

Épreuve d'analyse numérique

Projections convexes, décomposition, et algorithme de projections alternées

Dans tout ce sujet, on se place dans un espace de Hilbert : H désigne un espace vectoriel réel complet pour la norme, notée $\|\cdot\|$, associée à un produit scalaire, noté $\langle \cdot, \cdot \rangle$. Dans cet espace normé, la boule fermée de centre $x \in H$ et de rayon $R > 0$ est notée $B(x, R)$. Rappelons que le produit scalaire vérifie l'inégalité de Cauchy-Schwarz : pour tous x et y dans H ,

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Cette épreuve propose l'étude de propriétés géométriques des ensembles convexes fermés de H , en particulier des cônes convexes fermés, et des algorithmes de projections convexes alternées. Rappelons qu'une partie $C \subset H$ est dite convexe si

$$\forall x \in C, y \in C, \quad \forall t \in [0, 1], \quad tx + (1 - t)y \in C.$$

Les parties II et III sont indépendantes, mais utilisent des résultats de la partie I. La qualité de la présentation et de l'argumentation sera un élément important de la notation.

Partie I – Projection sur les convexes fermés

Rappelons le résultat fondamental des espaces de Hilbert. Soit C un convexe fermé de H ; alors, pour tout $x \in H$, il existe un unique élément de C , appelé la projection de x sur C et noté $p_C(x)$, tel que

$$\forall y \in C, \quad \|x - p_C(x)\| \leq \|x - y\|.$$

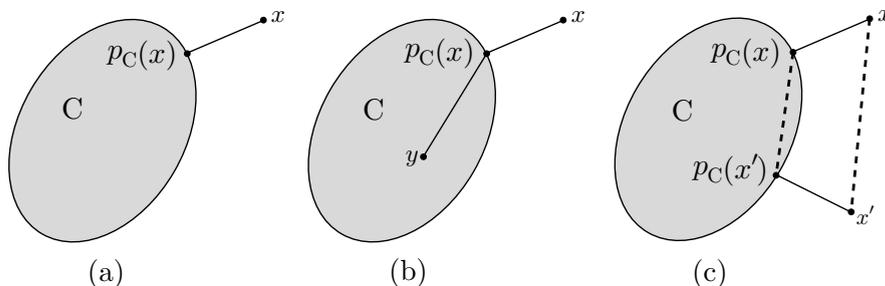
En d'autres termes, $p_C(x)$ est l'unique élément de C qui réalise la distance de x à C , notée $d_C(x)$:

$$d_C(x) = \min_{y \in C} \|x - y\| = \|x - p_C(x)\|.$$

De plus, $p = p_C(x)$ est caractérisé par les conditions

$$p \in C \quad \text{et} \quad \langle x - p, y - p \rangle \leq 0 \quad \text{pour tout } y \in C. \quad (1)$$

Les figures ci-dessous illustrent la projection sur un convexe fermé (a), la caractérisation angulaire (b), et la propriété de l'opérateur projection montrée à la question **2c** (c).



1) Théorème de projection. Dans cette question, nous démontrons le théorème de projection rappelé ci-dessus. Soient C un convexe fermé (non vide) de H et $x \in H$.

a) Montrer l'égalité du parallélogramme : pour tous u et v dans H , on a

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

b) Montrer que pour tous y et y' dans C , on a

$$\left\| \frac{y-y'}{2} \right\|^2 \leq \frac{1}{2}(\|y-x\|^2 + \|y'-x\|^2) - d_C(x)^2.$$

c) Montrer qu'il existe au plus un point qui réalise la distance $d_C(x)$.

d) Soit une suite $(y_n)_{n \in \mathbb{N}}$ de C telle que $\lim_{n \rightarrow +\infty} \|y_n - x\| = d_C(x)$. Montrer que la suite $(y_n)_{n \in \mathbb{N}}$ est de Cauchy. Conclure sur l'existence et l'unicité de la projection de x sur C .

2) Caractérisation angulaire de la projection et corollaires

a) Montrer l'équivalence entre $p = p_C(x)$ et p vérifie la propriété (1). [Indication : pour une des deux implications, on pourra introduire le point $y_t = (1-t)p + ty$ pour $t \in]0, 1[$.]

b) En déduire que, pour tous x et x' dans H , on a

$$\|p_C(x) - p_C(x')\|^2 \leq \langle p_C(x) - p_C(x'), x - x' \rangle.$$

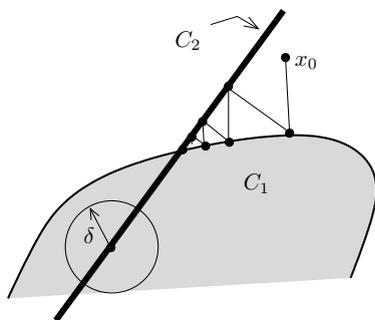
c) En déduire que l'opérateur de projection $p_C: H \rightarrow C \subset H$ est lipschitzien de rapport 1, c'est-à-dire que pour tous x et x' dans H , on a

$$\|p_C(x) - p_C(x')\| \leq \|x - x'\|.$$

d) En déduire que si $0 \in C$ alors $\|p_C(x)\| \leq \|x\|$ pour tout $x \in H$.

Partie II – Algorithme de projections alternées

De nombreux problèmes (en traitement du signal et de l'image par exemple) se modélisent de la manière suivante : trouver un élément qui satisfait à la fois deux types de contraintes « simples » ; plus précisément, trouver $x \in H$ tel que $x \in C_1 \cap C_2$ avec C_1 et C_2 deux convexes fermés de H (dont l'intersection est non vide) et pour lesquels on sait calculer p_{C_1} et p_{C_2} (mais pas $p_{C_1 \cap C_2}$).



Le prototype des algorithmes pour résoudre ce type de problème est la méthode « des projections alternées » que l'on étudie dans cette partie. L'idée de cet algorithme est de répéter le processus suivant : projeter un point sur C_1 , puis projeter le point obtenu sur C_2 , puis de nouveau sur C_1 , et ainsi de suite, comme l'illustre la figure ci-contre.

Dans cette partie, nous allons donc nous intéresser à la convergence de la suite de points de H définie par un itéré initial $x_0 \in H$ et par l'itération

$$x_{n+1} = p_{C_2}(p_{C_1}(x_n)),$$

avec C_1 et C_2 deux convexes fermés de H (dont l'intersection est non vide). Nous allons montrer que la suite $(x_n)_{n \in \mathbb{N}}$ s'approche de l'intersection $C_1 \cap C_2$ « à vitesse géométrique ». Plus précisément, nous allons montrer qu'il existe une constante $\alpha < 1$ (dépendante de « la géométrie » de l'intersection) et une constante $M > 0$, telles que, pour tout $n \in \mathbb{N}$,

$$d_{C_1 \cap C_2}(x_n) \leq M\alpha^n, \tag{2}$$

et nous en déduirons que la suite $(x_n)_{n \in \mathbb{N}}$ converge vers un point de l'intersection $C_1 \cap C_2$ avec cette même vitesse. Nous allons montrer ces propriétés sous l'hypothèse que C_2 intersecte l'intérieur de C_1 , noté $\text{int } C_1$. Pour simplifier cette étude de convergence, on suppose en fait que l'on a

$$0 \in (\text{int } C_1) \cap C_2,$$

ce qui se fait sans perte de généralité puisqu'il suffit de translater les deux ensembles d'un vecteur (inconnu) appartenant à l'intersection $(\text{int } C_1) \cap C_2$ non vide. Ainsi, dans toute cette partie, nous notons $\delta > 0$ un rayon tel que $B(0, \delta) \subset C_1$.

3) Montrer que pour tout $x \in C_1 \cap C_2$ et pour tout $m \in \mathbb{N}$

$$\|x_{m+1} - x\| \leq \|x_m - x\|. \quad (3)$$

En déduire que la suite $(x_n)_{n \in \mathbb{N}}$ est bornée. Dans la suite de cette partie **II**, on notera R un rayon tel que $x_n \in B(0, R)$ pour tout $n \in \mathbb{N}$.

4) Borne sur la distance à l'intersection. Nous allons montrer qu'il existe $\kappa > 0$ tel que

$$d_{C_1 \cap C_2}(y) \leq \kappa \max\{d_{C_1}(y), d_{C_2}(y)\} \quad \text{pour tout } y \in B(0, R). \quad (4)$$

a) Soit $u \in H$. Montrer que, si $\varepsilon > d_{C_1}(u)$, alors $w = \frac{\delta}{\delta + \varepsilon}u \in C_1$. [Indication : on pourra introduire le point $w_1 = \frac{\delta}{\varepsilon}(u - p_{C_1}(u))$.]

b) Soit $y \in B(0, R)$ et notons $u = p_{C_2}(y)$. Montrer que, si $\varepsilon > 2 \max\{d_{C_1}(y), d_{C_2}(y)\}$, alors $\varepsilon > d_{C_1}(u)$. En déduire que le point $w = \frac{\delta}{\delta + \varepsilon}u$ est dans l'intersection $C_1 \cap C_2$.

c) En déduire l'inégalité (4) avec $\kappa = (1 + 2R/\delta)$. [Indication : la question **2d** entre en jeu.]

5) Convergence de la suite des distances. Nous allons montrer à présent qu'il existe $\alpha < 1$ tel que

$$d_{C_1 \cap C_2}(x_{n+1}) \leq \alpha d_{C_1 \cap C_2}(x_n), \quad \text{pour tout } n \in \mathbb{N}. \quad (5)$$

a) Soit $x \in H$; vérifier tout d'abord que

$$d_{C_1 \cap C_2}(p_{C_2}(x))^2 \leq d_{C_1 \cap C_2}(x)^2 + d_{C_2}(x)^2 + 2\langle p_{C_1 \cap C_2}(x) - x, x - p_{C_2}(x) \rangle,$$

puis que

$$d_{C_1 \cap C_2}(p_{C_2}(x))^2 \leq d_{C_1 \cap C_2}(x)^2 - d_{C_2}(x)^2.$$

b) En déduire que, pour tout $x \in C_1 \cap B(0, R)$, alors

$$d_{C_1 \cap C_2}(p_{C_2}(x)) \leq \sqrt{1 - 1/\kappa^2} d_{C_1 \cap C_2}(x).$$

Vérifier qu'on a de même pour tout $x \in C_2 \cap B(0, R)$,

$$d_{C_1 \cap C_2}(p_{C_1}(x)) \leq \sqrt{1 - 1/\kappa^2} d_{C_1 \cap C_2}(x).$$

c) En déduire l'inégalité (5) avec $\alpha = 1 - 1/\kappa^2$. Conclure sur la convergence de la suite $(d_{C_1 \cap C_2}(x_n))_n$.

6) Convergence de la suite $(x_n)_n$. Pour tout $\ell \in \mathbb{N}$, notons $\delta_\ell = d_{C_1 \cap C_2}(x_\ell)$ et $p_\ell = p_{C_1 \cap C_2}(x_\ell)$. Considérons pour tout $n \in \mathbb{N}$

$$D_n = \bigcap_{\ell=0}^n B(p_\ell, \delta_\ell).$$

a) En utilisant l'inégalité (3) de la question **3a**, montrer que $x_n \in D_n$ pour tout $n \in \mathbb{N}$.

b) Montrer que le diamètre de D_n tend vers 0. On rappelle que le diamètre d'un ensemble $S \subset H$ est défini par

$$\text{diam } S = \sup\{\|x - y\| : x, y \in S\}.$$

c) Montrer que la suite $(x_n)_{n \in \mathbb{N}}$ est convergente et que sa limite \bar{x} est vérifiée, pour tout n ,

$$\|x_n - \bar{x}\| \leq 2 d_{C_1 \cap C_2}(x_n).$$

d) Montrer que $\bar{x} \in \bigcap_{n \in \mathbb{N}} D_n$, en déduire que $\bar{x} \in C_1 \cap C_2$. Conclure sur la convergence de la suite $(x_n)_{n \in \mathbb{N}}$ construite par l'algorithme des projections alternées.

7) Algorithme des projections moyennées Nous étudions maintenant une extension du résultat précédent à un algorithme proche : l'algorithme des « projections moyennées ». Soient N ensembles K_1, \dots, K_N convexes fermés de H . On considère la suite d'éléments de H définie par un itéré initial $y_0 \in H$ et l'itération

$$y_{n+1} = \frac{1}{N} \sum_{i=1}^N p_{K_i}(y_n). \quad (6)$$

Nous allons montrer que cette suite converge aussi vers un point de l'intersection $\bigcap_i C_i$, avec la même vitesse que la suite produite par projections alternées (sous une hypothèse sur l'intersection).

Considérons tout d'abord l'espace hilbertien H^N muni du produit scalaire produit défini par

$$\langle Y, Z \rangle_N = \sum_{i=1}^N \langle y_i, z_i \rangle \quad \text{pour tous } Y = (y_1, \dots, y_N) \in H^N \text{ et } Z = (z_1, \dots, z_N) \in H^N.$$

a) Soit Δ « la diagonale » de H^N définie par

$$\Delta = \{Y = (y, \dots, y) : y \in H\}.$$

Montrer que Δ est un sous-espace vectoriel fermé de H^N et que l'orthogonal de Δ est

$$\Delta^\perp = \left\{ Z = (z_1, \dots, z_N) \text{ tel que } \sum_{i=1}^N z_i = 0 \right\}.$$

b) Montrer que la projection de $Y = (y_1, \dots, y_N) \in H^N$ sur Δ est

$$p_\Delta(Y) = \frac{1}{N} \left(\sum_{i=1}^N y_i, \dots, \sum_{i=1}^N y_i \right)$$

c) Supposons que $0 \in \bigcap_{i=1}^N \text{int } K_i$. En utilisant la question **6d**, montrer qu'il existe un point $\bar{y} \in \bigcap_{i=1}^N K_i$ tel que la suite $(y_n)_{n \in \mathbb{N}}$ produite par (6) converge vers \bar{y} plus vite qu'une suite géométrique.

Partie III – Cônes convexes : décomposition, exemples

Dans cette dernière partie, les questions **9**, **10** et **11** sont indépendantes (et utilisent des résultats montrés à la question **8**).

Soit K une partie de H telle que $0 \in K$. On dit que K est un cône si, pour tout $x \in K$, la demi-droite d'origine 0 et de direction x est incluse dans K ; en d'autres termes,

$$\forall x \in K, \quad \forall t > 0, \quad tx \in K.$$

On appelle cône polaire de K , la partie de H définie par

$$K^\circ = \{y \in H \text{ tel que } \langle y, x \rangle \leq 0 \text{ pour tout } x \in K\}.$$

Remarquer, grâce à la caractérisation (1), que K° est exactement l'ensemble des points de H dont la projection sur K est 0.

8) Polaire et décomposition de Moreau. Dans toute cette question, on considère un cône K convexe et fermé dans H . Nous allons montrer le résultat suivant (théorème de décomposition de Moreau) : pour x, x_K et x_{K° trois éléments de H ; les propriétés suivantes sont équivalentes :

- (i) $x = x_K + x_{K^\circ}$ avec $x_K \in K$, $x_{K^\circ} \in K^\circ$ et $\langle x_K, x_{K^\circ} \rangle = 0$,
- (ii) $x_K = p_K(x)$ et $x_{K^\circ} = p_{K^\circ}(x)$.

- a) Montrer que K° est un cône convexe fermé dans H . Vérifier que, si K_1 et K_2 sont deux cônes convexes tels que $K_1 \subset K_2$, alors $K_1^\circ \supset K_2^\circ$. Vérifier aussi que si K est en fait un sous-espace vectoriel de H , alors K° est l'orthogonal de K .
- b) Soit $x \in H$; montrer que la caractérisation (1) s'écrit pour $p = p_K(x)$ par les trois conditions :

$$p \in K, \quad x - p \in K^\circ \quad \text{et} \quad \langle x - p, p \rangle = 0.$$

- c) Montrer que $(K^\circ)^\circ \supset K$, puis que $(K^\circ)^\circ = K$.
- d) Démontrer le résultat de décomposition énoncé ci-dessus.
- e) **Illustration** Considérons \mathbb{R}^n muni du produit scalaire canonique $\langle x, y \rangle_{\mathbb{R}^n} = \sum_{i=1}^n x_i y_i$. Montrer que $((\mathbb{R}_+)^n)^\circ = (\mathbb{R}_-)^n$. Donner (et démontrer) l'expression explicite de la projection d'un point $x \in \mathbb{R}^n$ sur $(\mathbb{R}_+)^n$.

9) Cônes polyédriques Soient m vecteurs a_1, \dots, a_m de \mathbb{R}^n ; on considère le sous-ensemble de \mathbb{R}^n

$$K_n = \left\{ x = \sum_{i=1}^m \alpha_i a_i \text{ avec } \alpha_i \geq 0 \text{ pour tout } i = 1, \dots, m \right\}.$$

- a) Montrer que K_n est un cône convexe. Dessiner K_2 dans le cas de deux vecteurs a_1, a_2 du plan \mathbb{R}^2 . Toujours dans ce cas, deviner et dessiner son cône polaire.
- b) Donner (et démontrer) l'expression du cône polaire de K_n .
- c) On admet que K_n est fermé. En déduire que, pour un vecteur $b \in \mathbb{R}^n$ quelconque, une et une seule des deux propriétés suivantes est vérifiée :
- il existe m réels $\alpha_i \geq 0$ tels que $b = \sum_{i=1}^m \alpha_i a_i$,
 - il existe $y \in \mathbb{R}^n$ tel que $\langle b, y \rangle_{\mathbb{R}^n} > 0$ et $\langle a_i, y \rangle_{\mathbb{R}^n} \leq 0$ pour tout $i = 1, \dots, m$.

10) Cône des matrices (semidéfinies) positives Soit une matrice X ; notons X_{ij} son coefficient à la ligne i et colonne j , et notons $\text{trace}(X)$ sa trace. Plaçons-nous dans $(S_n, \langle \cdot, \cdot \rangle)$ l'espace vectoriel des matrices (réelles) symétriques de taille n , muni du produit scalaire

$$\langle X, Y \rangle_{S_n} = \sum_{i,j=1}^n X_{ij} Y_{ij}, \quad \text{pour tous } X, Y \in S_n.$$

On identifie les vecteurs de \mathbb{R}^n avec les matrices colonnes de taille n .

- a) Montrer que $\langle X, Y \rangle = \text{trace}(XY)$ pour tout $X, Y \in S_n$.
- b) Montrer que l'ensemble des matrices (semidéfinies) positives

$$S_n^+ = \{X \in S_n \text{ tel que } \langle u, Xu \rangle_{\mathbb{R}^n} \geq 0 \text{ pour tout } u \in \mathbb{R}^n\}$$

est un cône convexe fermé de S_n .

c) **Visualisation en 3 dimensions** Pour cette question, considérons le cas $n = 2$. Vérifier que

$$S_2^+ = \left\{ A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in S_2 \text{ tel que } a + c \geq 0 \text{ et } ac - b^2 \geq 0 \right\}.$$

Montrer que l'application φ est une isométrie

$$\varphi: A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in S_2 \longmapsto \varphi(A) = \frac{1}{\sqrt{2}}(2b, c - a, c + a) \in \mathbb{R}^3.$$

Dessiner, dans \mathbb{R}^3 , le cône S_2^+ identifié à $\varphi(S_2^+)$.

- d) Montrer que, pour tous X et Y dans S_n^+ , on a $\langle X, Y \rangle \geq 0$.
- e) Montrer que $(S_n^+)^\circ = -S_n^+$.

f) Soit $X \in S_n$; considérons une décomposition spectrale $X = UDU^\top$ avec U une matrice orthogonale et D une matrice diagonale $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ avec les valeurs propres de X sur la diagonale. Montrer que

$$p_{S_n^+}(X) = U \text{Diag}(\max\{\lambda_1, 0\}, \dots, \max\{\lambda_n, 0\})U^\top.$$

11) Décomposition sur les « champs de gradient » des fonctions convexes Soit $I =]a, b[$ un intervalle de \mathbb{R} . L'objectif de cette question est de démontrer un résultat de décomposition d'une fonction de carré intégrable mettant en jeu le sous-ensemble des dérivées des fonctions convexes. On rappelle qu'une fonction $f: I \rightarrow \mathbb{R}$ est convexe si

$$\forall x \in I, y \in I, \quad \forall t \in [0, 1], \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Commençons par quelques rappels sur les fonctions convexes : pour les questions **a**, **b**, **c** et **d**, on fixe une fonction $f: I \rightarrow \mathbb{R}$ convexe, et nous voulons montrer que f est dérivable presque partout (au sens de Lebesgue).

a) Soient deux réels $x < y$ de I et $u \in]x, y[$. Montrer que

$$f(u) \leq \frac{y-u}{y-x}f(x) + \left(1 - \frac{y-u}{y-x}\right)f(y).$$

En déduire que

$$\frac{f(u) - f(x)}{u - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(u)}{y - u}.$$

Visualiser cette propriété sur un dessin.

b) Soit $x_0 \in I$; montrer que l'application « pente »

$$s(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

est croissante sur $I \setminus \{x_0\}$.

c) Soit $x_0 \in I$; montrer que f admet une dérivée à droite et une dérivée à gauche en x_0

$$f'_-(x_0) = \lim_{x \rightarrow x_0, x < x_0} s(x) \quad \text{et} \quad f'_+(x_0) = \lim_{y \rightarrow x_0, y > x_0} s(y),$$

qui satisfont $f'_-(x_0) \leq f'_+(x_0)$.

d) Montrer que l'ensemble des points où f n'est pas dérivable est au plus dénombrable.

Soit à présent L^2 l'espace de Hilbert des (classes de) fonctions de I dans \mathbb{R} mesurables de carré intégrable muni du produit scalaire

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

Considérons la partie de L^2

$$K = \{g \in L^2 \text{ tel que } g = f' \text{ avec } f \text{ une fonction convexe}\},$$

qui est bien définie grâce à la question **d**. Considérons aussi l'ensemble S constitué des fonctions $h: I \rightarrow \mathbb{R}$ qui « préservent la mesure de Lebesgue » (nous ne donnons pas son expression précise car nous ne l'utiliserons jamais explicitement). En fait, nous n'utiliserons cet ensemble que via le résultat suivant (admis) : $g \in K$ si et seulement si $\langle g, h - \text{id} \rangle_{L^2} \leq 0$ pour tout $h \in S$ (id est l'application identité de I).

e) Montrer que K est un cône convexe fermé dans L^2 .

f) Montrer que K° est l'adhérence dans L^2 de l'ensemble des éléments $\sum_{i=1}^N \alpha_i (h_i - \text{id})$ avec $N \in \mathbb{N}$, $\alpha_i \geq 0$ et $h_i \in S$.

g) Soit $\varphi \in L^2$; montrer qu'il existe une fonction f convexe (unique à une constante additive près) et un unique $h \in K^\circ$ pour lesquels on a la décomposition $\varphi = f' + h$ avec $\langle f', h \rangle_{L^2} = 0$.