Module B6
Projection sur un convexe

Dans ce module, sauf mention contraire, $\mathcal{X}$ et $\mathcal{Y}$ désignent des espaces de Hilbert, munis chacun d'un produit scalaire noté $\langle \cdot, \cdot \rangle$ et de norme associée notée $\| \cdot \|$, tandis que E désigne un espace euclidien, que l'on identifiera à $\mathbb{R}^n$ avec $n \in \mathbb{N}^*$, muni du produit scalaire usuel noté $\langle \cdot, \cdot \rangle$ également et de norme associée la norme euclidienne, notée $\| \cdot \|_2$.

1 Ensembles convexes

1.1 Définition et exemples

La définition des ensembles convexes repose sur la définition suivante :

Définition 1 (Combinaison convexe)

Soit $(x_1, x_2) \in \mathcal{X}^2$ et $\lambda \in [\,0\,;1\,]$. L'élément

$$\lambda\, x_1 + (1 - \lambda)\, x_2 \in \mathcal{X}$$

est appelé *combinaison convexe* de x_1 et x_2. On note $[\,x_1\,; x_2\,]$ l'ensemble des combinaisons convexes de x_1 et x_2.

Exercice

Soit $x_1 \leq x_2$ deux nombres réels. Vérifier que l'ensemble des combinaisons convexes de x_1 et x_2 est donné par l'intervalle fermé et borné $[\,x_1\,; x_2\,]$.

Notons que la somme des coefficients qui interviennent dans la combinaison convexe de deux points vaut $\lambda + (1 - \lambda) = 1$. On peut maintenant donner la définition suivante :

Définition 2 (Ensemble convexe)

Soit $\mathcal{C} \subset \mathcal{X}$. On dit que $\mathcal{C}$ est *convexe* si

$$\forall\, (x_1, x_2) \in \mathcal{C}^2, \forall\, \lambda \in [\,0\,;1\,], \qquad \lambda\, x_1 + (1 - \lambda)\, x_2 \in \mathcal{C}$$

Autrement dit, un ensemble convexe contient toutes les combinaisons convexes des paires de ses points. On peut simplement dire que $\mathcal{C}$ est *stable par combinaison convexe*.

La combinaison convexe est un cas particulier de la combinaison linéaire, dans laquelle la somme des coefficients scalaires positifs est contrainte à valoir 1. Aussi, tout ensemble stable par combinaison linéaire est stable par combinaison convexe (mais l'inverse est fausse en général !). Voici d'autres exemples d'ensembles convexes :

Exemple

Quelques exemples. Les ensembles suivants sont convexes :

- les (sous-)espaces vectoriels ;
- les singletons ;
- les demi-espaces d'un espace vectoriel ;
- les boules (fermées ou ouvertes).

Sur la droite réelle, les ensembles convexes sont exactement les intervalles (éventuellement réduits à un point) :

Exercice

Convexes dans $\mathbb{R}$. Montrer que les singletons et les intervalles (ouverts, fermés, semi-ouverts, bornés ou non) dans $\mathbb{R}$ sont des ensembles convexes.

Remarque : L'ensemble vide est également convexe.

Exercice

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe. Soit $\{x_i\}_{1\leq i\leq n} \in \mathcal{C}^n$ et $\{\lambda_i\}_{1\leq i\leq n} \in [\,0\,;1\,]^n$ tel que

$$\lambda_1 + \cdots + \lambda_n = 1$$

Montrer que l'élément défini par

$$\lambda_1\, x_1 + \cdots + \lambda_n\, x_n$$

appartient à l'ensemble $\mathcal{C}$. *Indication : on pourra raisonner par récurrence.*

1.2 Propriétés

On va maintenant s'intéresser aux opérations préservant la convexité. Ainsi, au lieu de démontrer la convexité d'un ensemble à partir de la définition, on peut appliquer les résultats qui suivent si l'ensemble considéré est construit à partir d'ensembles convexes connus (comme la boule, les intervalles...).

Proposition 1 (Somme et différence)

Soit $\mathcal{C}_1$ et $\mathcal{C}_2$ deux ensembles convexes de $\mathcal{X}$. Alors

$$\mathcal{C}_1 + \mathcal{C}_2 = \left\{x_1 + x_2 \in \mathcal{X} \mid x_1 \in \mathcal{C}_1 \text{ et } x_2 \in \mathcal{C}_2\right\}$$

et

$$\mathcal{C}_1 - \mathcal{C}_2 = \left\{x_1 - x_2 \in \mathcal{X} \mid x_1 \in \mathcal{C}_1 \text{ et } x_2 \in \mathcal{C}_2\right\}$$

sont des ensembles convexes de $\mathcal{X}$.

Démonstration :

- **Somme.** Soit z et z' deux points de $\mathcal{C}_1 + \mathcal{C}_2$. Par définition de la somme de deux ensembles, il existe donc $(x_1, x_1') \in \mathcal{C}_1^2$ et $(x_2, x_2') \in \mathcal{C}_2^2$ tels que

$$z = x_1 + x_2 \qquad \text{et} \qquad z' = x_1' + x_2'$$

Soit $\lambda \in [\,0\,;1\,]$. Par convexité des ensembles $\mathcal{C}_1$ et $\mathcal{C}_2$, on a $\lambda\, x_1 + (1-\lambda)\, x_1' \in \mathcal{C}_1$ et $\lambda\, x_2 + (1-\lambda)\, x_2' \in \mathcal{C}_2$. Par conséquent, la somme de ces deux points appartient à $\mathcal{C}_1 + \mathcal{C}_2$. Or, puisque

$$(\lambda\, x_1 + (1-\lambda)\, x_1') + (\lambda\, x_2 + (1-\lambda)\, x_2') = \lambda\,(x_1 + x_2) + (1-\lambda)\,(x_1' + x_2')$$

on en déduit que $\lambda\,(x_1 + x_2) + (1-\lambda)\,(x_1' + x_2') \in \mathcal{C}_1 + \mathcal{C}_2$.

- **Différence.** On démontre de même que $\mathcal{C}_1 - \mathcal{C}_2$ est convexe. ■

EXEMPLE

Translation. Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe et $a \in \mathcal{X}$. L'ensemble

$$\mathcal{C} + a = \{x + a \mid x \in \mathcal{C}\}$$

est convexe.

Proposition 2 (Intersection)

Soit $\mathcal{I} \subset \mathbb{R}$ et $\mathcal{C}_i$ un ensemble convexe de $\mathcal{X}$ pour tout $i \in \mathcal{I}$. Alors

$$\bigcap_{i\in\mathcal{I}} \mathcal{C}_i = \{x \in \mathcal{X} \mid \forall\, i \in \mathcal{I}, x \in \mathcal{C}_i\}$$

est un ensemble convexe de $\mathcal{X}$.

DÉMONSTRATION : Laissée au lecteur.

EXEMPLE

Rectangle. Soit $a, b, c, d \in \mathcal{X}$. L'ensemble

$$[\,a\,;b\,] \times [\,c\,;d\,] = \left\{(t\,a + (1-t)\,b, u\,c + (1-u)) \mid (t, u) \in [\,0\,;1\,]^2\right\}$$

est convexe. On en déduit que tout rectangle du plan est convexe.

EXERCICE

L'union finie d'ensembles convexes n'est pas convexe. Soit $(a_1, a_2) \in \mathcal{X}^2$ tels que $a_1 \neq a_2$. On pose $\mathcal{C}_1 = \{a_1\}$ et $\mathcal{C}_2 = \{a_2\}$. Les ensembles $\mathcal{C}_1$ et $\mathcal{C}_2$ sont convexes. Vérifier que leur union, c'est-à-dire l'ensemble

$$\mathcal{C}_1 \cup \mathcal{C}_2 = \{x \in \mathcal{X} \mid x \in \mathcal{C}_1 \text{ ou } x \in \mathcal{C}_2\}$$

n'est pas convexe.

Proposition 3 (Image directe)

Soit $\mathcal{C}$ un ensemble convexe de $\mathcal{X}$ et $A : \mathcal{X} \to \mathcal{Y}$ une application linéaire. Alors

$$A(\mathcal{C}) = \left\{A\,x \in \mathcal{Y} \mid x \in \mathcal{C}\right\}$$

est un ensemble convexe de $\mathcal{Y}$.

DÉMONSTRATION : Soit y_1 et y_2 deux points de $A(\mathcal{C})$. Par définition de l'image directe, il existe donc $(x_1, x_2) \in \mathcal{C}^2$ tel que

$$y_1 = A\,x_1 \qquad \text{et} \qquad y_2 = A\,x_2$$

Soit $\lambda \in [\,0\,;1\,]$. Puisque $\mathcal{C}$ est convexe, on a $\lambda\,x_1 + (1-\lambda)\,x_2 \in \mathcal{C}$. L'application A est linéaire, donc $A\big(\lambda\,x_1 + (1-\lambda)\,x_2\big) = \lambda\,A\,x_1 + (1-\lambda)\,A\,x_2$, ce qui implique

que $\lambda\, y_1 + (1-\lambda)\, y_2 \in A(\mathcal{C})$. ■

Exemple

Homothétie et rotation. Toute homothétie et rotation du plan ou de l'espace préserve la convexité.

Proposition 4 (Image réciproque)

Soit $\mathcal{C}$ un ensemble convexe de $\mathcal{Y}$ et $A : \mathcal{X} \to \mathcal{Y}$ une application linéaire. Alors

$$A^{-1}(\mathcal{C}) = \Big\{ x \in \mathcal{X} \mid A\,x \in \mathcal{C} \Big\}$$

est un ensemble convexe de $\mathcal{X}$.

Démonstration : Soit x_1 et x_2 deux points de $A^{-1}(\mathcal{C})$. Par définition de l'image réciproque, on a

$$(A\,x_1, A\,x_2) \in \mathcal{C}^2$$

Soit $\lambda \in [\,0\,;1\,]$. Par convexité de $\mathcal{C}$, le point $\lambda\, A\,x_1 + (1-\lambda)\, A\,x_2$ appartient à $\mathcal{C}$. L'application A est linéaire donc $\lambda\, A\,x_1 + (1-\lambda)\, A\,x_2 = A\big(\lambda\, x_1 + (1-\lambda)\, x_2\big)$ Il s'ensuit que $\lambda\, x_1 + (1-\lambda)\, x_2 \in A^{-1}(\mathcal{C})$. ■

Exemple

Soit $G : \mathcal{X} \to \mathbb{R}$ une application affine, c'est-à-dire une application de la forme ; il existe donc $(a, b) \in \mathcal{X} \times \mathbb{R}$ tel que

$$\forall\, x \in \mathcal{X}, \qquad G(x) = L(x) + b$$

où $L : \mathcal{X} \to \mathbb{R}$ est linéaire et $b \in \mathcal{X}$ un vecteur. Alors les ensembles suivants

$$\Big\{ x \in \mathcal{X} \mid G(x) = 0 \Big\} \qquad \text{et} \qquad \Big\{ x \in \mathcal{X} \mid G(x) \leq 0 \Big\}$$

sont convexes, car ils peuvent être vus comme les images réciproques par la forme linéaire L du singleton $\{-b\}$ et de l'intervalle fermé $]\,-\infty\,;-b\,]$, qui sont tous les deux des ensembles convexes de $\mathbb{R}$.

2 Projection sur un convexe

2.1 Définition et existence

Une première application d'optimisation convexe importante est celle du problème de la projection orthogonale sur un convexe fermé. Dans ce problème, on considère un convexe $\mathcal{C}$ et un point x_0 en-dehors de ce convexe. On s'intéresse alors à l'existence, puis à l'unicité, d'un point dans le convexe minimisant la distance avec le point initial x_0.

Définition 3 (Projection sur un convexe)

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe et $x_0 \in \mathcal{X}$. On appelle *projection* de x_0 sur $\mathcal{C}$ tout point noté $\operatorname{proj}_{\mathcal{C}}(x_0)$ de $\mathcal{C}$ défini (s'il existe) par

$$\forall\, x \in \mathcal{C}, \qquad \|\operatorname{proj}_{\mathcal{C}}(x_0) - x_0\| \leq \|x - x_0\|$$

REMARQUE : Les points $\mathrm{proj}_{\mathcal{C}}(x_0)$ sont les solutions (si elles existent) du problème

$$\text{Minimiser } \|x - x_0\| \text{ sous les contraintes } x \in \mathcal{C}$$

On peut écrire ce problème de manière équivalente

$$\text{Minimiser } \frac{1}{2}\|x - x_0\|^2 \text{ sous les contraintes } x \in \mathcal{C} \qquad (\mathcal{P}_{\mathrm{proj}})$$

L'intérêt d'une telle réécriture est d'obtenir un problème dont la fonction objectif est différentiable, contrairement au problème introduit dans la définition 3 ; le facteur multiplicatif 1/2 permet d'éviter le facteur multiplicatif pour le gradient de la fonction objectif, qui vaut

$$x \mapsto x - x_0$$

Théorème 1

Soit $\mathcal{X}$ un espace de HILBERT. Soit $\mathcal{C} \subset \mathcal{X}$. Si $\mathcal{C}$ est convexe, non vide et fermé, alors la projection de tout point x_0 sur $\mathcal{C}$ existe et est unique.

REMARQUE : Si $\mathcal{X}$ est un espace euclidien, alors l'existence a déjà été démontrée dans le module **B4 : Optimisation sous contraintes**.

DÉMONSTRATION :

- **Existence.** Soit $(x_k)_{k\in\mathbb{N}^*}$ une suite minimisante d'éléments du problème de projection, c'est-à-dire vérifiant $x_k \in \mathcal{C}$ pour tout $k \in \mathbb{N}^*$ et

$$\lim_{k\to+\infty} \frac{1}{2}\|x_0 - x_k\|^2 = \inf_{x\in\mathcal{C}} \frac{1}{2}\|x_0 - x\|^2$$

Notons $\alpha \in \mathbb{R}^+$ cette quantité. On rappelle l'identité du parallélogramme

$$\forall\,(a,b) \in (\mathbb{R}^n)^2 \qquad \|a+b\|^2 + \|a-b\|^2 = 2\left(\|a\|^2 + \|b\|^2\right)$$

On l'applique à $a = x_0 - x_k$ et $b = x_0 - x_j$ (avec $(j,k) \in (\mathbb{N}^*)^2$) :

$$4\left\|x_0 - \frac{x_j + x_k}{2}\right\|^2 + \|x_j - x_k\|^2 = 2\left(\|x_0 - x_j\|^2 + \|x_0 - x_k\|^2\right)$$

L'ensemble $\mathcal{C}$ est convexe et $(x_j, x_k) \in \mathcal{C}^2$ donc $(x_j + x_k)/2 \in \mathcal{C}$. En particulier, par définition de α, on a

$$\frac{1}{2}\left\|x_0 - \frac{x_j + x_k}{2}\right\|^2 \geq \alpha$$

Ainsi $$0 \leq \|x_j - x_k\|^2 \leq 2\left(\|x_0 - x_j\|^2 + \|x_0 - x_k\|^2\right) - 8\,\alpha$$

En faisant tendre j et k vers $+\infty$ dans cette relation, le terme de droite convergeant par définition de la suite minimisante vers 0, on en déduit par encadrement que la suite $(x_k)_{k\in\mathbb{N}^*}$ est une suite de CAUCHY. Puisque $U = \mathbb{R}^n$ est complet, elle est convergente. Notons $\mathrm{proj}_{\mathcal{C}}(x_0)$ la limite de $(x_k)_{k\in\mathbb{N}}$; on a que $\mathrm{proj}_{\mathcal{C}}(x_0) \in \mathcal{C}$ car $\mathcal{C}$ est fermé par hypothèse. Par ailleurs, la norme est continue, donc $x \mapsto \|x\|^2/2$ aussi (par composition) et

$$\lim_{k\to+\infty} \frac{1}{2}\|x_0 - x_k\|^2 = \frac{1}{2}\|x_0 - \mathrm{proj}_{\mathcal{C}}(x_0)\|^2 = \alpha$$

Par définition de la borne inférieure, on a donc montré que

$$\forall\, x \in \mathcal{C}, \qquad \frac{1}{2}\|x_0 - \mathrm{proj}_{\mathcal{C}}(x_0)\|^2 \leq \frac{1}{2}\|x_0 - x\|^2$$

Autrement dit, $\mathrm{proj}_{\mathcal{C}}(x_0)$ est une projection de x_0 sur $\mathcal{C}$.

- **Unicité.** Il s'agit de la conséquence de la stricte convexité de la fonction

$$x \mapsto \frac{1}{2}\|x - x_0\|^2$$

et de la proposition 2 du module **B4 : Optimisation sous contraintes**. ■

2.2 Propriétés

Puisque la projection est définie de manière unique lorsque le convexe est non vide et fermé, on peut la considérer comme une application.

Proposition 5

Soit $\mathcal{X}$ un espace de Hilbert. Soit $\mathcal{C} \subset \mathcal{X}$ un convexe non vide et fermé. Alors l'application $\mathrm{proj}_{\mathcal{C}} : \mathcal{X} \to \mathcal{X}$ vérifie l'inégalité

$$\forall\,(x_1, x_2) \in \mathcal{X}^2, \qquad \|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\| \le \|x_1 - x_2\|$$

Autrement dit, $\mathrm{proj}_{\mathcal{C}}$ est *lipschitzienne* de constante 1.

Démonstration : Soit $(x_1, x_2) \in \mathcal{X}^2$. On commence par écrire

$$\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2) = \mathrm{proj}_{\mathcal{C}}(x_1) - x_1 + x_1 - x_2 + x_2 - \mathrm{proj}_{\mathcal{C}}(x_2)$$

Ensuite, puisque $\|a\|^2 = \langle a, a\rangle$, on obtient le développement suivant :

$$\begin{aligned}\|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\|^2 = {} & \langle \mathrm{proj}_{\mathcal{C}}(x_1) - x_1, \mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\rangle \\ & + \langle x_1 - x_2, \mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\rangle \\ & + \langle x_2 - \mathrm{proj}_{\mathcal{C}}(x_2), \mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\rangle\end{aligned}$$

La proposition 6 (qui se démontre de manière indépendante) assure que

$$\langle x_1 - \mathrm{proj}_{\mathcal{C}}(x_1), x_2 - \mathrm{proj}_{\mathcal{C}}(x_1)\rangle \le 0 \qquad \text{et} \qquad \langle x_2 - \mathrm{proj}_{\mathcal{C}}(x_2), x_1 - \mathrm{proj}_{\mathcal{C}}(x_2)\rangle \le 0$$

ce qui implique que

$$\|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\|^2 \le \langle x_1 - x_2, \mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\rangle$$

On applique alors l'inégalité de Cauchy-Schwarz, ce qui donne

$$\|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\|^2 \le \|x_1 - x_2\|\,\|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\|$$

Si $\mathrm{proj}_{\mathcal{C}}(x_1) \neq \mathrm{proj}_{\mathcal{C}}(x_2)$, alors on peut diviser par $\|\mathrm{proj}_{\mathcal{C}}(x_1) - \mathrm{proj}_{\mathcal{C}}(x_2)\| > 0$; sinon, alors la relation à démontrer devient

$$0 \le \|x_1 - x_2\|$$

ce qui est évidemment vrai. ■

En particulier, on a démontré le résultat suivant :

Corollaire 1

Soit $\mathcal{X}$ un espace de Hilbert. Soit $\mathcal{C} \subset \mathcal{X}$ un convexe non vide et fermé. Alors l'application $\mathrm{proj}_{\mathcal{C}} : \mathcal{X} \to \mathcal{X}$ est continue.

2.3 Exemples

Dans ce paragraphe, on va donner quelques exemples élémentaires de projection sur un convexe fermé non vide.

Exemple

Projection sur un segment de $\mathbb{R}$. Soit $a > b$ deux réels. Soit $x_0 \in \mathbb{R}$. On s'intéresse à la projection de x_0 sur le segment

$$\mathcal{C} = [\,a\,;b\,]$$

c'est-à-dire au point $x^* \in \mathcal{C}$ tel que

$$\forall\, x \in [\,a\,;b\,], \qquad |x^* - x_0| \leq |x - x_0|$$

Ce point existe et est unique car le segment est convexe, fermé et non vide (il contient a). Le terme de gauche est minoré par 0 ; par ailleurs, il s'annule si et seulement si $x^* = x_0$. On en déduit que, si $x_0 \in [\,a\,;b\,]$, alors $x^* = x_0$. On suppose donc que $x_0 \notin [\,a\,;b\,]$. Autrement dit, on a soit $x_0 < a$, soit $x_0 > b$.

- **Si $x_0 < a$:** dans ce cas, pour tout $x \in [\,a\,;b\,]$, on a

  $$x_0 < a \leq x \leq b \qquad \text{soit} \qquad 0 < a - x_0 \leq x - x_0 \leq b - x_0$$

  En particulier, on a

  $$\forall\, x \in [\,a\,;b\,], \qquad |a - x_0| \leq |x - x_0|$$

  de sorte que $x^* = a$.
- **Si $x_0 > b$:** dans ce cas, pour tout $x \in [\,a\,;b\,]$, on a

  $$a \leq x \leq b < x_0 \qquad \text{soit} \qquad a - x_0 \leq x - x_0 \leq b - x_0 < 0$$

  En particulier, on a

  $$\forall\, x \in [\,a\,;b\,], \qquad |b - x_0| \leq |x - x_0|$$

  de sorte que $x^* = b$.

Finalement,
$$\operatorname{proj}_{[\,a\,;b\,]}(x_0) = \begin{cases} x_0 & \text{si } a \leq x_0 \leq b \\ a & \text{si } a > x_0 \\ b & \text{si } x_0 > b \end{cases}$$

Le caractère fermé de l'ensemble sur lequel on projette est essentiel :

Contre-exemple

Projection sur un ouvert. Soit $a > b$ deux réels. Soit $x_0 \in \mathbb{R}$. On s'intéresse à la projection de x_0 sur l'intervalle ouvert

$$\mathcal{C} =]\,a\,;b\,[$$

Si $x_0 < a$, alors on a pour tout $x\,]\,a\,;b\,[$,

$$|a - x_0| < |x - x_0| = x - x_0$$

Notons pour commencer que $\mathcal{C}$ n'est pas fermé (mais il est convexe et non vide). On suppose que la projection de x_0 sur $\mathcal{C}$ existe, et notons-la x^*. Par définition, on a $a < x^* < b$ et

$$\forall\, x \in]\,a\,;b\,[, \qquad |x^* - x_0| \leq |x - x_0|$$

On pose $x = (x^* - a)/2$. On a $x \in \,]\,a\,;x^*\,[\, \subset \,]\,a\,;b\,[$ et

$$|x^* - x_0| = x^* - x_0 = x^* - x + x - x_0 = |x^* - x| + |x - x_0| \geq |x - x_0|$$

ce qui est absurde. On en déduit que la projection de x_0 sur $\mathcal{C}$ n'existe pas.

EXEMPLE

Projection sur un disque. Soit $C \in \mathbb{R}^2$ et $R > 0$. Soit $x_0 \in \mathbb{R}^2$. On s'intéresse à la projection de x_0 sur le disque

$$\mathcal{C} = \overline{\mathcal{B}}(C,R) = \left\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 \leq R \right\}$$

c'est-à-dire au point $x^* \in \mathcal{C}$ tel que

$$\forall\, x \in \overline{\mathcal{B}}(C,R), \qquad \|x^* - x_0\|_2^2 \leq \|x - x_0\|_2^2$$

Ce point existe et est unique car le disque est convexe, fermé et non vide (il contient C). On remarque que, si $x_0 \in \overline{\mathcal{B}}(C,R)$, alors $x^* = x_0$. On suppose donc maintenant que $x_0 \notin \overline{\mathcal{B}}(C,R)$. Autrement dit, on a $\|x_0 - C\|_2 > R$. Soit $x \in \overline{\mathcal{B}}(C,R)$. L'inégalité de CAUCHY–SCHWARZ assure que

$$\begin{aligned}
\|x - x_0\|_2^2 &= \|x - C - (x_0 - C)\|_2^2 \\
&= \|x - C\|_2^2 - 2\,\langle x - C, x_0 - C\rangle + \|x_0 - C\|_2^2 \\
&\geq \|x - C\|_2^2 - 2\,\|x - C\|_2\,\|x_0 - C\|_2 + \|x_0 - C\|_2^2 \\
\|x - x_0\|_2^2 &\geq (\|x - C\|_2 - \|x_0 - C\|_2)^2
\end{aligned}$$

Or, le terme de droite est atteint sur le disque : en effet, si on pose

$$x^* = C + R\frac{x_0 - C}{\|x_0 - C\|_2}$$

on a $\|x^* - C\|_2 = R$ et

$$\begin{aligned}
\left\| C + R\frac{x_0 - C}{\|x_0 - C\|_2} - x_0 \right\|_2^2 &= \left\| R\frac{x_0 - C}{\|x_0 - C\|_2} - (x_0 - C) \right\|_2^2 \\
&= \left(\frac{R - \|x_0 - C\|_2}{\|x_0 - C\|_2} \right)^2 \|x_0 - C\|_2^2 \\
\left\| C + R\frac{x_0 - C}{\|x_0 - C\|_2} - x_0 \right\|_2^2 &= (R - \|x_0 - C\|_2)^2
\end{aligned}$$

On en déduit que

$$\mathrm{proj}_{\overline{\mathcal{B}}(C,R)}(x_0) = \begin{cases} x_0 & \text{si } \|x_0 - C\|_2 \leq R \\ C + R\dfrac{x_0 - C}{\|x_0 - C\|_2} & \text{si } \|x_0 - C\|_2 > R \end{cases}$$

Lorsque la projection considérée se fait sur un ensemble défini par des contraintes d'égalité et / ou d'inégalité, on peut envisager d'appliquer le théorème de KARUSH-KUHN-TUCKER vu dans le module **B5 : Théorème de KARUSH–KUHN–TUCKER** :

EXEMPLE

Projection sur un disque. Soit $x_0 \in \mathbb{R}^2$. La projection de x_0 sur le disque

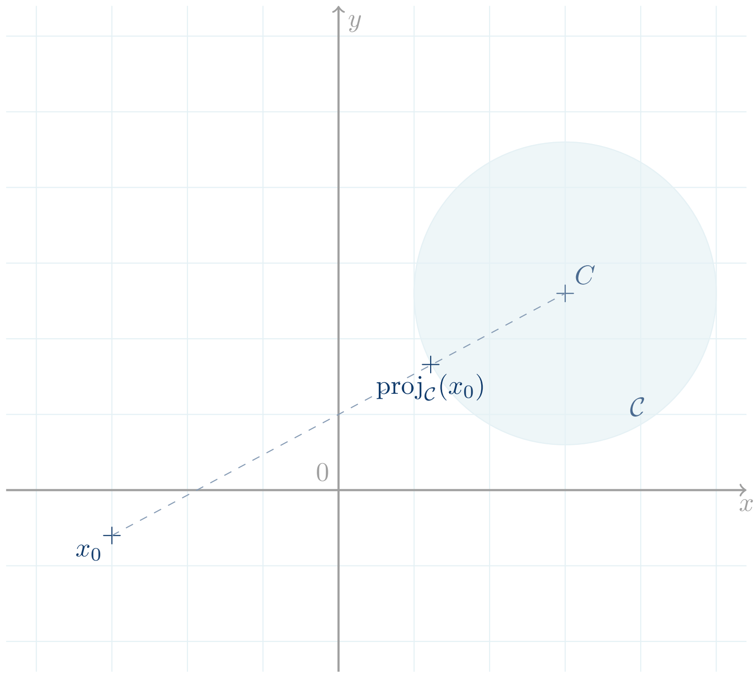

FIGURE 1 – Projection sur un disque fermé non vide.

$\overline{\mathcal{B}}(C, R)$ s'écrit comme le problème d'optimisation sous contraintes d'inégalité :

$$\text{Minimiser } \frac{1}{2}\,\|x - x_0\|_2^2 \text{ sous les contraintes } \|x - C\|_2^2 \le R^2$$

Le lagrangien associé est donné par

$$\forall\,(x, \mu) \in \mathbb{R}^2 \times \mathbb{R}^+, \qquad \mathcal{L}(x; \mu) = \frac{1}{2}\,\|x - x_0\|_2^2 + \mu\,(\|x - C\|_2^2 - R^2)$$

La contrainte est qualifiée et les conditions KKT s'écrivent

$$\begin{cases} \|x - C\|_2^2 - R^2 \le 0 & \text{(admissibilité)} \\ x - x_0 + \mu\,(x - C) = 0 & \text{(premier ordre)} \\ \mu\,(\|x - C\|_2^2 - R^2) = 0 & \text{(complémentarité)} \\ \mu \ge 0 \end{cases}$$

La condition du premier ordre assure que

$$x = \frac{x_0 + \mu\,C}{1 + \mu}$$

Si $\mu = 0$, alors $x = x_0$; ce point est admissible si $x_0 \in \mathcal{C}$. Si $x_0 \notin \mathcal{C}$, alors $\mu > 0$ et la condition de complémentarité assure que

$$R^2 = \left\| \frac{x_0 + \mu\,C}{1 + \mu} - C \right\|_2^2 = \left\| \frac{x_0 - C}{1 + \mu} \right\|_2^2 = \frac{\|x_0 - C\|_2^2}{(1 + \mu)^2}$$

On en déduit que $$\mu = \frac{\|x_0 - C\|_2}{R} - 1$$

qui est bien strictement positif si $x_0 \notin \mathcal{C}$ et la condition du premier ordre devient

$$x = \frac{R\,x_0 + \|x_0 - C\|_2\,C - R\,C}{\|x_0 - C\|_2} = C + R\,\frac{x_0 - C}{\|x_0 - C\|_2}$$

La convexité de l'ensemble sur lequel on projette est importante pour garantir l'unicité du projeté :

CONTRE-EXEMPLE

Projection sur un ensemble non convexe. Soit $C \in \mathbb{R}^2$ et $R > 0$. Soit $x_0 \in \mathbb{R}^2$. On s'intéresse à la projection de x_0 sur le cercle

$$\mathcal{C} = \left\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 = R \right\}$$

Notons que le cercle n'est pas convexe (mais il est fermé et non vide). Si $x_0 = C$, alors pour tout $x \in \mathcal{C}$, on a $\|x - C\|_2 = R$, de sorte que tous les points de $\mathcal{C}$ minimisent la distance de C au cercle. Ainsi, la projection n'est pas unique.

3 Théorèmes de séparation

3.1 Inéquation variationnelle pour la projection

Montrons que la projection sur un convexe fermé est caractérisée par une inégalité :

Proposition 6

Soit $\mathcal{X}$ un espace de HILBERT. Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe, non vide et fermé et $x_0 \in \mathcal{X}$. Alors la projection de x_0 sur $\mathcal{C}$ est l'unique point de $\mathcal{C}$ qui vérifie

$$\forall\, x \in \mathcal{C}, \qquad \langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - \mathrm{proj}_{\mathcal{C}}(x_0)\rangle \leq 0$$

DÉMONSTRATION : Il s'agit d'une conséquence de la proposition 7 du module **B4 : Optimisation sous contraintes**, qui assure que les solutions du problème convexe

$$\text{Minimiser } f(x) = \frac{1}{2}\,\|x - x_0\|^2 \text{ sous les contraintes } x \in \mathcal{C}$$

sont les solutions du problème d'inéquation variationnelle

$$\text{Trouver } x^* \in \mathcal{C} \qquad \text{tel que} \qquad \forall\, x \in \mathcal{C}, \quad \langle \nabla f(x^*), \mathrm{proj}_{\mathcal{C}}(x_0) - x\rangle \geq 0$$

(car f est différentiable convexe et $\mathcal{C}$ est convexe). On conclut en remarquant que $x^* = \mathrm{proj}_{\mathcal{C}}(x_0)$ et que $\nabla f(x^*) = x^* - x_0$. ■

3.2 Séparation d'un point et d'un convexe fermé

La proposition 6 va nous permettre de démontrer deux théorèmes de séparation, qui stipulent l'existence d'un hyperplan affine séparant un point et un ensemble convexe, c'est-à-dire partageant l'espace $\mathcal{X}$ en deux demi-espaces ouverts contenant l'un le convexe donné et l'autre le point considéré (qui par hypothèse n'appartient pas à l'ensemble convexe). On commence par le cas d'un convexe fermé :

Théorème 2

Soit $\mathcal{X}$ un espace de HILBERT. Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe fermé et non vide, et $x_0 \notin \mathcal{C}$. Alors il existe $a, b \in \mathcal{X} \times \mathbb{R}$ tel que

$$\langle a, x_0 \rangle > b \qquad \text{et} \qquad \forall\, x \in \mathcal{C}, \qquad \langle a, x \rangle < b$$

L'ensemble
$$\mathcal{H} = \Big\{ x \in \mathcal{X} \mid \langle a, x \rangle = b \Big\}$$
définit un hyperplan affine. On dit alors que $\mathcal{H}$ *sépare strictement* le point x_0 et l'ensemble C, car x_0 et $\mathcal{C}$ sont chacun contenus dans deux demi-espaces strictement différents (définis par $\mathcal{H}$).

DÉMONSTRATION : On définit
$$f : \left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathbb{R} \\ x & \mapsto & \langle x_0 - \operatorname{proj}_{\mathcal{C}}(x_0), x \rangle \end{array} \right.$$
La fonction f est linéaire et continue. Remarquons que, d'après la proposition 6, on a d'une part
$$\forall\, x \in \mathcal{C}, \qquad f(x) \leq f\big(\operatorname{proj}_{\mathcal{C}}(x_0)\big)$$
D'autre part, on a
$$f(x_0) = \langle x_0 - \operatorname{proj}_{\mathcal{C}}(x_0), x_0 \rangle = \|x_0 - \operatorname{proj}_{\mathcal{C}}(x_0)\|^2 + f\big(\operatorname{proj}_{\mathcal{C}}(x_0)\big)$$
Puisque $x_0 \notin \mathcal{C}$ et que $\operatorname{proj}_{\mathcal{C}}(x_0) \in C$ par définition de la projection, il s'ensuit que $x_0 - \operatorname{proj}_{\mathcal{C}}(x_0) \neq 0$ et que
$$f(x_0) > f\big(\operatorname{proj}_{\mathcal{C}}(x_0)\big)$$
On a donc démontré que
$$\forall\, x \in \mathcal{C}, \qquad f(x_0) > f\big(\operatorname{proj}_{\mathcal{C}}(x_0)\big) \geq f(x)$$
Considérons le point $y = (x_0 + \operatorname{proj}_{\mathcal{C}}(x_0))/2$. On a $y \neq x_0$ et $y \neq \operatorname{proj}_{\mathcal{C}}(x_0)$. Par ailleurs, par linéarité de f,
$$f(y) = \frac{1}{2}\left(f(x_0) + f(\operatorname{proj}_{\mathcal{C}}(x_0)\right) \qquad \text{et} \qquad f(x_0) > f(y) > f\big(\operatorname{proj}_{\mathcal{C}}(x_0)\big)$$
Il s'ensuit que $\qquad \forall\, x \in \mathcal{C}, \qquad \underbrace{f(y)}_{=\,b} > f(x) = \langle \underbrace{x_0 - \operatorname{proj}_{\mathcal{C}}(x_0)}_{=\,a}, x \rangle$

tandis que $\qquad f(x_0) = \langle x_0 - \operatorname{proj}_{\mathcal{C}}(x_0), x_0 \rangle = \langle a, x_0 \rangle > f(y) = b$ ■

REMARQUE : Notons que, avec les notations de la preuve ci-dessus,
$$b = \langle a, y \rangle = \frac{1}{2}\,\langle x_0 - \operatorname{proj}_{\mathcal{C}}(x_0), x_0 + \operatorname{proj}_{\mathcal{C}}(x_0) \rangle = \frac{1}{2}\left(\|x_0\|^2 - \|\operatorname{proj}_{\mathcal{C}}(x_0)\|^2\right)$$
si bien que l'hyperplan séparateur déterminé à la proposition précédente est l'ensemble des points $x \in \mathcal{X}$ tels que
$$\langle x_0 - \operatorname{proj}_{\mathcal{C}}(x_0), x \rangle = \frac{1}{2}\left(\|x_0\|^2 - \|\operatorname{proj}_{\mathcal{C}}(x_0)\|^2\right)$$
On remarquera par ailleurs que le point $y = (x_0 + \operatorname{proj}_{\mathcal{C}}(x_0))/2$ appartient à l'hyperplan séparateur.

EXEMPLE

Droite séparant un point et un disque du plan. Soit $C \in \mathbb{R}^2$ et $R > 0$. On considère le disque fermé non vide

$$\mathcal{C} = \overline{\mathcal{B}}(C,R) = \left\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 \leq R \right\}$$

Soit $x_0) \in \mathbb{R}^2 \setminus \mathcal{C}$. D'après la remarque précédente, la droite d'équation

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - \frac{1}{2}\,(x_0 + \mathrm{proj}_{\mathcal{C}}(x_0)) \right\rangle = 0$$

sépare le disque $\mathcal{C}$ du point x_0. On voit qu'il s'agit de l'unique droite passant par le milieu du segment $[\,x_0\,;\mathrm{proj}_{\mathcal{C}}(x_0)\,]$ et perpendiculaire à la droite reliant les points x_0 et $\mathrm{proj}_{\mathcal{C}}(x_0)$.

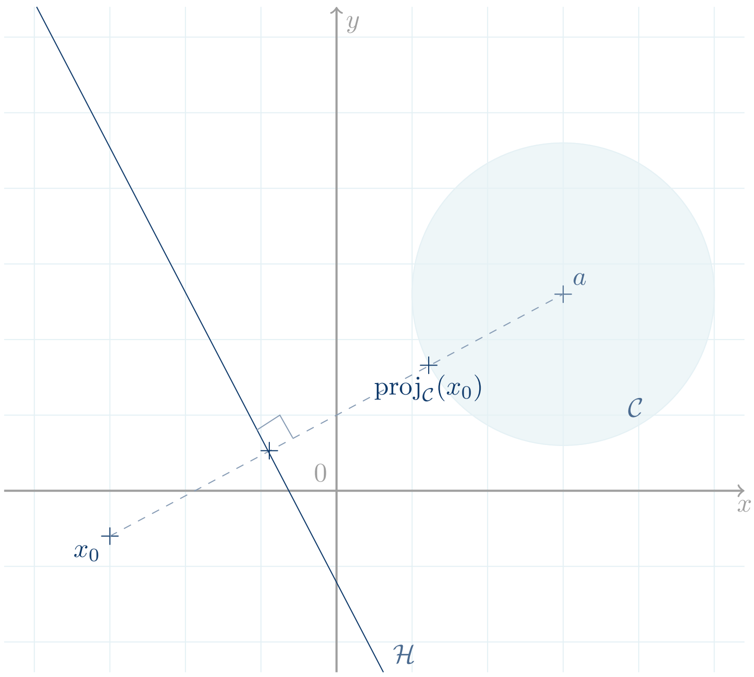

FIGURE 2 – Séparation d'un point et d'un disque fermé non vide.

REMARQUE : La figure 2 suggère qu'il peut exister une infinité d'hyperplans séparateurs.

3.3 Théorème de HAHN–BANACH

On s'intéresse maintenant au cas d'un ensemble convexe ouvert $\mathcal{C}$. Si le point x_0 n'appartient pas à l'adhérence de $\mathcal{C}$, alors on peut appliquer la proposition précédente à $\mathrm{Adh}(\mathcal{C})$. L'hyperplan séparateur donné dans cette proposition définit donc $a \in \mathcal{X}$ et $b \in \mathbb{R}$ tel que

$$\langle a, x_0 \rangle > b \qquad \text{et} \qquad \forall\, x \in \mathrm{Adh}(\mathcal{C}), \quad \langle a, x \rangle < b$$

et on a en particulier $\qquad \forall\, x \in \mathcal{C}, \qquad \langle a, x \rangle < b$

puisque $\mathcal{C} \subset \mathrm{Adh}(\mathcal{C})$. La question de l'existence d'un hyperplan strictement séparateur se pose donc lorsque $x_0 \in \mathrm{Adh}(\mathcal{C})$. Il est clair qu'il n'est pas possible de séparer x_0

et $\mathcal{C}$ de manière stricte à l'aide d'un hyperplan. En effet, si c'était possible, alors il existerait $a \in \mathcal{X}$ et $b \in \mathbb{R}$ tel que

$$\langle a, x_0 \rangle > b \qquad \text{et} \qquad \forall\, x \in \mathcal{C}, \quad \langle a, x \rangle < b$$

Puisque $x_0 \in \mathrm{Adh}(\mathcal{C})$, il existe une suite $(x_k)_{k\in\mathbb{N}^*}$ tel que $x_k \in \mathcal{C}$ pour tout $k \in \mathbb{N}^*$ et

$$\lim_{k\to+\infty} x_k = x_0$$

La continuité du produit scalaire assure alors par passage à la limite que

$$b > \lim_{k\to+\infty} \langle a, x_k \rangle = \langle a, x_0 \rangle > b$$

ce qui constitue une contradiction. On va cependant montrer qu'il est possible de séparer de manière non stricte tout convexe ouvert et un point n'appartenant pas à ce convexe. On va admettre le résultat suivant :

Proposition 7

Soit $\mathcal{X}$ un espace de Hilbert. Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble convexe ouvert et non vide, et $x_0 \notin \mathcal{C}$. Alors il existe $a, b \in \mathcal{X} \times \mathbb{R}$ tel que

$$\langle a, x_0 \rangle = b \qquad \text{et} \qquad \forall\, x \in \mathcal{C}, \qquad \langle a, x \rangle < b$$

Démonstration : Admis.

Autrement dit, il existe un hyperplan affine

$$\mathcal{H} = \Big\{ x \in \mathcal{X} \mid \langle a, x \rangle = b \Big\}$$

passant par le point x_0 et tel que l'ensemble C est contenu dans un des deux demi-espaces ouverts définis par $\mathcal{H}$.

Exemple

Droite passant par un point donné et définissant un demi-plan contenant strictement un disque du plan. Soit $C \in \mathbb{R}^2$ et $R > 0$. On considère le disque ouvert non vide

$$\mathcal{C} = \overline{\mathcal{B}}(C, R) = \Big\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 < R \Big\}$$

d'adhérence
$$\mathrm{Adh}(\mathcal{C}) = \Big\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 = R \Big\}$$

Soit $x_0 \in \mathbb{R}^2 \setminus \mathcal{C}$. Si $x_0 \notin \mathrm{Adh}(\mathcal{C})$, alors on a montré dans l'exemple précédent que la droite d'équation

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - \frac{1}{2}\left(x_0 + \mathrm{proj}_{\mathcal{C}}(x_0)\right) \right\rangle = 0$$

sépare le disque $\mathcal{C}$ du point x_0. Considérons la droite parallèle à cette droite et passant par x_0

$$\langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - x_0 \rangle = 0$$

Par construction, cette droite, que nous noterons $\mathcal{H}$, passe par le point x_0. Par ailleurs, pour tout $x \in \mathbb{R}^2$ tel que

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - \frac{1}{2}\left(x_0 + \mathrm{proj}_{\mathcal{C}}(x_0)\right) \right\rangle < 0$$

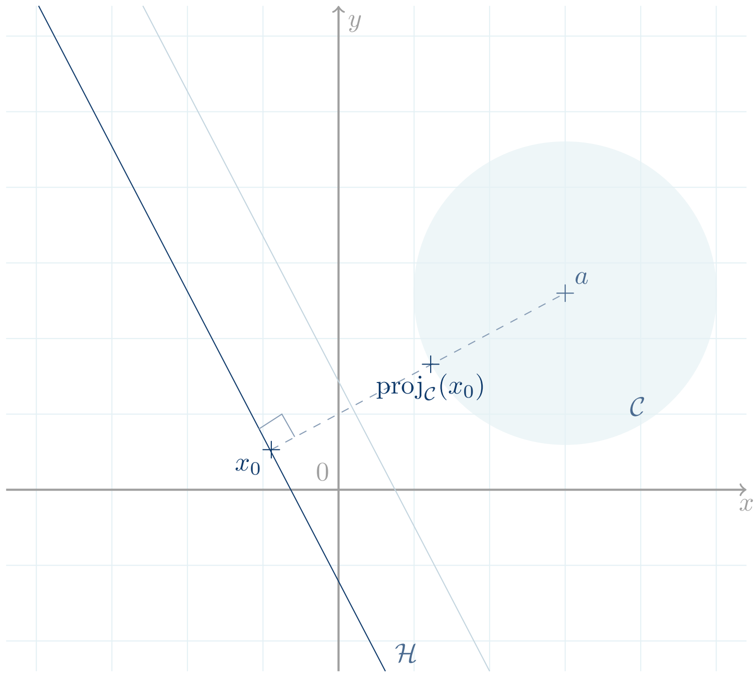

FIGURE 3 – Droite passant par un point donné et définissant un demi-plan contenant strictement un disque du plan. Cas où $x_0 \notin \mathrm{Adh}(\mathcal{C})$.

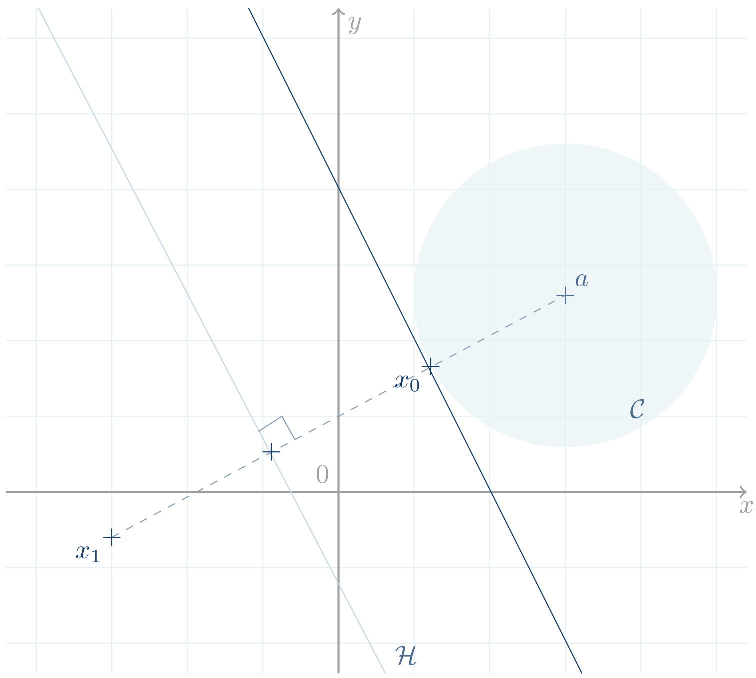

FIGURE 4 – Droite passant par un point donné et définissant un demi-plan contenant strictement un disque du plan. Cas où $x_0 \in \mathrm{Adh}(\mathcal{C})$.

on a $$\langle x_0 - \mathrm{proj}_{\mathcal{C}}(x_0), x - x_0\rangle < -\frac{1}{2}\,\|x_0 - \mathrm{proj}_{\mathcal{C}}(x_0)\|_2^2 < 0$$

de sorte que $\mathcal{C}$ est contenu strictement dans un demi-plan ouvert délimité par $\mathcal{H}$. Supposons maintenant $x_0 \in \mathrm{Adh}(\mathcal{C})$, c'est-à-dire que $\|x_0 - C\|_2 = R$. Soit $k \in \mathbb{N}^*$. Considérons le disque fermé non vide

$$\mathcal{C}_k = \overline{\mathcal{B}}(C, R - 1/k) = \left\{ x \in \mathbb{R}^2 \mid \|x - C\|_2 < R - \frac{1}{k} \right\}$$

On note $x_k = \mathrm{proj}_{\mathcal{C}_k}(x_0)$. Il est facile de montrer que les x_k pour $k \in \mathbb{N}$ sont alignés, et qu'ils définissent la même droite $\mathcal{H}$

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}_k}(x_0), x - x_0 \right\rangle = 0 = \left\langle x_0 - \mathrm{proj}_{\mathcal{C}_1}(x_0), x - x_0 \right\rangle$$

passant par x_0 et telle que $\mathcal{C}_k$ soit strictement contenu dans le demi-plan

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}_1}(x_0), x - x_0 \right\rangle < 0$$

Soit $x \in \mathcal{C}$. Il existe $k_0 \in \mathbb{N}^*$ tel que $x \in \mathcal{C}_{k_0}$. Aussi, on a bien

$$\left\langle x_0 - \mathrm{proj}_{\mathcal{C}_1}(x_0), x - x_0 \right\rangle < 0$$

On peut enfin démontrer une version du théorème de Hahn–Banach :

Théorème 3 (Hahn–Banach)

Soit $\mathcal{C}_1$ et $\mathcal{C}_2$ deux convexes disjoints et non vides. On suppose que $\mathcal{C}_2$ est **ouvert**. Alors il existe $a \in \mathcal{X}$ et $b \in \mathbb{R}$ tel que

$$\forall\,(x_1, x_2) \in \mathcal{C}_1 \times \mathcal{C}_2, \qquad \langle a, x_2\rangle \leq b \leq \langle a, x_1\rangle$$

Remarque : On dit que l'hyperplan $\mathcal{H}$ défini par

$$\langle a, x\rangle = b$$

sépare strictement les deux ensembles $\mathcal{C}_1$ et $\mathcal{C}_2$. On voit donc que le théorème de Hahn–Banach généralise la proposition 7.

Démonstration : On définit l'ensemble suivant :

$$\mathcal{C} = \mathcal{C}_2 - \mathcal{C}_1 = \bigcup_{x_1 \in \mathcal{C}_1} (\mathcal{C}_2 - x_1)$$

D'après la proposition 1 l'ensemble $\mathcal{C}$ est convexe. Puisque chacun des $\mathcal{C}_2 - x_1$ est un ensemble ouvert (c'est la translatée de $\mathcal{C}_2$), il s'ensuit par union que $\mathcal{C}$ est ouvert. Enfin, $\mathcal{C}$ est non vide. Aussi, on peut appliquer la proposition 7 : puisque $\mathcal{C}_1$ et $\mathcal{C}_2$ sont disjoints, pour tout $(x_1, x_2) \in \mathcal{C}_1 \times \mathcal{C}_2$ on a $x_1 \neq x_2$, donc $x_2 - x_1 \neq 0$. Autrement dit, $0 \notin \mathcal{C}$, et il existe donc un hyperplan $\mathcal{H}$ passant par 0 et définissant un demi-espace ouvert contenant $\mathcal{C}$. Ainsi, il existe $(a, b_0) \in \mathcal{X} \times \mathbb{R}$ tel que

$$\langle a, 0\rangle = b_0 \qquad \text{et} \qquad \forall\, x \in \mathcal{C}, \quad \langle a, x\rangle < b_0$$

La première égalité assure que $b_0 = 0$, tandis que la seconde inégalité s'écrit

$$\forall\,(x_1, x_2) \in \mathcal{C}_1 \times \mathcal{C}_2, \qquad \langle a, x_2 - x_1\rangle < 0 \qquad \text{soit} \qquad \langle a, x_2\rangle < \langle a, x_1\rangle$$

Posons $f(x) = \langle a, x\rangle$. Soit $x_2 \in \mathcal{C}_2$. On vient de montrer que

$$\forall\, x_1 \in \mathcal{C}_1, \qquad f(x_1) \geq \langle a, x_2\rangle$$

En passant à la borne inférieure, on obtient que

$$\inf_{x_1 \in \mathcal{C}_1} f(x_1) \geq \langle a, x_2 \rangle$$

ce qui implique que $b = \inf_{x_1 \in \mathcal{C}_1} f(x_1)$ est finie. ■