$

les Lemmes 3.16 et 1.20 assurent l'existence d'une sous suite $(x_{n_j})_j$ de $(x_n)_n$ telle que

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| = 0. \quad (3.26)$$

$$\begin{aligned} (3.25) \Rightarrow \langle y - x_{n_j}, x - x_{n_j} \rangle &\leq \sum_{k=n_j-(r-1)}^{n_j} \langle x_k - x_{n_j}, e_k \rangle \\ &\leq \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| \end{aligned}$$

$$\begin{aligned}
\Rightarrow \limsup_j \langle y - x_{n_j}, x - x_{n_j} \rangle &\leq \limsup_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| \\
&= \lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| = 0.
\end{aligned}$$

□

Avec tous ces instruments établis dans les lemmes précédents, on est prêt et capable de démontrer le résultat principale de cette partie donné par le théorème de Boyle et Dykstra.

Preuve. le Lemme 3.17 assure l'existence d'une sous suite x_{n_j} de (x_n) telle que

$$\limsup_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0 \text{ pour tout } y \in K. \quad (3.27)$$

comme x_n est bornée par le Lemme 3.14, et d'après le Théorème 1.24 il existe une sous suite de $(x_{n_j})_j$ que l'on note aussi $(x_{n_j})_j$ et un élément $y_0 \in H$ tels que

$$x_{n_j} \xrightarrow{w} y_0 \quad (3.28)$$

de plus, selon (1.5) on obtient

$$\|y_0\| \leq \liminf_j \|x_{n_j}\| = \lim_j \|x_{n_j}\|. \quad (3.29)$$

Comme il ya un nombre finie d'ensembles $K_i, i = 1, \dots, r$ alors un nombre infini des termes $(x_{n_j})_j$ doivent appartenir à un seul ensemble K_{i_0} . Selon les données, l'ensemble K_{i_0} est convexe fermé, par conséquent, il est faiblement fermé grâce au Théorème 1.23 ce qui assure l'inclusion $y_0 \in K_{i_0}$.

On utilise le fait que $(x_n - x_{n-1})_n$ converge vers 0, ou déduite que toutes les suites $(x_{n_j+1})_j, (x_{n_j+2})_j, (x_{n_j+3})_j \dots$ convergent faiblement vers y_0 . En effet

$$\begin{aligned}
\langle f, y_0 - x_{n_j+1} \rangle &= \langle f, y_0 - x_{n_j} + x_{n_j} - x_{n_j+1} \rangle \\
&= \langle f, y_0 - x_{n_j} \rangle + \langle f, x_{n_j} - x_{n_j+1} \rangle \\
&= \underbrace{[\langle f, y_0 \rangle - \langle f, x_{n_j} \rangle]}_{=0} - \underbrace{\langle f, x_{n_j} - x_{n_j+1} \rangle}_{=0}
\end{aligned}$$

et par suite $y_0 \in K_i$ pour tout $i = 1, \dots, r$ ce qui exprime $y_0 \in K$.

Pour tout $y \in K$, (3.29) et (3.27) impliquent que

$$\begin{aligned}
\langle y - y_0, x - y_0 \rangle &= \langle y, x \rangle - \langle y, y_0 \rangle - \langle y_0, x \rangle + \|y_0\|^2 \\
&\leq \lim_j (\langle y, x \rangle - \langle y, x_{n_j} \rangle - \langle x_{n_j}, x \rangle + \|x_{n_j}\|^2) \\
&= \lim_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0
\end{aligned}$$

ce qui équivale à (selon (2.8))

$$y_0 = P_K(x). \quad (3.30)$$

De plus, en posant $y = y_0$ dans les inégalités ci-dessus, on obtient

$$\lim_j \|x_{n_j}\|^2 = \|y_0\|^2 \quad (3.31)$$

et

$$\lim_j \langle y_0 - x_{n_j}, x - x_{n_j} \rangle = 0 \quad (3.32)$$

en tenant compte de (3.28), (3.31) et le Lemme 1.22 on obtient

$$\lim_{j \rightarrow \infty} \|x_{n_j} - y_0\| = 0. \quad (3.33)$$

donc

$$\|x_{n_j} - P_K(x)\| = \|x_{n_j} - y_0\| \rightarrow 0. \quad (3.34)$$

Pour terminer la preuve, il faut montrer que la suite (x_n) converge vers y_0 . De l'équation (3.24) avec $y = y_0$ et $n = n_j$, on trouve

$$\langle y_0 - x_{n_j}, x - x_{n_j} \rangle = \sum_{k=n_j-(r-1)}^{n_j} \langle y_0 - x_k, e_k \rangle + \sum_{k=n_j-(r-1)}^{n_j} \langle x_k - x_{n_j}, e_k \rangle \quad (3.35)$$

le premier membre de (3.35) tend vers zéro lorsque $j \rightarrow \infty$ grâce à (3.32), et la deuxième somme du second membre tend vers zéro par (3.25). Par conséquent

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} \langle y_0 - x_k, e_k \rangle = 0. \quad (3.36)$$

en utilisant les Lemme 3.13 et 3.11 avec $m = n_j$ et $y = y_0$ respectivement, on trouve que pour tout $n \geq n_j$

$$\|x_{n_j} - y_0\|^2 \geq \|x_n - y_0\|^2 - 2 \sum_{k=n_j-(r-1)}^{n_j} \langle e_k, x_k - y_0 \rangle,$$

ou

$$\|x_n - y_0\|^2 \leq \|x_{n_j} - y_0\|^2 + 2 \sum_{k=n_j-(r-1)}^{n_j} \langle e_k, x_k - y_0 \rangle. \quad (3.37)$$

Mais le deuxième termes du second membre de (3.37) tendent vers zéro quand $j \rightarrow \infty$ grâce à (3.33) et (3.36), donc

$$\lim_n \|x_n - y_0\| = 0,$$

ce qui termine la preuve. □

3.3 Application : Systèmes linéaires

Dans cette section, nous donnons un exemple où on peut appliquer la méthode des projections alternées. Pour cela, on considère le système linéaire suivant

$$Ax = b, A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R}), x = (x_i), b = (b_i) \in \mathbb{R}^n$$

On suppose que ce système possède une solution unique (A inversible). On note $a_i, i = \overline{1, n}$ la $i^{\text{ème}}$ ligne de la matrice A , c-à-d,

$$a_i = (a_{i1} \ a_{i2} \ \cdots \ a_{in}), i = \overline{1, n}$$

Soit $H_i \subset \mathbb{R}^n$ l'hyperplan défini par

$$H_i = \{x \in \mathbb{R}^n : \langle a_i, x \rangle = b_i\}, i = \overline{1, n}$$

Il est clair que les hyperplans $H_i, i = \overline{1, n}$ sont convexes fermés. En effet, soient $x, y \in H_i$ et $\lambda \in [0, 1]$, alors

$$\begin{aligned} \langle a_i, \lambda x + (1 - \lambda)y \rangle &= \lambda \langle a_i, x \rangle + (1 - \lambda) \langle a_i, y \rangle \\ &= \lambda b_i + (1 - \lambda)b_i = b_i \Rightarrow \lambda x + (1 - \lambda)y \in H_i. \end{aligned}$$

par conséquent, la projection sur chaque H_i est bien définie. D'autre part, on a

$$\begin{aligned} Ax = b &\Leftrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} x = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \\ &\Leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} x = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Leftrightarrow \begin{cases} \langle a_1, x \rangle = b_1 \\ \vdots \\ \langle a_n, x \rangle = b_n \end{cases} \\ &\Leftrightarrow x \in \bigcap_{i=1}^n H_i \end{aligned}$$

On déduit que résoudre le système linéaire $Ax = b$ est équivalent à chercher le point d'intersection des hyperplans $H_i, i = \overline{1, n}$, ce qui nous conduit à utiliser la méthode des projections alternées pour trouver ce dernier point. La suite générée a donc la forme suivante

$$x_0 \in \mathbb{R}^n, \quad x_m = \left(P_{H_n} P_{H_{n-1}} \cdots P_{H_1} \right)^m (x_0), m \geq 1$$

Comme la dimension de l'espace \mathbb{R}^n est finie alors selon le Théorème de Bregman 3.9, cette dernière suite converge fortement vers le point d'intersection qui représente aussi la solution du système $Ax = b$.

Afin que l'on puisse appliquer cette méthode, il est clair que l'on doit tout d'abord calculer la projection sur chaque hyperplan $H_i, i = \overline{1, n}$ ce qu'est donné par la proposition suivante

Proposition 3.18. Soient $z \in H, c \in \mathbb{R}$ et soit H_z l'hyperplan défini par

$$H_z = \{y \in H : \langle z, y \rangle = c\}$$

alors

$$P_{H_z}(x) = x + \frac{c - \langle z, x \rangle}{\langle z, z \rangle} z \quad (3.38)$$

Preuve. On suppose dans un premier temps que $c = 0$, c-à-d, $H_z = \{y \in H : \langle z, y \rangle = 0\}$ ce qui implique que H_z est un s.e.v. Pour tout $y \in H_z$ on a

$$\left\langle x - \left(x + \frac{-\langle z, x \rangle}{\langle z, z \rangle} z \right), y \right\rangle = \frac{\langle z, x \rangle}{\langle z, z \rangle} \underbrace{\langle z, y \rangle}_{=0} = 0$$

ce qui exprime bien (3.38) grâce à la caractérisation de la projection sur un s.e.v donnée par (2.12). Supposons maintenant que $c \neq 0$. Soit $x_0 \in H_z$ et soit H_0 l'hyperplan défini par $H_0 := H_z - x_0$, alors

$$H_0 = \{y \in H : \langle z, y \rangle = 0\}$$

en effet,

$$\begin{aligned} v \in H_0 &\Leftrightarrow v \in H_z - x_0 \Leftrightarrow v + x_0 \in H_z \Leftrightarrow \langle z, v + x_0 \rangle = c \\ &\Leftrightarrow \langle z, v \rangle + \underbrace{\langle z, x_0 \rangle}_{=c} = c \\ &\Leftrightarrow \langle z, v \rangle = 0 \Leftrightarrow v \in \{y \in H : \langle z, y \rangle = 0\} \end{aligned}$$

donc

$$\begin{aligned} P_{H_z}(x) &= P_{H_0+x_0}(x) \stackrel{(2.9)}{=} P_{H_0}(x - x_0) + x_0 \\ &\stackrel{\text{cas } c=0}{=} x - x_0 - \frac{\langle z, x - x_0 \rangle}{\langle z, z \rangle} z + x_0 = x + \frac{\langle z, x_0 \rangle - \langle z, x \rangle}{\langle z, z \rangle} z \\ &= x + \frac{c - \langle z, x \rangle}{\langle z, z \rangle} z. \end{aligned}$$

□

En utilisant cette dernière proposition on arrive à définir P_{H_i} comme suit

$$P_{H_i}(x) = x + \frac{b_i - \langle a_i, x \rangle}{\langle a_i, a_i \rangle} a_i = x + \frac{b_i - \sum_{j=1}^n a_{ij} x_j}{\sum_{j=1}^n a_{ij}^2} a_i$$

Exemple 3.19. On considère le système linéaire $Ax = b$ tel que

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

alors

$$\begin{aligned} H_1 &= \{y \in \mathbb{R}^3 : \langle a_1, y \rangle = b_1\} \\ &= \left\{ y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} : y_1 + 2y_2 + y_3 = 1 \right\} \end{aligned}$$

ce qui donne

$$\begin{aligned} P_{H_1}(x_0) &= x_0 + \frac{b_1 - \langle a_1, x_0 \rangle}{\langle a_1, a_1 \rangle} a_1 \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1 - 4}{6} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} =: z_1 \end{aligned}$$

$$\begin{aligned} H_2 &= \{y \in \mathbb{R}^3 : \langle a_2, y \rangle = b_2\} \\ &= \left\{ y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, 4y_1 - y_3 = 2 \right\} \end{aligned}$$

$$\begin{aligned} P_{H_2}(z_1) &= z_1 + \frac{b_2 - \langle a_2, z_1 \rangle}{\langle a_2, a_2 \rangle} a_2 \\ &= \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} + \frac{2 - \frac{3}{2}}{17} \begin{pmatrix} 4 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{21}{34} \\ 0 \\ \frac{8}{17} \end{pmatrix} =: z_2 \end{aligned}$$

$$\begin{aligned} H_3 &= \{y \in \mathbb{R}^3 : \langle a_3, y \rangle = b_3\} \\ &= \left\{ y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, -y_1 + 2y_2 + 2y_3 = 1 \right\} \end{aligned}$$

$$\begin{aligned} P_{H_3}(z_2) &= z_2 + \frac{b_3 - \langle a_3, z_2 \rangle}{\langle a_3, a_3 \rangle} a_3 \\ &= \begin{pmatrix} \frac{21}{34} \\ 0 \\ \frac{8}{17} \end{pmatrix} + \frac{1 - \frac{11}{34}}{9} \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} \end{aligned}$$

donc

$$x_1 = \begin{pmatrix} \frac{83}{153} \\ \frac{23}{153} \\ \frac{95}{153} \end{pmatrix}$$

Supposons à présent que l'on veut résoudre un système d'inéquations linéaires $Ax \leq b$ formé par une matrice $A \in \mathcal{M}_n(\mathbb{R})$ et un vecteur $b = (b_i) \in \mathbb{R}^n$. Ce système vérifie les inégalité

$$\langle a_i, x \rangle \leq b_i, i = \overline{1, n}$$

où a_i représente la $i^{\text{ème}}$ ligne de A . De façon analogue à celle utilisée pour les système linéaires, on trouve que

$$Ax \leq b \Leftrightarrow x \in \bigcap_{i=1}^n \widehat{H}_i$$

où $\widehat{H}_i, i = \overline{1, n}$ sont les demi-espaces définis par

$$\widehat{H}_i := \{y \in \mathbb{R}^n : \langle a_i, y \rangle \leq b_i\}, i = \overline{1, n}$$

Pour chercher **une** solution x , on commence par un point initial $x_0 \in \mathbb{R}^n$, ensuite, on définit la suite des projections alternées comme suit

$$x_m = \left(P_{\widehat{H}_n} P_{\widehat{H}_{n-1}} \cdots P_{\widehat{H}_1} \right)^m (x_0), \quad m \geq 1$$

Si on veut de plus, que ce point x soit le point de l'intersection des \widehat{H}_i le plus proche à x_0 , on utilise l'algorithme Dykstra.

Bibliographie

- [1] D. Azé, *Éléments d'analyse convexe et variationnelle*, Éditions Ellipses, Paris, 1997.
- [2] S. Boyd, J. Dattorro, *Alternating projections*. alternating projection.pdf, 2003. Retrieved 9th April 2012.
- [3] J.P. Boyle, R.L. Dykstra, A method for finding projections onto the intersections of convex sets in Hilbert spaces, In *Advances in Order Restricted Statistical Inference*, Lecture Notes in Statistics, 37, pages 28–47, Springer, Berlin, 1986.
- [4] L.M. Bregman, The method of successive projection for finding a common point of convex sets, *Soviet Mathematics*, 688-692, 1965.
- [5] H. Brézis, *Analyse fonctionnelle*, Éditions Dunod, 2005.
- [6] Y. Censor, A. Gibali, and S. Reich, A von Neumann alternating method for finding common solutions to variational inequalities. *Nonlinear Analysis : Theory, Methods and Applications*, 2012.
- [7] P. Diaconis, K. Khare, L. Saloff-Coste, Stochastic alternating projections. *Illinois Journal of Mathematics*, 54, 963-979, 2010.
- [8] F.Dutsch The angle betwine supspace of a Hilbert space, In *Approximation theory wavelets and applications*, 107-130.
- [9] R.L. Dykstra, An algorithm for restricted least-squares regression, *J. Amer. Statist. Assoc*, 78, 837–842, 1983.
- [10] R. Escalante and M. Raydan, Dykstra's algorithm for a constrained least-squares matrix problem. *Numerical linear algebra with applications*, 3, 459-471, 1998.
- [11] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucia, J. Pelant, V. Zizler, *Functional Analysis and Infinite-Dimensional Geometry*, Springer-Verlag, New York, NY, 2001.
- [12] S.D. Flåm, J. Zowe, Relaxed outer projections, weighted averages and convex feasibility, *BIT*, 30(1990), 289-300.

-
- [13] K. Friedrichs, On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. *Trans. Amer. Math. Soc.*, 41 :321–364, 1937.
- [14] I. Halperin, The product of projection operators, *Acta Sci. Math. (Szeged)*, 23 :96–99,1962.
- [15] J.B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of convex analysis*, Springer. 2004.
- [16] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM Journal on Numerical Analysis*, 964-979, 1979.
- [17] Y. Sonntag, *Topologie et analyse fonctionnelle : cours de licence*, Ellipses 1998.
- [18] N.Z. Shor, *Minimisation Methods for Non-Differentiable Functions*, Springer-Verlag, Berlin, 1985. Springer Ser. Comput. Math. Vol.3.
- [19] J. von Neumann, *Functional Operators Vol. II, The Geometry of Orthogonal Spaces*, 22. Princeton University Press, 1950.