Notes de cours - Préparation à l'agrégation
Convexité et applications
ENS Rennes

Rozenn Texier-Picard*

Contents

*ENS Rennes, av. Robert Schuman, F-35170 Bruz, France ; rozenn.texier@ens-rennes.fr

1 Introduction

Ce cours, rédigé pour des étudiants préparant l'agrégation de mathématiques, a pour objet d'aborder de façon transversale les résultats fondamentaux liés à la notion de convexité en mathématiques. Il reste le plus souvent à un niveau élémentaire, et tente de faire le lien entre des résultats de nature géométrique et d'autres de nature analytique même si, par commodité, il est présenté en deux parties séparées. De plus, les notions sont autant que possible illustrées par des exemples et des applications, de difficulté variable. De même, des exercices de type vrai ou faux ? permettent de se poser beaucoup de petites questions qui souvent ne figurent pas dans les livres. Si l'intuition géométrique permet souvent d'y répondre, celle-ci trouve vite ses limites en dimension infinie...

De nombreux ouvrages de géométrie énoncent la plupart des résultats présentés ici, en se limitant le plus souvent à la dimension finie (voir la bibliographie). Lorsque c'est abordable et intéressant, nous nous poserons la question de la généralisation à la dimension infinie, qui a souvent des applications remarquables en analyse fonctionnelle.

Les résultats énoncés dans ce cours pourront être présentés à l'oral de l'agrégation, soit dans des leçons spécifiquement concernées par la convexité, soit dans beaucoup d'autres leçons d'algèbre et géométrie (formes linéaires et hyperplans, applications affines, angles et distances...) ou d'analyse (applications différentiables, problèmes d'extrêmums, continuité et différentiabilité, suites et séries de fonctions...) Dans cette optique, la plupart des résultats intéressants sont accompagnés de références bibliographiques permettant de retrouver les preuves.

Rappelons que la notion de convexité est très souvent un cadre idéal pour les problèmes de minimisation ; elle a donc une grande importance dans de nombreuses applications, de la physique à l'économie, en passant par à peu près toutes les sciences.

Les paragraphes les plus difficiles sont signalés par une astérisque. Ils peuvent être omis lors d'une première lecture.

Pour signaler une erreur, proposer un autre résultat ou une référence, ou pour toute question, on peut contacter l'auteur à l'adresse : `rozenn.texier@ens-rennes.fr`

2 Parties convexes d'un espace vectoriel

Dans toute cette partie, E sera un espace vectoriel réel, de dimension finie ou infinie. On pourra, lorsque c'est nécessaire, munir E d'une topologie (on se limitera essentiellement à des topologies métriques) compatible avec la structure d'espace vectoriel, c'est-à-dire rendant la translation et les homothéties continues.

2.1 Définitions et le théorème fondamental

2.1.1 Définition

Définition 2.1.1. Partie convexe

Soit $C \subset E$. Alors C est convexe si

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

Remarque : la définition s'étend de façon évidente au cas où l'espace E est affine. Pour simplifier les notations, et parce que la plupart des applications concerneront des espaces vectoriels, on se restreint à ce cadre.

Exemples faciles :

- un sous-espace affine est convexe,
- le simplexe de $\mathbb{R}^k$: $\Delta_k = \left\{ x \in (\mathbb{R}^+)^k, \sum_{i=1}^k x_i = 1 \right\}$,
- les intervalles de $\mathbb{R}$ sont convexes ; réciproquement, on vérifie aisément que tout convexe de $\mathbb{R}$ est un intervalle.

Proposition 2.1.2. Premières propriétés

Les propriétés ci-dessous se vérifient à partir de la définition.

- Une intersection quelconque de convexes est convexe.
- Le produit cartésien de deux convexes $A \subset E$ et $B \subset F$ est un convexe de l'espace produit $E \times F$.
- La somme de Minkowski de deux convexes est convexe.

Définition 2.1.3. Somme de Minkowski

Soient A et B deux parties de E. On définit leur somme de Minkowski, notée $A+B$, par la relation

$$A + B = \{a + b, a \in A, b \in B\}.$$

Cette notation sera souvent reprise dans la suite du cours. Elle ne doit bien sûr pas être confondue avec la réunion de A et B.

Exemple 2.1.4 Somme de Minkowski

- si $B = \{b\}$ est un singleton, alors $A + B$ n'est autre que l'image de A par la translation de vecteur b.
- si $B = B(0, r)$ est une boule ouverte, alors $A + B$ correspond au convexe A "épaissi" avec une épaisseur r.

2.1.2 Théorème fondamental de la géométrie affine *

Théorème 2.1.5. Théorème fondamental de la géométrie affine

Soit E un espace affine de dimension finie $N \geq 2$ ou de dimension infinie. Alors on a les résultats suivants :

i toute bijection de E dans E qui envoie un convexe sur un convexe est une bijection affine ;

ii toute bijection de E dans E qui envoie trois points alignés sur trois points alignés est une bijection affine.

Remarque 2.1.6
Ce résultat est clairement faux en dimension 1 puisque toute application continue envoie un convexe sur un convexe, d'après le théorème des valeurs intermédiaires.

La preuve de ce théorème est trop longue pour constituer un développement. Pour des références, voir Berger T.1 p. 77, ou Bourbaki, Espaces vectoriels topologiques, exercice 6 p. 183 et Algèbre, exercice 7 p. 201.

On se contente ici de montrer que la deuxième affirmation implique la première. En effet, soit f une bijection vérifiant les hypothèses du point 1, et soit $g = f^{-1}$. On va montrer que g envoie trois points alignés sur trois points alignés. Pour cela, il suffit de montrer que

$$\forall x, y, z \in E, z \in [x, y] \Rightarrow g(z) \in [g(x), g(y)].$$

D'après notre hypothèse, puisque le segment $[g(x), g(y)]$ est convexe, son image par f l'est. Or cette image contient $x = f(g(x))$ et $y = f(g(y))$ donc elle doit contenir le segment $[x, y]$. En particulier, elle contient le point z. De plus, $z = f(g(z))$ et f est bijective. Il en découle que $g(z) \in [g(x), g(y)]$. En admettant la propriété ii, on en déduit que g est affine, et donc f également.

2.2 Enveloppe convexe

2.2.1 Définitions

Définition 2.2.1. Enveloppe convexe

Soit $S \subset E$, on appelle enveloppe convexe de S, et on note $co(S)$, l'intersection de tous les convexes contenant S. Si E est muni d'une topologie métrique, on appelle enveloppe convexe fermée de S, notée $\bar{co}(S)$, l'intersection de tous les convexes fermés contenant S.

Une intersection de convexes étant convexe, il apparaît que $co(S)$ est également le plus petit convexe contenant S.

Proposition 2.2.2. Enveloppe convexe fermée

On a la relation : $\overline{co(S)} = \bar{co}(S)$, i.e. l'enveloppe convexe fermée est la fermeture de l'enveloppe convexe.

Preuve : exercice. Il faut démontrer au préalable que la fermeture d'un convexe est encore convexe.

Définition 2.2.3. Combinaison convexe

Soit $k \in \mathbb{N}^*$ et soient $x_1, .., x_k$ des points de E. On appelle *combinaison convexe* de $x_1, ..., x_k$ tout barycentre à coefficients positifs des points $x_1, ..., x_k$, i.e. tout point x tel que

$$\exists (t_1, ..., t_k) \in \Delta_k,\ x = \sum_{i=1}^{k} t_i x_i.$$

Remarquons qu'une combinaison convexe est toujours une combinaison *finie* de points de E.

Proposition 2.2.4. Enveloppe convexe

Soit $S \subset E$. Alors l'enveloppe convexe de S est l'ensemble des *combinaisons convexes* de points de S, i.e.

$$co(S) = \left\{ x \in E, \exists k \in \mathbb{N}^*, \exists \alpha \in \Delta_k, \exists (x_1, \cdots, x_k) \in S^k, x = \sum_{i=1}^{k} \alpha_i x_i \right\}.$$

Preuve : on note, pour $k \in \mathbb{N}^*$,

$$C_k = \left\{ x \in E, \exists \alpha \in \Delta_k, \exists (x_1, \cdots, x_k) \in S^k, x = \sum_{i=1}^{k} \alpha_i x_i \right\}$$

et on pose $C = \cup_{k \in \mathbb{N}^*} C_k$. On montre alors les points suivants :

- C est convexe (à partir de la définition),

- $S = C_1 \subset C$,
- Soit K un convexe contenant S. Alors K contient C_1. Par récurrence sur k, on montre que K contient tous les $C_k, k \geq 1$. Ainsi, K contient C.

Le théorème de Carathéodory montre qu'en dimension finie N, on peut se limiter à des combinaisons convexes à $N+1$ points de S pour définir son enveloppe convexe. Citons au préalable un exemple classique, dont la preuve se trouve par exemple dans [11].

Exemple 2.2.5 Théorème de Gauss-Lucas

Soit P un polynôme non constant à coefficients complexes. Alors les racines de P' sont dans l'enveloppe convexe des racines de P.

Preuve : On note $P(X) = \prod_{i=1}^{d}(X-\zeta_i)^{m_i}$. On a alors

$$\frac{P'(X)}{P(X)} = \sum_{i=1}^{d} \frac{m_i}{X-\zeta_i}. \tag{1}$$

Soit z une racine de P'. On veut montrer qu'elle est dans l'enveloppe convexe des ζ_i. Si $P(z)=0$, c'est clair. Supposons donc le contraire et écrivons l'égalité (1) au point z. On obtient :

$$0 = \sum_{i=1}^{d} \frac{m_i}{z-\zeta_i}$$

soit finalement

$$0 = \sum_{i=1}^{d} \frac{m_i(\bar{z}-\bar{\zeta}_i)}{|z-\zeta_i|^2}. \tag{2}$$

Notons

$$t_i = \frac{m_i}{|z-\zeta_i|^2} > 0, \quad T = \sum_i t_i, \quad \alpha_i = \frac{t_i}{T}.$$

Par passage au conjugué dans l'égalité (2) il vient

$$0 = \sum_i t_i(z-\zeta_i)$$

soit encore :

$$z = \sum_{i=1}^{d} \alpha_i \zeta_i,$$

ce qui termine la preuve.

2.2.2 Le théorème de Carathéodory

C'est un théorème classique sur la convexité, à connaître et à méditer ! La preuve se trouve par exemple dans [3, 7, 11].

Théorème 2.2.6. Carathéodory

Soit $S \subset \mathbb{R}^N$, alors tout point de $co(S)$ est combinaison convexe de $N+1$ points de S, i.e.

$$co(S) = \left\{ x \in E, \exists \alpha \in \Delta_{N+1}, \exists (x_1, \cdots, x_{N+1}) \in S^{N+1}, x = \sum_{i=1}^{N+1} \alpha_i x_i \right\} = C_{N+1}.$$

Preuve : il suffit de montrer l'inclusion $co(S) \subset C_{N+1}$. Soit donc $x \in co(S)$ et soit $k \geq 1$, $\alpha \in \Delta_k$, tels que $x = \sum_{i=1}^k \alpha_i x_i$. Si $k \leq N+1$, alors on a bien $x \in C_{N+1}$. Supposons maintenant $k > N+1$, et $\alpha_1, ..., \alpha_k > 0$.

Puisque $x_1, ..., x_k$ sont dans un espace affine de dimension N, et $k > N+1$, alors les points sont affinement dépendants, i.e. les vecteurs $x_2 - x_1, ..., x_k - x_1$ sont linéairement dépendants. Ainsi, il existe $\lambda_2, ..., \lambda_k$ non tous nuls tels que

$$\sum_{i=2}^k \lambda_i (x_i - x_1) = 0.$$

En posant $\lambda_1 = -\sum_{i=1}^k \lambda_i$, on a

$$\sum_{i=1}^k \lambda_i x_i = 0, \quad \sum_{i=1}^k \lambda_i = 0.$$

De cette propriété, on déduit que x peut se réécrire comme combinaison des x_i sous la forme

$$x = \sum_{i=1}^k (\alpha_i - t\lambda_i) x_i, \text{ avec } \sum_{i=1}^k (\alpha_i - t\lambda_i) = 1, \ \forall t \in \mathbb{R}.$$

L'idée est maintenant de jouer sur le nombre t pour que les $\alpha_i - t\lambda_i$ soient tous positifs ou nuls, et que l'un d'eux soit nul. Ainsi, on se sera ramené à une combinaison convexe à $k-1$ éléments.

Nécessairement l'un des λ_i au moins est strictement positif. Notons alors

$$t = \min \left\{ \frac{\alpha_i}{\lambda_i}, i = 1, ..., k, \lambda_i > 0 \right\},$$

et $\alpha_i' = \alpha_i - t\lambda_i$. Alors par construction, $(\alpha_1', ..., \alpha_k') \in \Delta_k$, et l'un des coefficients α_i' est nul (celui qui réalise le minimum) définissant t. On vérifie alors que x est combinaison convexe de $k-1$ éléments de S. En effet, on a

$$x = \sum_{i=1}^k \alpha_i' x_i.$$

On peut réitérer ce procédé tant que la famille $(x_1, ..., x_k)$ est affinement dépendante, donc tant que $k > N+1$. On se ramène donc finalement à une combinaison convexe de $N+1$ points de S.

Corollaire 2.2.7. Enveloppe convexe d'un compact

Soit S un compact de $\mathbb{R}^N$. Alors son enveloppe convexe est compacte.

Preuve : il suffit de vérifier que l'application ϕ, définie sur le compact $\Delta_{N+1} \times S^{N+1}$ par

$$\phi : \left| \begin{array}{lcl} \Delta_{N+1} \times S^{N+1} & \to & co(S) \\ (\alpha_1, ..., \alpha_{N+1}, x_1, ..., x_{N+1}) & \mapsto & \sum_i \alpha_i x_i \end{array} \right.$$

est surjective et continue.

Remarque 2.2.8

Attention, en dimension infinie, l'enveloppe convexe d'un compact n'est pas nécessairement fermée, donc a fortiori pas compacte ! Prenons par exemple un espace de Hilbert H muni d'une base hilbertienne $(e_n, n \in \mathbb{N})$. On pose pour tout $n \geq 0$, $f_n = \frac{e_n}{n+1}$. Soit K le compact défini par

$$K = \{f_n, n \in \mathbb{N}\} \cup \{0\}.$$

(C'est bien un compact puisqu'il est formé des termes d'une suite (f_n) convergente et de sa limite 0.) On définit maintenant une suite $(v_n)_{n\geq 1}$ dans $co(K)$ en posant

$$v_1 = \frac{f_1}{2}, \quad v_2 = \frac{f_1}{2} + \frac{f_2}{4}, \quad v_3 = \frac{f_1}{2} + \frac{f_2}{4} + \frac{f_3}{8}$$

et plus généralement,

$$v_n = \sum_{i=1}^{n} \frac{f_i}{2^i}.$$

Alors la suite (v_n) converge dans H vers la limite

$$v = \sum_{i=1}^{\infty} \frac{f_i}{2^i}.$$

Or v n'est pas combinaison convexe de points de K, sinon il serait combinaison d'un nombre *fini* de f_k, ce qui contredit l'unicité de la décomposition sur la base hilbertienne $(e_k, k \in \mathbb{N})$. On a donc montré que $co(K)$ n'est pas fermée.

On peut montrer en revanche que l'enveloppe convexe fermée d'un compact est compacte, voir [9].

2.2.3 Vrai ou faux ?

Voici à titre d'exercices une liste de questions plus ou moins naturelles, qui peuvent aider à prendre du recul par rapport aux notions abordées jusqu'à présent.

i VRAI OU FAUX ? $C \subset E$ est convexe si et seulement si la demi-somme de deux points de C est toujours dans C.

ii VRAI OU FAUX ? Si C est un convexe de E métrique, alors l'intérieur et l'adhérence de C sont convexes également.

iii VRAI OU FAUX ? Soit $S \subset \mathbb{R}^N$. Alors tout point de $co(S)$ est combinaison convexe de deux éléments de S.

iv VRAI OU FAUX ? L'enveloppe convexe d'un fermé de $\mathbb{R}^N$ est fermée.

v VRAI OU FAUX ? Soit $S \subset \mathbb{R}^N$. Alors $\bar{co}(S) = co(\bar{S})$.

vi VRAI OU FAUX ? Soit $C \subset \mathbb{R}^N$ un convexe d'intérieur vide. Alors C est inclus dans un hyperplan affine de $\mathbb{R}^N$.

vii VRAI OU FAUX ? Tout convexe fermé borné de $\mathbb{R}^N$ est homéomorphe à la boule unité de $\mathbb{R}^N$.

Et voici les réponses.

i FAUX. Par exemple $\mathbb{Q}$ n'est pas convexe dans $\mathbb{R}$. Par contre, le résultat est vrai si on impose de plus que C est fermé. La preuve utilise la densité des rationnels diadiques dans $[0, 1]$.

ii VRAI. S'appuyer sur des dessins pour les démonstrations, qui sont élémentaires.

iii FAUX, sauf pour $N = 1$ (c'est alors exactement le théorème de Carathéodory). Par exemple si $N \geq 2$, et si S est formé de trois points non alignés, les combinaisons convexes de deux éléments donnent seulement les bords du triangle.

iv FAUX. Contre-exemple dans $\mathbb{R}^2$: $S = \{(0, 0)\} \cup (\mathbb{R} \times 1)$, $co(S) = \{(0, 0)\} \cup (\mathbb{R} \times]0, 1])$.

Par contre, le résultat est vrai si on ajoute l'hypothèse S borné, d'après le corollaire 2.2.7.

v FAUX, le même contre-exemple que ci-dessus convient. Par contre, c'est vrai en dimension finie si on rajoute l'hypothèse : S borné. En effet, $\bar{S}$ est alors compact et on peut utiliser le corollaire 2.2.7. Remarquons que même si S est compact c'est faux en dimension infinie, cf remarque 2.2.8.

vi VRAI. Il suffit de montrer la contraposée, à savoir : pour tous points $x_1, ..., x_{N+1}$ affinement indépendants de $\mathbb{R}^N$, $co(x_1, ..., x_{N+1})$ est d'intérieur non vide. Cela se montre par récurrence sur $N \geq 1$. C'est clair pour $N = 1$. Supposons la propriété démontrée pour N points de $\mathbb{R}^{N-1}$, et considérons $N + 1$ points affinement indépendants de $\mathbb{R}^N$. Les points $x_1, ..., x_N$ engendrent un sous-espace affine de dimension $N - 1$ donc leur enveloppe convexe contient une boule ouverte de dimension $N - 1$. Ainsi $co(x_1, ..., x_{N+1})$ contient le cône de sommet x_{N+1} et de base cette boule. Il est donc d'intérieur non vide.

vii FAUX, par exemple l'ensemble vide est convexe fermé borné ; plus généralement, il existe des convexes fermés bornés d'intérieur vide, par exemple un segment dans le plan. Par contre, si on impose que le convexe soit d'intérieur non vide, la propriété est vraie, voir section suivante.

2.3 Jauge d'un convexe

Pour cette section, on peut se référer à [3]. Dans tout ce paragraphe, on suppose que E est un espace vectoriel normé.

2.3.1 Définition et propriétés

Soit $C \subset E$ un convexe tel que 0 est dans l'intérieur de C, et soit $x \in E$. On remarque que

$$\left\{\alpha > 0, \frac{x}{\alpha} \in C\right\}$$

est un intervalle de la forme $(A, +\infty[$ (ouvert ou fermé en A selon les cas).

Définition 2.3.1. Jauge d'un convexe

Soit $C \subset E$ un convexe tel que 0 est dans l'intérieur de C. On appelle jauge d'un convexe l'application $\rho_C : E \to \mathbb{R}^+$ définie par

$$\rho_C(x) = \inf\left\{\alpha > 0, \frac{x}{\alpha} \in C\right\}.$$

Exemples :

- pour $C = E$, la jauge est identiquement nulle ;
- pour $C = B(0, r)$, on a $\rho_C(x) = \frac{\|x\|}{r}$;
- pour C borné, on montre facilement que $\rho_C(x) = 0 \iff x = 0$ (exercice).

Proposition 2.3.2. Propriétés de la jauge

La fonction ρ_C vérifie les propriétés suivantes.

i elle est positivement homogène : $\forall x \in E, \forall \lambda \geq 0, \rho_C(\lambda x) = \lambda \rho_C(x)$;

ii $\{x \in E, \rho_C(x) = 1\} = \partial C$; $\{x \in E, \rho_C(x) < 1\} = \mathring{C}$;

iii elle est sous-additive : $\forall x, y \in E, \rho_C(x + y) \leq \rho_C(x) + \rho_C(y)$;

iv elle est continue.

Preuve :

i La propriété i est immédiate d'après la définition.

ii Il est clair que si $x \in C$, alors $\rho_C(x) \leq 1$. Montrons maintenant que $\{x \in E, \rho_C(x) < 1\} \subset C$. Soit donc x tel que $\rho_C(x) < 1$. Par définition, il existe $\alpha \in]0, 1[$ tel que $\frac{x}{\alpha} \in C$. On a aussi $0 \in C$ et par convexité $x \in C$. Inversement, si $x \in \mathring{C}$, alors on a immédiatement (par la propriété i) $\rho_C(x) < 1$, et si $x \notin \bar{C}$, alors $\rho_C(x) > 1$. Ainsi, si $\rho_C(x) \leq 1$, on a $x \in \bar{C}$. On a donc montré

$$\mathring{C} \subset \{x \in E, \rho_C(x) < 1\} \subset C \subset \{x \in E, \rho_C(x) \leq 1\} \leq \bar{C}. \qquad (3)$$

Avant de finir la démonstration de la propriété ii, examinons les propriétés iii et iv.

iii Montrons la sous-additivité. L'idée de la preuve est présentée sur la figure 3. Soient $\epsilon > 0$ fixé, $x, y \in E$. On note

$$\bar{x} = \frac{x}{\rho_C(x) + \epsilon}, \quad \bar{y} = \frac{y}{\rho_C(y) + \epsilon}.$$

Figure 1: Preuve de la sous-additivité.

On a d'après le point 1, $\rho_C(\bar{x}) < 1, \rho_C(\bar{y}) < 1$ donc d'après (3), on en déduit que $\bar{x} \in C, \bar{y} \in C$. Posons alors

$$\alpha = \frac{\rho_C(x) + \epsilon}{\rho_C(x) + \rho_C(y) + 2\epsilon}, \quad \bar{z} = \alpha \bar{x} + (1 - \alpha)\bar{y}.$$

Alors $\bar{z} \in C$ donc $\rho_C(\bar{z}) \leq 1$. D'autre part, on obtient après calcul

$$\bar{z} = \frac{x + y}{\rho_C(x) + \rho_C(y) + 2\epsilon}.$$

On obtient donc

$$\rho_C(x + y) \leq \rho_C(x) + \rho_C(y) + 2\epsilon,$$

ce qui donne bien le résultat attendu en faisant tendre ϵ vers 0.

iv Montrons la continuité de ρ_C. Elle repose sur le fait que 0 est intérieur à C. En effet, soit $r > 0$ tel que la boule fermée $\bar{B}(0,r)$ soit incluse dans C. Alors

$$\forall y \in E, \frac{y}{\|y\|} r \in C,$$

d'où on déduit aisément que

$$\forall y \in E,\ \rho_C(y) \leq \frac{\|y\|}{r}. \tag{4}$$

De plus, en utilisant la sous-additivité,

$$\rho_C(x+y) \leq \rho_C(x) + \rho_C(y), \quad \rho_C(x) \leq \rho_C(x+y) + \rho_C(-y),$$

de sorte que

$$|\rho_C(x+y) - \rho_C(x)| \leq \max(\rho_C(y), \rho_C(-y)) \leq \frac{\|y\|}{r}.$$

On a ainsi montré le caractère lipschitzien (donc continu) de la jauge.

Nous pouvons maintenant terminer la preuve de la propriété ii. De la continuité de ρ_C, on déduit que $\{x \in E, \rho_C(x) < 1\}$ est un ouvert de E et $\{x \in E, \rho_C(x) \leq 1\}$ est un fermé de E. Ainsi, la relation (3) donne immédiatement que

$$\mathring{C} = \{x \in E, \rho_C(x) < 1\},\ \bar{C} = \{x \in E, \rho_C(x) \leq 1\}.$$

En considérant $\partial C = \bar{C} \setminus \mathring{C}$, on obtient la propriété ii.

2.3.2 Applications

La jauge d'un convexe est un élément essentiel dans la démonstration des théorèmes de séparation dans le cas EVN. Nous la retrouverons donc dans la suite du cours. En attendant, elle permet aussi de démontrer le résultat ci-dessous.

Corollaire 2.3.3. Homéomorphisme avec une boule

Soit C un convexe borné d'intérieur non vide d'un espace vectoriel normé E. Alors il existe un homéomorphisme $f : E \to E$ qui envoie $\mathring{C}$ sur la boule ouverte $B(0,1)$ et la frontière ∂C sur la sphère $\partial B(0,1)$.

Preuve : quitte à translater C, on suppose que 0 est un point intérieur ; on définit alors l'homéomorphisme $f : E \to E$ par

$$f(x) = \begin{cases} 0, & \text{si } x = 0, \\ \rho_C(x)\frac{x}{\|x\|}, & \text{sinon.} \end{cases}$$

La fonction f est clairement continue en tout $x \neq 0$, elle est aussi continue en 0 grâce à l'inégalité (4). De plus, f est bijective, d'inverse

$$f^{-1} : x \mapsto \begin{cases} 0, & \text{si } x = 0, \\ \frac{\|x\|}{\rho_C(x)} x, & \text{sinon.} \end{cases}$$

De même que pour f, f^{-1} est continue en dehors de 0, et elle est continue en 0 grâce à l'inégalité

$$\forall y \in E,\ \rho_C(y) \geq \frac{\|y\|}{R}$$

où R est le rayon d'une boule centrée en 0 et contenant C. La propriété ii ci-dessus permet de conclure.

2.4 Projection sur un convexe fermé, applications

Il s'agit d'un résultat central en théorie de la convexité, et qui doit être maîtrisé, au moins en dimension finie pour la norme euclidienne et dans les espaces de Hilbert.

2.4.1 Dimension finie, norme euclidienne

Théorème 2.4.1. Théorème de projection (cas euclidien)

On munit $\mathbb{R}^N$ de la norme euclidienne, notée $\|x\|$ et du produit scalaire canonique, noté $\langle x, y\rangle$. Soit C un convexe fermé non vide de $\mathbb{R}^N$ et soit $x \in \mathbb{R}^N$. Alors

i il existe un unique point $y \in C$ tel que $\|y - x\| = \inf_{z\in C} \|z - x\|$. Ce point y s'appelle projeté de x sur C et sera noté $y = p_C(x)$.

ii Le point y est caractérisé par la propriété :

$$\forall z \in C, \langle x - y, z - y\rangle \leq 0. \tag{5}$$

Figure 2: Projection sur un convexe fermé non vide.

Preuve :

i L'existence provient essentiellement d'un argument de compacité. Puisque C est fermé, on se ramène à minimiser la fonction continue $z \mapsto \|z - x\|$ sur le compact non vide $C \cap \bar{B}(x, \|x - z_0\|)$, où z_0 est un point quelconque de C. (La dimension finie est essentielle ici pour avoir la compacité de la boule !)

L'unicité provient de la convexité de C et du théorème de Pythagore (la structure euclidienne est ici essentielle) : si deux projetés différents existent, alors par le théorème de Pythagore leur milieu est meilleur, et il est toujours dans C, ce qui contredit la définition du projeté.

ii Montrons la caractérisation (5).

$\Rightarrow$: soit y le projeté de x et soit $z \in C$. On pose $z_t = y + t(z - y) \in C$, $t \in]0, 1[$. On a donc par définition de y,

$$\|z_t - x\|^2 = t^2\|z - y\|^2 + 2t\langle z - y, y - x\rangle + \|y - x\|^2 \geq \|y - x\|^2.$$

En divisant par $t > 0$ et en faisant tendre t vers 0, on obtient l'inégalité recherchée.

$\Leftarrow$: soient y vérifiant la caractérisation et soit $z \in C$. On a :

$$\begin{aligned} 0 &\geq \langle x - y, z - y\rangle \\ &= \langle x - y, x - y + z - x\rangle \\ &= \|x - y\|^2 + \langle x - y, z - x\rangle \\ &\geq \|x - y\|^2 - \|x - y\|\|z - x\| \text{ (inégalité de Cauchy Schwarz)} \end{aligned}$$

D'où le résultat.

Figure 3: Un exemple avec une infinité de projetés et une norme non euclidienne.

Proposition 2.4.2. Continuité de la projection

L'application $x \mapsto p_C(x)$ est 1-lipschitzienne, i.e.

$$\forall x_1, x_2 \in \mathbb{R}^n \times \mathbb{R}^n, \|p_C(x_1) - p_C(x_2)\| \leq \|x_1 - x_2\|.$$

Preve : on utilise la caractérisation du projeté.

$$\begin{aligned} \langle x_2 - p_C(x_2), p_C(x_1) - p_C(x_2)\rangle &\leq 0, \\ \langle x_1 - p_C(x_1), p_C(x_2) - p_C(x_1)\rangle &\leq 0, \end{aligned}$$

ce qui donne après sommation des deux inégalités :

$$\|p_C(x_1) - p_C(x_2)\|^2 \leq \langle x_1 - x_2, p_C(x_1) - p_C(x_2)\rangle \leq \|p_C(x_1) - p_C(x_2)\|\|x_1 - x_2\|.$$

D'où le résultat.

2.4.2 Vrai ou faux ?

i VRAI OU FAUX ? En dimension finie, quelle que soit la norme considérée, il existe toujours un unique projeté sur un convexe fermé non vide C.

ii VRAI OU FAUX ? Si C est un fermé non vide de $\mathbb{R}^N$ ayant la propriété de projeté unique, alors il est convexe.

iii VRAI OU FAUX ? Soit C un convexe compact non vide de $\mathbb{R}^N$. Alors $\partial C = p_C(\mathbb{R}^N \setminus C)$.

Les réponses :

i FAUX. En cas de changement de norme, toujours en dimension finie, on garde l'existence d'une solution mais on perd l'unicité. Prenons par exemple $\mathbb{R}^2$ muni de la norme $\|\cdot\|_\infty$. Alors le point $(0, 1)$ a une infinité de projetés sur la droite d'équation $y = 0$, voir figure 3.

ii VRAI, ce résultat a été démontré indépendamment par L.N.H. Bunt (néerlandais, 1934) et Theodore Samuel Motzkin (israélo-américain,1935), mais est connu sous le nom de théorème de Motzkin. Voir par exemple [11] pour une démonstration.

Théorème 2.4.3. Théorème de Motzkin

Soit A un fermé non vide de $\mathbb{R}^N$, vérifiant

$$\forall x \in \mathbb{R}^N, \exists! y \in A, d(x, A) = \|x - y\|.$$

Alors A est convexe.

Figure 4: Une idée de la preuve du théorème de Motskin

Figure 5: Projection d'une sphère sur un convexe compact

On présente ici simplement l'idée de la preuve sur un dessin (voir Figure 4). Si A n'est pas convexe, alors on peut trouver $x, y \in A$, $z \in]x, y[$ et $\epsilon > 0$ tels que $A \cap B(z, \epsilon) = \emptyset$. (On utilise ici le caractère fermé de A). Considérons toutes les boules fermées contenant $B(z, \epsilon)$ et dont l'intérieur ne rencontre pas A. La clé de la démonstration consiste à prouver qu'il existe parmi elles une boule B_0 de rayon maximal. Par unicité du projeté de son centre sur A, le bord de B_0 rencontre A en un seul point. Mais alors en la décalant un peu on peut obtenir une boule fermée de rayon identique, qui ne rencontre pas A. Cette boule étant compacte, elle est à distance strictement positive de A, ce qui contredit la maximalité du rayon de B_0.

iii VRAI, voir [10]. L'inclusion $p_C(\mathbb{R}^N \setminus C) \subset \partial C$ est claire. Pour l'autre inclusion, soit B une boule ouverte contenant C, et soit S le bord de B. On va montrer que $\partial C \subset p_C(S)$. La preuve est illustrée par la figure 5.

Soit donc $x \in \partial C$, et $i \geq 1$. Soit $x_i \in B \setminus C$ tel que $|x - x_i| < \frac{1}{i}$. On note $u_i = p_C(x_i)$. Alors

$$|x - u_i| \leq |x - x_i| < \frac{1}{i}.$$

Soit maintenant y_i le point d'intersection de la sphère S et de la demi-droite $[u_i x)$. Alors on a $p_C(y_i) = u_i$. De la suite (y_i) on extrait une sous-suite convergeant vers un $y \in S$. Par continuité de la projection, il vient

$$p_C(y) = \lim_{i \to +\infty} p_C(y_i) = \lim_{i \to +\infty} u_i = x.$$

On a donc montré le lemme suivant.

Lemme 2.4.4. Projection d'une sphère sur un convexe compact

Soit C un convexe compact non vide de $\mathbb{R}^N$. Soit B une boule ouverte contenant C, et soit S sa frontière. Alors le bord de C est l'image de S par p_C.

2.4.3 Le cas de la dimension infinie

Attention, en dimension infinie les choses se compliquent ! On a vu que la preuve de l'existence utilisait la compacité des boules, elle ne fonctionne donc qu'en dimension finie. Toutefois le résultat s'étend partiellement à des espaces de dimension infinie.

Théorème 2.4.5. Projection, cas hilbertien

Soit H un espace de Hilbert, de norme notée $\|x\|$ et de produit scalaire noté $\langle x, y\rangle$. Soit C un convexe fermé non vide de H et soit $x \in H$. Alors

i il existe un unique point $y \in C$ tel que $\|y - x\| = \inf_{z\in C} \|z - x\|$. Ce point y s'appelle projeté de x sur C et sera noté $y = p_C(x)$.

ii Le point y est caractérisé par la propriété :

$$\forall z \in C, \langle x - y, z - y\rangle \leq 0. \tag{6}$$

Preuve : la preuve de l'unicité est la même qu'en dimension finie. Pour l'existence, il faut travailler un peu plus finement. Voir par exemple [3] pour la preuve.

Soit $x \in H$ et $(y_n)_{n\in\mathbb{N}}$ une suite minimisante. On va montrer que cette suite est de Cauchy, elle converge donc et par continuité de la distance, sa limite est le projeté de x sur C. Pour montrer la propriété de Cauchy, on part de l'identité du parallélogramme :

$$\forall u, v \in H,\ 2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2.$$

En choisissant $u = x - y_m$ et $v = x - y_n$, n, m étant fixés, on obtient :

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

Notons, pour $k \in \mathbb{N}$, $d_k = \|x - y_k\|$. Puisque la suite (y_k) est minimisante, on a

$$\lim_{k\to+\infty} d_k = \inf_{z\in C} \|x - z\| =: d.$$

On a :

$$\frac{y_n + y_m}{2} \in C \text{ donc } \left\|x - \frac{y_n + y_m}{2}\right\| \geq d.$$

On en déduit :

$$\left\|\frac{y_n - y_m}{2}\right\|^2 \leq \frac{1}{2}(d_m^2 + d_n^2) - d^2 \underset{n,m\to+\infty}{\longrightarrow} 0.$$

On obtient la convergence vers y qui est donc le projeté de x sur C.

Remarque 2.4.6 Cas des espaces de Banach

Dans un espace de Banach réflexif, on peut montrer l'existence du projeté grâce à la compacité faible-* des boules. On perd en revanche l'unicité. Dans un espace de Banach quelconque, on n'a pas nécessairement existence du projeté, comme le montre le contre-exemple ci-dessous.

Exemple 2.4.7 Un espace de Banach sans existence du projeté

On considère $E = \{f \in C^0([0, 1], \mathbb{R}), f(0) = 0\}$, muni de la norme $\|.\|_\infty$. Il s'agit bien d'un espace de Banach (hyperplan fermé dans C^0). Considérons le sous-espace

$$C = \{f \in E, \int_0^1 f = 0\}.$$

Figure 6: Un espace de Banach sans théorème de projection

L'intégration de 0 à 1 étant une opération continue pour la norme considérée, C est un hyperplan fermé de E. Cependant, la fonction $f_0 = id_{[0,1]}$ n'a pas de projeté sur C. En effet, on montre aisément que

$$\forall f \in C, \|f - f_0\| > \frac{1}{2}.$$

D'autre part, on construit comme sur la figure 5 une suite de fonctions f_n affines par morceaux continues, vérifiant

$$\|f_n - f_0\| \leq \frac{1}{2} + \frac{1}{n}.$$

Ainsi,

$$\inf_{f \in C} \|f - f_0\| = \frac{1}{2}$$

n'est pas atteint. (Notons que le projeté naturel de f_0 serait la fonction $f = f_0 - \frac{1}{2}$, en pointillés sur la figure, mais elle ne vérifie pas la contrainte $f(0) = 0$ et donc n'est pas dans l'espace E.)

Ce théorème de projection, souvent absent des livres de géométrie, a de nombreuses applications en analyse. Nous en citons ici quelques exemples.

2.4.4 Application 1 : optimisation sous contrainte convexe

L'optimisation est un domaine omniprésent dans les applications, et beaucoup de problèmes conduisent à minimiser une fonction sur un convexe fermé non vide K de $\mathbb{R}^N$. On a alors la caractérisation suivante :

Proposition 2.4.8. Minimisation sous contrainte convexe

Soit K un convexe fermé non vide de $\mathbb{R}^N$ et soit $f : K \to \mathbb{R}$ différentiable. Si $x_0 \in K$ est un point de minimum de f sur K, alors pour tout $\rho > 0$, on a

$$x_0 = p_K(x_0 - \rho \nabla f(x_0)).$$

Si $K = \mathbb{R}^N$, on retrouve la caractérisation usuelle : $\nabla f(x_0) = 0$. Dans le cas d'un convexe non trivial, cette caractérisation est à la base de l'algorithme du gradient avec projection pour la recherche du point de minimum, donné par la formule :

$$u_0 \in K, u_{n+1} = p_K(u_n - \rho \nabla f(x_n)).$$

Pour démontrer la proposition, on prend $x \in K$ et on note $x_t = tx + (1-t)x_0$, $t \in [0,1]$. Alors pour tout t, $x_t \in K$, et pour t assez petit on a $f(x_t) \geq f(x_0)$. La formule de Taylor donne alors

$$\langle \nabla f(x_0), x - x_0 \rangle \geq 0,$$

ou encore

$$\langle x_0 - \rho \nabla f(x_0) - x_0, x - x_0 \rangle \leq 0,$$

ce qui est la caractérisation du projeté.

2.4.5 Application 2 : de Brouwer à Schauder *

On admet ici le théorème de point fixe de Brouwer (1910, Luitzen Jan Egbertus Brouwer).

Théorème 2.4.9. Brouwer

Soit B la boule unité fermée de $\mathbb{R}^N$. Alors toute application continue $f : B \to B$ a un point fixe.

Remarque 2.4.10

i De ce théorème, on déduit aisément que la même propriété est valable en remplaçant la boule par un compact homéomorphe à la boule, par exemple un compact convexe non vide.

ii Rappelons que le théorème de Brouwer est faux en dimension infinie, comme le montre le contre-exemple ci-dessous :

$$f: \begin{array}{ccc} B(0,1) \subset l^2(\mathbb{N}) & \to & l^2(\mathbb{N}) \\ x = (x_n)_{n\in\mathbb{N}} & \mapsto & f(x) = (\sqrt{1-\|x\|^2}, x_0, x_1, \ldots). \end{array}$$

Cependant, on va à partir du théorème de projection, obtenir le théorème de point fixe de Schauder (1930).

Théorème 2.4.11. Schauder

i Soit C un convexe compact non vide d'un espace vectoriel normé E, de dimension finie ou infinie. Alors toute application continue $f : C \to C$ a un point fixe.

ii Soit C un convexe fermé non vide d'un espace vectoriel normé E. Alors toute application continue $f : C \to C$, d'image relativement compacte, a un point fixe.

Le point 1 est plus simple d'énoncé, cependant le point 2 (qui implique clairement le 1) est le plus utile pour les applications. En effet, en dimension infinie, imposer à un convexe d'être compact est une hypothèse forte. (En particulier, en dimension infinie, un compact est nécessairement d'intérieur vide, sinon il contiendrait une boule.) On démontre donc directement le 2.

La preuve de ce théorème peut faire l'objet d'un développement, qui se placera utilement dans de nombreuses leçons. Il est alors conseillé de se placer dans le cas hilbertien, où la preuve est moins technique. Ne pas hésiter à l'illustrer par quelques dessins.

Preuve dans le cas hilbertien : (voir [5]). Elle consiste en deux étapes :

- on commence par le cas où E est de dimension finie,
- dans le cas de la dimension infinie, on se ramène à la dimension finie par compacité, et on construit des applications f_n sur un convexe C_n de dimension finie,

Figure 7: Preuve du théorème de Schauder - Étape 1.

- enfin, on passe à la limite pour montrer l'existence d'un point fixe pour f.

- **Étape 1 :** cas où E est de dimension finie.

  Puisque $\overline{f(C)}$ est compact, il est inclus dans une boule fermée B, elle-même compacte. On définit l'application
  $$g = f \circ p_C \ : B \to f(C) \subset B.$$
  (Voir Figure 7.) Alors g est continue, donc d'après le théorème de Brouwer elle admet un point fixe $x = g(x) \in B$. Or $x = f \circ p_C(x) \in f(C) \subset C$ donc $x \in C$ et $x = f(x)$.

- **Étape 2 :** cas où E est de dimension infinie.

  On va se ramener à la dimension finie, en considérant des intersections de C avec des sous-espaces de dimension finie, de plus en plus gros.

  Fixons $n \geq 1$ et recouvrons $\overline{f(C)}$ par des boules ouvertes de rayon $\frac{1}{n}$. Par compacité, on peut en extraire un recouvrement fini :
  $$\exists k \in \mathbb{N}^*, \exists x_1, ..., x_k \in C, \quad \overline{f(C)} \subset \cup_{j=1}^{k} B\left(x_j, \frac{1}{n}\right).$$
  (Pour alléger l'écriture, on ne fait pas apparaître la dépendance de k et des x_i par rapport à l'entier n fixé.)

  Considérons le sous-espace affine (de dimension finie) E_n engendré par les $x_j, 1 \leq j \leq k$ et notons $C_n = C \cap E_n$. C'est un convexe fermé. Cependant, il n'a pas de raison d'être stable par f et on ne peut lui appliquer directement le 1. Il faut modifier la fonction f en composant à gauche par une application qui ramène dans C_n.

  On note p_n, la projection sur C_n. Alors l'application $f_n = p_n \circ f : C_n \to C_n$ est continue, d'image relativement compacte, donc par l'étape 1, elle a un point fixe $a_n = p_n(f(a_n)) \in C_n$.

- **Étape 3 :** passage à la limite.

  Il reste à faire tendre n vers l'infini et montrer que la suite (a_n), à une sous-suite près, converge vers un point fixe de f.

  La suite $(f(a_n))$ reste dans le compact $\overline{f(C)}$ donc a une sous-suite qui converge vers $l \in C$. On a, pour tout n,
  $$\|f(a_n) - a_n\| = \|f(a_n) - p_n \circ f(a_n)\| = d(f(a_n), C_n).$$
  Or
  $$f(a_n) \in f(C) \subset \cup_{j=1}^{k} B\left(x_j, \frac{1}{n}\right).$$
  Ainsi,
  $$d(f(a_n), C_n) \leq \min_{1 \leq j \leq k} \|f(a_n) - x_j\| \leq \frac{1}{n}.$$

Figure 8: Preuve du théorème de Schauder - Étape 3.

(Voir figure 8.) On en déduit que
$$\|f(a_n) - a_n\| \underset{n\to+\infty}{\longrightarrow} 0,$$
donc $a_n \underset{n\to+\infty}{\longrightarrow} l$. De plus, f étant continue, on a $f(a_n) \underset{n\to+\infty}{\longrightarrow} f(l)$ donc $f(l) = l$.

Remarque 2.4.12 Cas d'un evn.
Attention, ici on ne peut plus utiliser la projection ! On va donc définir à la main une application p_n, continue, qui envoie C dans C_n et vérifie
$$\forall x \in f(C), \|p_n(x) - x\| \leq \frac{1}{n}.$$
On construit p_n comme combinaison convexe des x_i, en posant
$$p_n(x) = \frac{\sum_{i=1}^k \phi_{i,n}(x)x_i}{\sum_{i=1}^k \phi_{i,n}(x)}, \text{ avec } \phi_{i,n}(x) = \sup\left(\frac{1}{n} - \|x - x_i\|, 0\right).$$
Le reste de la démonstration est identique au cas hilbertien.

Comme exemple d'utilisation de ce théorème, on peut retrouver le théorème de Cauchy-Peano sans passer par les solutions approchées par la méthode d'Euler. On se ramène à un argument de point fixe, comme dans la preuve classique du théorème de Cauchy-Lipschitz, en transformant le problème de Cauchy en équation intégrale. La compacité découle du théorème d'Arzela-Ascoli, comme dans la preuve classique du théorème de Cauchy-Peano.

Corollaire 2.4.13. (Cauchy Peano)

Soit E un espace vectoriel de dimension *finie*, et $f : E \to E$ une application continue. Alors, pour tous $t_0 \in \mathbb{R}, y_0 \in E$, le problème de Cauchy
$$y' = f(y), \quad y(t_0) = y_0 \tag{7}$$
admet une solution (non nécessairement unique ici !) sur un voisinage de t_0.

Idée de la preuve : y est solution locale du problème de Cauchy (7) est équivalent à
$$y \in C^0([t_0 - \alpha, t_0 + \alpha]), \forall t, y(t) = y_0 + \int_{t_0}^t f(y(s))ds,$$
pour un $\alpha > 0$ assez petit. On se ramène à un problème de point fixe pour l'application F définie sur $X = C^0([t_0 - \alpha, t_0 + \alpha])$ par
$$F(y)(t) = y_0 + \int_{t_0}^t f(y(s))ds.$$

On choisit pour convexe

$$C = \{y \in X, \|y(t) - y_0\| \leq r\},$$

pour un réel $r > 0$ à définir de façon à avoir la compacité de $\overline{F(C)}$. (Remarquons que C lui-même n'est pas compact.)

2.4.6 Application 3 : le théorème de Stampacchia *

Ce théorème généralise le théorème bien connu de Lax-Milgram. Les rapports du jury d'agreg le citent comme un développement classique, que les candidats ne savent souvent pas bien motiver. Attention donc à être en mesure d'en donner quelques applications.

Théorème 2.4.14. Stampacchia

Soit H un espace de Hilbert, K un convexe fermé non vide. Soit $a : H \times H \to \mathbb{R}$ une forme bilinéaire continue et coercive, i.e.

$$\exists \alpha, C > 0, \quad \forall u, v \in H, |a(u,v)| \leq C\|u\|\|v\|, \quad a(v,v) \geq \alpha \|v\|^2.$$

Alors, pour toute forme linéaire $\phi \in H'$, il existe un unique $u \in K$ tel que

$$\forall v \in K,\ a(u, v - u) \geq \langle \phi, v - u \rangle_{H',H}. \tag{8}$$

Remarque 2.4.15 Stampacchia vs Lax-Milgram

Sous les mêmes hypothèses sur a, le théorème de Lax-Milgram permet d'affirmer que, pour ϕ donnée, il existe un unique $w \in H$ vérifiant :

$$\forall v \in H,\ a(w, v) = \langle \phi, v \rangle_{H',H}.$$

Ici, on veut imposer une contrainte supplémentaire à la solution qui est de rester dans le convexe K. Si c'est le cas pour w, alors bien sûr $w = u$ et l'inégalité (8) devient en fait une égalité. Mais dans le cas contraire, u sera un point différent de w et on perdra donc l'égalité. Par exemple, si la forme a est tout simplement le produit scalaire, alors u sera le projeté de w sur K.

Notons également que ce théorème est souvent appliqué dans le cas où K est un sous-espace affine. Dans ce cas, l'inégalité devient une égalité. Voir par exemple l'application : EDP avec données non homogènes.

Preuve : (voir par exemple [3]).

On note $\langle \cdot, \cdot \rangle$ le produit scalaire dans H. D'après le théorème de Riesz-Fréchet, il existe un unique $f \in H$ tel que

$$\forall v \in H,\ \langle \phi, v \rangle_{H',H} = \langle f, v \rangle.$$

De plus, à u fixé, $v \mapsto a(u, v)$ est linéaire continue donc il existe un unique élément de H, noté Au, tel que

$$\forall v \in H, a(u, v) = \langle Au, v \rangle.$$

L'opérateur A est clairement linéaire continu de H dans lui-même. Le problème se ramène donc à trouver $u \in K$ tel que

$$\langle Au - f, v - u \rangle \geq 0, \forall v \in K,$$

où de façon équivalente, en fixant $\rho > 0$,

$$\langle \rho f - \rho Au + u - u, v - u\rangle \leq 0, \forall v \in K,$$

ce qui signifie précisément que u est le projeté sur K de $\rho f - \rho Au + u$.

Pour montrer l'existence et l'unicité de u, il suffit donc de vérifier que, pour un $\rho > 0$ bien choisi, l'application S de K dans lui-même définie par

$$S(v) = p_K(v + \rho f - \rho Av)$$

est contractante. Soient donc v_1 et $v_2 \in H$, on a :

$$\begin{array}{rcl}\|Sv_1 - Sv_2\|^2 & \leq & \|v_1 - v_2 - \rho A(v_1 - v_2)\|^2 \\ & \leq & \|v_1 - v_2\|^2 + \rho^2 C^2 \|v_1 - v_2\|^2 - 2\rho\langle A(v_1 - v_2), v_1 - v_2\rangle \\ & \leq & (1 + \rho^2 C^2 - 2\rho\alpha)\|v_1 - v_2\|^2.\end{array}$$

avec

$$\forall \rho \in \left]0, \frac{2\alpha}{C^2}\right[, 1 + \rho^2 C^2 - 2\rho\alpha < 1.$$

Exemple 2.4.16 EDO avec donnée non homogène

Une application typique du théorème de Lax-Milgram concerne les EDO ou EDP elliptiques avec données homogènes au bord. Par exemple, on considère le problème

$$\left\{\begin{array}{rcll} -u'' + u & = & f, & x \in I =]0, 1[, \\ u(0) & = & 0, & \\ u(1) & = & 0, & \end{array}\right.$$

où $f \in L^2(I)$. L'existence et l'unicité d'une solution faible découlent du théorème de Lax-Milgram dans l'espace de Hilbert $H_0^1(I)$, pour les applications

$$a(u, v) = \int_I (u'v' + uv), \quad \phi(v) = \int_I fv.$$

Considèrons maintenant le problème

$$\left\{\begin{array}{rcll} -u'' + u & = & f, & x \in I =]0, 1[, \\ u(0) & = & \alpha, & \\ u(1) & = & \beta, & \end{array}\right.$$

où $f \in L^2(I)$ et α, β sont des réels donnés. L'espace des fonctions H^1 vérifiant la condition au bord n'est plus un espace de Hilbert, mais un sous-espace affine de l'espace $H = H^1(I)$. L'existence et l'unicité d'une solution faible découlent du théorème de Stampacchia dans l'espace de Hilbert H, pour les applications a et ϕ définies comme ci-dessus, et le convexe

$$K = \{v \in H, v(0) = \alpha, v(1) = \beta\}.$$

Remarque 2.4.17 Minimisation d'une énergie
Soient H, K, a, ϕ comme dans le théorème de Stampacchia. On suppose de plus que a est symétrique. On peut alors lui associer une énergie définie par

$$J(v) = \frac{1}{2}a(v,v) - \langle \phi, v\rangle_{H',H}.$$

Alors on a équivalence entre les problèmes

$$\begin{cases} u \in K, \\ \forall v \in K,\ a(u, v-u) \geq \langle \phi, v-u\rangle_{H',H}. \end{cases}$$

et

$$\begin{cases} u \in K, \\ \forall v \in K,\ J(v) \geq J(u). \end{cases}$$

Ainsi, dans ce cadre, le théorème de Stampacchia peut être réinterprété comme un théorème d'existence et d'unicité pour un problème de minimisation de l'énergie, avec contrainte convexe $u \in K$.

Exemple 2.4.18 Problème de l'obstacle
Un exemple typique d'utilisation concerne le problème de l'obstacle en élasticité linéaire, qui consiste à trouver la forme que prend un élastique fixé en deux points $(0,0)$ et $(1,0)$ et astreint à rester au-dessus d'un obstacle dont le bord supérieur a pour équation

$$y = \psi(x), x \in [0,1].$$

Si $y = u(x)$ est l'équation de la forme prise par l'élastique, on se ramène au problème de minimiser l'énergie élastique

$$J(v) = \int_0^1 v'(x)^2 dx$$

sous la contrainte

$$v \in H_0^1(0,1), \quad v \geq \psi\ p.p.$$

Le théorème de Stampacchia fournit ainsi une solution unique à ce problème.

Figure 9: Le problème de l'obstacle

2.5 Hyperplans d'appui et théorèmes de séparation

Désormais, E est un espace vectoriel normé. (Le lecteur pourra se limiter à la dimension finie ou au cadre hilbertien pour plus de facilité.)

Définition 2.5.1. Hyperplan affine

On dit que H est un hyperplan affine (resp. un hyperplan affine fermé) de E s'il existe une forme linéaire (resp. une forme linéaire continue) $f : E \to \mathbb{R}$ et un réel α tels que

$$H = \{x \in E, f(x) = \alpha\} = H_{f,\alpha}.$$

Définition 2.5.2. Séparation large

Soit $H = H_{f,\alpha}$ un hyperplan affine fermé de E. On dit que H sépare deux sous-ensembles A et B au sens large si, quitte à remplacer (f, α) par $(-f, -\alpha)$, on a

$$\forall (a, b) \in A \times B, f(a) \leq \alpha \leq f(b),$$

ou, de manière équivalente,

$$\sup_{a \in A} f(a) \leq \alpha \leq \inf_{b \in B} f(b).$$

Définition 2.5.3. Séparation stricte

Soit H un hyperplan affine fermé de E. On dit que H sépare deux sous-ensembles A et B au sens strict si, quitte à remplacer (f, α) par $(-f, -\alpha)$, on a

$$\exists \epsilon > 0, \quad \forall (a, b) \in A \times B, f(a) \leq \alpha - \epsilon < \alpha + \epsilon \leq f(b).$$

ou, de manière équivalente,

$$\sup_{a \in A} f(a) < \alpha < \inf_{b \in B} f(b).$$

Exercice : vérifier les équivalences dans les deux définitions ci-dessus.

2.5.1 Hyperplan d'appui

Définition 2.5.4. Hyperplan d'appui

- Soit C un convexe de E et H un hyperplan affine fermé de E. On dit que H est un hyperplan d'appui à C si C est inclus dans un demi-espace de frontière H, et si H rencontre le bord de C.
- Si de plus $x_0 \in H \cap \partial C$, on dit que H est un hyperplan d'appui à C en x_0.

Proposition 2.5.5. Existence d'un hyperplan d'appui en tout point du bord (dimension finie)

Soit C un convexe non vide de $\mathbb{R}^N$, $C \neq \mathbb{R}^N$. Alors, pour tout $x \in \partial C$, il existe un hyperplan d'appui à C en x.

Preuve :
On s'appuie sur le lemme 2.4.4 démontré page 15.

- Si C est borné : alors il existe $y \notin C$ tel que $x = p_C(y)$. L'hyperplan passant par x et orthogonal à (xy) convient.
- Si C n'est pas borné, il existe un hyperplan d'appui H en x à $C \cap B(x, 1)$. Soit H^- le demi-espace fermé de frontière H, contenant $C \cap B(x, 1)$, et H^+ l'autre demi-espace. S'il existe $z \in C \setminus H^-$, alors $[x, z] \subset C$. Or $]x, z] \subset H^+$. Pour $y \in [x, z] \cap B(x, 1)$ on a une contradiction. Ainsi, H est un hyperplan d'appui à C.

Remarque 2.5.6 Cas de la dimension infinie
Ce résultat est faux. Contre-exemple : un hyperplan dense (noyau d'une forme linéaire non continue) est convexe mais n'admet pas d'hyperplan d'appui en un point du bord. Sous l'hypothèse C non dense, le résultat sera vrai dans un Banach réflexif (en utilisant la compacité faible-* des boules)...

2.5.2 Théorèmes de séparation : le cas de la dimension finie

Théorème 2.5.7. Théorèmes de séparation : dimension finie

Soient A et B deux convexes disjoints non vides de $\mathbb{R}^N$.

i Si A est fermé et B compact, alors il existe un hyperplan fermé qui sépare A et B au sens strict.

ii Sans aucune hypothèse topologique, il existe un hyperplan fermé qui sépare A et B au sens large.

Remarque 2.5.8 Cas de deux convexes fermés
Si A et B sont deux convexes fermés disjoints, alors il n'y a pas en général d'hyperplan qui les sépare au sens strict. Un contre-exemple est fourni par l'exemple ci-dessous, illustré par la figure 10.

$$A = \{(x, y) \in (\mathbb{R}^+)^2, xy \geq 1\}, \quad B = \mathbb{R}\times]-\infty, 0].$$

Pour la preuve du théorème, voir par exemple [10]. On peut aussi trouver la preuve du seul point 1 dans [7].

Figure 10: Un contre-exemple au théorème de séparation stricte.

On commence par le cas où A est fermé et où $B = \{x_0\}$ est un singleton, avec $x_0 \notin A$.

Lemme 2.5.9.

Soit A un convexe fermé non vide, et $B = \{x_0\}$, avec $x_0 \notin A$. Alors on peut séparer A et B au sens strict par un hyperplan fermé.

Preuve : le convexe A étant fermé et non vide, on peut définir la projection p_A sur A. Alors le vecteur $s = x_0 - p_A(x_0)$ vérifie

$$\langle s, x_0 \rangle > \sup_{y \in A} \langle s, y \rangle.$$

Le lemme est donc démontré.

Prenons maintenant A fermé et B compact. On note

$$A - B = \{a - b, a \in A, b \in B\}.$$

C'est un convexe (somme de deux convexes), fermé (voir lemme ci-après) non vide, et qui ne contient pas 0 puisque A et B sont disjoints. D'après le lemme ci-dessus, il existe $s \in E$ tel que

$$\langle s, 0 \rangle = 0 > \sup_{(a,b) \in A \times B} \langle s, a - b \rangle.$$

On a clairement,

$$\sup_{(a,b) \in A \times B} \langle s, a - b \rangle \leq \sup_{a \in A} \langle s, a \rangle - \inf_{b \in B} \langle s, b \rangle.$$

On a donc

$$\sup_{(a,b) \in A \times B} \langle s, a - b \rangle = \sup_{a \in A} \langle s, a \rangle - \inf_{b \in B} \langle s, b \rangle < 0.$$

Le point 1 est démontré.

Pour le point 2, on utilise de nouveau la remarque suivante : séparer A et B est équivalent à séparer $A - B$ et $\{0\}$.

- Si $A - B$ est fermé, alors on peut séparer $A - B$ et $\{0\}$ au sens strict par un hyperplan fermé.
- Sinon, deux cas : soit 0 est sur le bord de $A - B$. Alors il existe un hyperplan d'appui à $A - B$ passant par 0 (voir la proposition 2.5.5), et cet hyperplan sépare $A - B$ et $\{0\}$ au sens large. Soit 0 n'est pas sur le bord de $A - B$. Dans ce cas, on peut séparer l'adhérence de $A - B$ et $\{0\}$ au sens strict par un hyperplan fermé, et donc on sépare a fortiori $A - B$ et $\{0\}$.

On a utilisé dans la preuve le lemme ci-dessous, dont la preuve est laissée en exercice.

Lemme 2.5.10. Somme d'un fermé et d'un compact

Si A est fermé et B compact, alors $A + B$ est fermé.

Remarquons que le résultat est faux si l'hypothèse B compact est remplacée par B fermé. Contre-exemple :

$$A = \{(x, y) \in (\mathbb{R}^+)^2, xy \geq 1\}, \quad B = \mathbb{R} \times \{0\}, \quad A + B = \mathbb{R} \times]0, +\infty[.$$

2.5.3 Théorème de séparation : cas des evn*

Théorème 2.5.11. Théorèmes de séparation

Soit E un evn. Soient A et B deux convexes disjoints non vides de E.

i Si A est ouvert, alors il existe un hyperplan fermé qui sépare A et B au sens large.

ii Si A est fermé et B compact, alors il existe un hyperplan fermé qui sépare A et B au sens strict.

Ici, la preuve repose sur le théorème de Hahn-Banach, qui est un résultat difficile. (Il repose sur le lemme de Zorn, voir par exemple [4].)

Théorème 2.5.12. Hahn-Banach

Soient E un $\mathbb{R}$-ev, et $p : E \to \mathbb{R}$ une application positivement homogène et sous-additive. Soit S un sous-espace vectoriel de E, et soit $f : S \to \mathbb{R}$ une application linéaire vérifiant :

$$\forall x \in S, f(x) \leq p(x).$$

Alors il existe une forme linéaire F sur E telle que $F_{|S} = f$ et

$$\forall x \in E, F(x) \leq p(x).$$

Voyons comment de ce théorème on déduit les résultats de séparation. Comme dans le cas euclidien, on va commencer par le cas où B est un singleton, puis dans le cas général on s'y ramènera en faisant des sommes de Minkowski.

Étape 1. Cas où A est ouvert et $B = \{x_0\}$ est un singleton.
Si A contient le point 0, il suffit dans ce cas d'appliquer le théorème de Hahn Banach avec

$$S = \mathbb{R}x_0, \quad f(tx_0) = t, \quad p = \rho_A$$

où ρ_A désigne la jauge du convexe A. On a vu au paragraphe 2.3 qu'elle vérifie les propriétés demandées.

On obtient par le théorème l'existence d'une forme linéaire F qui prolonge f et vérifie

$$\forall x \in E, F(x) \leq \rho_A(x).$$

En particulier, on a

$$\forall x \in A, F(x) \leq \rho_A(x) \leq 1 = F(x_0).$$

Ainsi, l'hyperplan d'équation $F(x) = 1$ sépare A et B au sens large. Il reste à vérifier que cet hyperplan est fermé, i.e. que F est continue. Cela découle de la propriété de Lipschitz pour la jauge :

$$\forall x \in E, |F(x)| \leq \max(p(x), p(-x)) \leq M\|x\|.$$

Dans le cas où $0 \notin A$, le résultat de séparation reste valable, il suffit d'effectuer une translation pour se ramener au cas précédent.

Étape 2. Séparation large quand A est ouvert.
On pose $C = A - B$, alors C est convexe. De plus, il est ouvert comme réunion d'ouverts :

$$C = \cup_{b \in B}(A - b),$$

et on a $0 \notin C$. D'après l'étape 1, on peut donc séparer le singleton $\{0\}$ et C au sens large par un hyperplan fermé, i.e.

$$\exists f \in E', \forall (a,b) \in A \times B, 0 \leq f(a) - f(b).$$

On pose $\alpha = \sup_{a \in A} f(a)$, alors l'hyperplan d'équation $f(x) = \alpha$ sépare A et B au sens large.

Étape 3. Séparation stricte pour A fermé, B compact.
On pose, pour $\epsilon > 0$,

$$A_\epsilon = A + B(0,\epsilon), B_\epsilon = B + B(0,\epsilon).$$

Comme précédemment, A_ϵ et B_ϵ sont des ouverts de E. De plus, ils sont convexes, non vides, et disjoints si ϵ est assez petit (en effet, la distance d'un fermé à un compact disjoint est strictement positive). Alors par l'étape 2, on peut séparer A_ϵ et B_ϵ au sens large par un hyperplan fermé H. Ainsi, H sépare A et B strictement.

2.5.4 Vrai ou faux ?

i VRAI OU FAUX ? On peut toujours séparer (au sens large) deux convexes fermés disjoints non vides par un hyperplan fermé.

ii VRAI OU FAUX ? Le seul convexe dense est l'espace entier.

iii VRAI OU FAUX ? Soit C un convexe compact non vide de E evn. Alors pour toute $l \in E'$, C admet un hyperplan d'appui parallèle à $Ker(l)$.

iv VRAI OU FAUX ? Soit C un convexe fermé borné non vide de E evn. Alors pour toute $l \in E'$, C admet un hyperplan d'appui parallèle à $Ker(l)$.

Les solutions :

i C'est FAUX en dimension infinie, mais VRAI en dimension finie, d'après le point 2 du théorème 2.5.7. En dimension finie, on peut aussi le démontrer à partir du point 1, voir [7]. L'idée est d'intersecter l'un des convexes avec des boules de plus en plus grosses pour se ramener à un convexe compact, et appliquer le théorème de séparation stricte. Puis en passant à la limite quand le rayon des boules tend vers l'infini, on obtient la séparation large.

Pour un contre-exemple en dimension infinie, voir Bourbaki, EVT, II p.83 exercice 10. (C'est assez tordu.)

ii C'est VRAI en dimension finie comme corollaire du point précédent. C'est bien sûr FAUX en dimension infinie, il suffit de prendre pour cela le noyau d'une forme linéaire non continue. Cf le lemme 2.5.13 rappelé ci-dessous.

iii VRAI. En effet, l est continue sur le compact C, donc elle est bornée et atteint ses bornes. Ainsi,

$$\exists x_1, x_2 \in C, \forall y \in C, l(x_1) \leq l(y) \leq l(x_2).$$

Les hyperplans d'équations $l(y) = l(x_1)$ et $l(y) = l(x_2)$ sont donc des hyperplans d'appui à C parallèles à $Ker(l)$.

iv FAUX en dimension infinie. Pour un contre-exemple, on prend E comme au point ii, et on choisit comme convexe la boule fermée B. On définit une forme linéaire l par

$$\forall u \in E,\ l(u) = \sum_{n \in \mathbb{N}} 2^{-n} u_n.$$

Elle vérifie

$$\forall u \in E,\ |l(u)| \leq \|u\|_\infty \sum_{n \in \mathbb{N}} 2^{-n} = 2\|u\|_\infty.$$

Ainsi, on a :

$$\forall u \in C, -2 \leq l(u) \leq 2.$$

On va montrer que $\{l(u), u \in C\} =]-2, 2[$, de sorte que l'inf et le sup ne sont pas atteints. Tout d'abord, il n'existe pas $u \in C$ telle que $l(u) = 2$, car cela imposerait $\forall n \in \mathbb{N}, u_n = 1$ ce qui contredit l'appartenance à E. Même chose pour $l(u) = -2$. D'autre part, on peut construire une suite minimisante

$$(u^p)_{p \in \mathbb{N}} \in E^{\mathbb{N}}, l(u^p) \underset{p \to +\infty}{\longrightarrow} 2.$$

Il suffit pour cela de poser

$$u_n^p = \begin{cases} -1, & \text{si } n \leq p, \\ 0, & \text{sinon.} \end{cases}$$

Considérons les hyperplans parallèles à $Ker(l)$, i.e. les

$$H_\alpha = \{x \in E, l(x) = \alpha\}, \alpha \in \mathbb{R}.$$

Ils sont donc de deux types :

- si $-2 < \alpha < 2$, alors C rencontre les deux demi-espaces ouverts de frontière H_α, donc H_α ne peut être un hyperplan d'appui à C,
- si $|\alpha| \geq 2$, alors C ne rencontre pas H_α donc H_α ne peut être un hyperplan d'appui à C.

Lemme 2.5.13. Noyau d'une forme linéaire

Soit E un evn et l une forme linéaire sur E. Alors son noyau est soit fermé, soit dense.

Voir par exemple [4] pour une démonstration.

2.5.5 Applications *

Corollaire 2.5.14.

Soit C un convexe non vide et non dense dans E evn. Alors $\bar{C}$ est l'intersection des demi-espaces fermés contenant C.

Preuve : si $x \notin \bar{C}$ alors on peut séparer x et $\bar{C}$ au sens strict par un hyperplan. Ainsi, le demi-espace fermé contenant C, et de frontière cet hyperplan, ne contient pas x.

Remarque 2.5.15

- Ce corollaire permet notamment de montrer qu'un convexe fermé d'un evn E est fermé pour la topologie faible.
- Le corollaire a un pendant analytique : toute fonction f convexe et semi-continue inférieurement (i.e. d'épigraphe fermé) est égale au suprêmum des fonctions affines qui la minorent. Voir par exemple [10]. (Attention, pour montrer ce résultat, il faut montrer que $epi(f)$ est l'intersection des demi-espaces *non-verticaux* qui le contiennent.) L'intérêt de ce résultat est de fournir une stratégie pour montrer des résultats sur les fonctions convexes sci : on les montre sur les fonctions affines qui les minorent (facile) puis on passe au sup.

Corollaire 2.5.16. Caractérisation des sous-espaces denses

Soit $F \subset E$ un sous-espace vectoriel. Si F n'est pas dense, alors il est inclus dans un hyperplan fermé, i.e

$$\exists f \in E', f \neq 0, \text{ telle que } \forall x \in F, f(x) = 0.$$

(On utilise en général la contraposée de ce résultat pour montrer qu'un sous-espace est dense.)

Preuve : supposons F non dense, et soit $x_0 \notin \bar{F}$. On peut alors séparer strictement x_0 et $\bar{F}$. Donc il existe $f \in E'$ et $\alpha \in \mathbb{R}$ tels que

$$\forall x \in \mathbb{R}, f(x) < \alpha < f(x_0).$$

Or $f(F)$ est un sous-espace vectoriel de $\mathbb{R}$ donc c'est soit $\mathbb{R}$, soit $\{0\}$. Le premier cas est exclu, donc $f(F) = \{0\}$.

Corollaire 2.5.17. Lemme de Farkas-Minkowski

Soient H un espace de Hilbert, $k \in \mathbb{N}$ et $a_1, ..., a_k \in H$. On note

$$K = \{x \in H, \langle a_i, x\rangle \leq 0, 1 \leq i \leq k\}.$$

Soit $p \in H$ vérifiant :

$$\forall x \in K, \langle p, x\rangle \geq 0. \tag{9}$$

Alors

$$\exists \lambda_1, ..., \lambda_k \geq 0, \; p = -\sum_{i=1}^{k} \lambda_i a_i.$$

Preuve (voir par exemple [1]) :
On note
$$C = \left\{ x = \sum_{i=1}^{k} \lambda_i a_i, \lambda_i \geq 0, 1 \leq i \leq k \right\}.$$
La preuve se fait en deux étapes : tout d'abord on montre que C est un cône convexe fermé. Ensuite, on suppose par l'absurde que $-p \notin C$. On peut donc séparer au sens strict $\{-p\}$ et C par un hyperplan fermé, ce qui conduit à une contradiction avec l'hypothèse (9).

Étape 1 : C est un cône convexe fermé.
Il est clair que C est un cône, au sens : $\forall x \in C, \forall t \geq 0, tx \in C$. On vérifie aisément que C est convexe. Il reste à montrer que C est fermé. Pour cela, on distingue deux cas.

- Cas où les a_i sont linéairement indépendants : soit
$$x_n = \sum_{i=1}^{k} \lambda_i^n a_i$$
une suite d'éléments de C qui converge vers une limite x. L'espace vectoriel engendré par les a_i étant fermé, il existe $\lambda_1, ..., \lambda_k$ tels que
$$x = \sum_{i=1}^{k} \lambda_i a_i.$$
Par l'indépendance linéaire des a_i, on en déduit que pour tout i, $\lambda_i = \lim_{n \to +\infty} \lambda_i^n \geq 0$ donc $x \in C$.

- Cas où les a_i sont liés : le raisonnement précédent ne s'applique plus. Par contre, on va montrer que C est une réunion finie de cônes convexes engendrés par des vecteurs linéairement indépendants. Ainsi, C sera fermé comme union finie de fermés.
Soit $(\mu_1, ..., \mu_k) \neq (0, ..., 0)$ tel que
$$\sum_{i=1}^{k} \mu_i a_i = 0.$$
On va montrer que
$$C = \cup_{1 \leq j \leq k} C_j, \qquad \text{où } C_j = \left\{ \sum_{i \neq j} \lambda_i a_i, \lambda_i \geq 0, 1 \leq i \leq k \right\}.$$
Il est clair que les C_j sont inclus dans C. Réciproquement, soit
$$x = \sum_{i=1}^{k} \lambda_i a_i \in C.$$
Quitte à remplacer $(\mu_1, ..., \mu_k)$ par son opposé, on peut supposer que l'un des μ_i au moins est strictement négatif. On pose alors
$$t = \min \left\{ \frac{-\lambda_i}{\mu_i}, \mu_i < 0 \right\} = \frac{-\lambda_j}{\mu_j}.$$

On a alors

$$x = \sum_{i=1}^{k} \lambda_i a_i = \sum_{i=1}^{k} (\lambda_i + t\mu_i) a_i = \sum_{i \neq j} (\lambda_i + t\mu_i) a_i.$$

avec pour tout $i \neq j$, $\lambda_i + t\mu_i \geq 0$ par construction. Ainsi, $x \in C_j$. On réitère le procédé sur chaque C_j jusqu'à se ramener à des cônes engendrés par des vecteurs linéairement indépendants. Ainsi, C est bien une réunion finie de cônes fermés, donc il est fermé.

Étape 2 : théorème de séparation.

Supposons maintenant que $-p \notin C$. On peut alors séparer par un hyperplan affine fermé le compact $\{-p\}$ et le convexe fermé C au sens strict. Ainsi,

$$\exists u \in H, \exists \alpha \in \mathbb{R}, \forall v \in C, \langle v, u \rangle < \alpha < \langle -p, u \rangle.$$

Puisque $0 \in C$, on a $\alpha > 0$. Par la propriété de cône, on a de plus :

$$\forall t \geq 0, \forall v \in C, \langle tv, u \rangle < \alpha,$$

en particulier, en divisant par t et en faisant tendre t vers $+\infty$,

$$\forall v \in C, \langle v, u \rangle \leq 0.$$

On en déduit que $u \in K$. Or on a montré que

$$\langle u, p \rangle < -\alpha < 0$$

ce qui contredit l'hypothèse (9). Ainsi, on a $-p \in C$ ce qui donne bien le résultat annoncé.

2.6 Points extrêmaux et applications

2.6.1 Points extrêmaux

Définition 2.6.1. Points extrêmaux

Soit C un convexe de E, et $x \in C$. x est dit extrêmal si $C \setminus \{x\}$ est convexe, ou de façon équivalente :

$$\nexists (y_1, y_2, t) \in C \times C \times]0, 1[, \text{ tels que } x = ty_1 + (1-t)y_2.$$

On notera $Ext(C)$ l'ensemble des points extrêmaux de C.

Exemple 2.6.2 Premiers exemples

Les points extrêmaux d'un polyèdre convexe sont les sommets. Les points extrêmaux d'une boule euclidienne sont les points de la sphère. Un sous-espace affine non trivial de E n'a pas de points extrêmaux.

Théorème 2.6.3. Minkowski

Un convexe compact de $\mathbb{R}^N$ est l'enveloppe convexe de ses points extrêmaux.

Pour montrer ce théorème, on a besoin du lemme suivant.

Lemme 2.6.4.

Soit C un convexe de E et H un hyperplan d'appui à C. Alors

$$Ext(C \cap H) = Ext(C) \cap H$$

- L'inclusion $\supset$ est claire. En effet, si $u \in Ext(C) \cap H$, alors on a

$$\forall y_1, y_2 \in C, y_1 \neq y_2, \forall t \in]0,1[, ty_1 + (1-t)y_2 \neq u.$$

  C'est donc vrai a fortiori si on prend $y_1, y_2 \in C \cap H$.

- Montrons maintenant l'inclusion inverse. Soit $u \in Ext(C \cap H)$ et supposons qu'on ait :

$$\exists y_1, y_2 \in C, y_1 \neq y_2, \exists t \in]0, 1\ ty_1 + (1-t)y_2 = u.$$

  Les points y_1 et y_2 sont dans le demi-espace fermé de frontière H qui contient C. Si l'un des deux est dans le demi-espace ouvert, alors u également, ce qui contredit l'hypothèse. Ainsi, y_1 et y_2 sont dans H et donc $u \notin Ext(C \cap H)$.

Preuve du théorème : soit C un convexe compact non vide de $\mathbb{R}^N$. Quitte à se placer dans le sous-espace affine engendré par C, on suppose que C est d'intérieur non vide. On raisonne par récurrence sur la dimension N.

- Cas où $N = 0$. Alors C est un singleton, et le résultat est clair.
- Soit $N \geq 1$, fixé. Supposons le résultat vrai en dimension strictement inférieure à N. Soit C un convexe compact non vide de $\mathbb{R}^N$ et soit $x \in C$.
  - Si x est sur le bord de C : alors il existe H hyperplan d'appui à C en x. On se ramène à la dimension $N-1$ en considérant le convexe $C \cap H$. Par l'hypothèse de récurrence, x est combinaison convexe de points extrêmaux de $C \cap H$. Il reste à montrer que ces points sont aussi des points extrêmaux de C, ce qui découle du lemme.
  - Si x est intérieur à C : alors il existe $x' \in C, x' \neq x$. On remarque que la droite (xx') coupe ∂C en deux points exactement. En effet, $(xx') \cap C$ est un convexe compact de dimension 1, donc c'est un segment $[y, z]$, avec $y, z \in \partial C$. D'après le cas précédent, y et z sont combinaisons convexes de points extrêmaux de C. Ainsi, $x = ty + (1-t)z$ l'est également.

Exemple 2.6.5 Matrices bi-stochastiques

Soit $N \geq 1$, une matrice $S \in \mathbb{R}^{N \times N}$ est dite bi-stochastique si ses coefficients sont positifs ou nuls et si la somme des coefficients de chaque ligne et de chaque colonne de S vaut 1. Notons Ω_N l'ensemble de ces matrices bi-stochastiques. Alors les points extrêmaux de Ω_N sont les matrices de permutation, i.e. les matrices de la forme

$$P = [\delta_{\sigma(i),j}]_{1 \leq i,j \leq N}, \quad \sigma \in \mathcal{S}_N.$$

La preuve est laissée en exercice (voir [8], p. 200). On en déduit que Ω_N est l'enveloppe convexe de l'ensemble des matrices de permutation. (Théorème de Birkoff). Ce résultat permet d'obtenir des inégalités matricielles (voir [8]).

Remarque 2.6.6 Application du théorème
Pour minimiser une fonction concave (par exemple affine) sur un polyèdre convexe, il suffit de comparer les valeurs de la fonction aux sommets du polyèdre. Cette remarque est à l'origine des algorithmes de type simplexe en programmation linéaire.

Remarque 2.6.7 Cas de la dimension infinie
L'énoncé ci-dessus est faux (on rappelle que l'enveloppe convexe d'un compact n'est pas toujours fermée). Toutefois, il est vrai dans un evn si on remplace "enveloppe convexe" par "enveloppe convexe fermée". C'est le théorème de Krein-Milman qui a de nombreuses applications en analyse fonctionnelle. Voir par exemple [9].

2.6.2 Vrai ou faux ?

i VRAI OU FAUX ? Tout convexe compact admet au moins un point extrêmal.

ii VRAI OU FAUX ? Tout convexe fermé borné admet au moins un point extrêmal.

iii VRAI OU FAUX ? Soit C un convexe compact non vide de E evn, et soit H un hyperplan d'appui à C. Alors H contient un point extrêmal de C.

iv VRAI OU FAUX ? Soit C un convexe fermé borné non vide de E evn, et soit H un hyperplan d'appui à C. Alors H contient un point extrêmal de C.

Les solutions :

i VRAI d'après les théorèmes de Minkowski (dim finie) et Krein Milman (dim infinie).

ii FAUX en dimension infinie. Contre-exemple : soit

$$E = \left\{ u \in \mathbb{R}^{\mathbb{N}}, u_n \underset{n\to+\infty}{\longrightarrow} 0 \right\},$$

muni de la norme infini. Soit B la boule unité fermée de E. Alors B n'a pas de point extrêmal. En effet, soit $u \in B$, $N \in \mathbb{N}$ tel que

$$\forall n > N,\ |u_n| < \frac{1}{2}.$$

On construit des suites v et w en posant

$$\forall n \in \mathbb{N}, v_n = \begin{cases} u_n, & \text{si } n < N, \\ u_n + \frac{1}{2n}, & \text{sinon.} \end{cases} \qquad w_n = \begin{cases} u_n, & \text{si } n < N, \\ u_n - \frac{1}{2n}, & \text{sinon.} \end{cases}$$

Alors on a : $v, w \in B, v \neq w, u = (v + w)/2$. Donc u n'est pas extrêmal.

iii VRAI. En effet, $C \cap H$ est convexe compact non vide donc il admet un point extrêmal d'après le i. C'est donc un point extrêmal de C d'après le lemme.

iv FAUX. Pour un contre-exemple, on prend E et B comme ci-dessus et $H = \{u \in E, u_0 = 1\}$. Alors H est un hyperplan d'appui à B, et cependant B n'a pas de point extrêmal.

Par contre, le résultat est vrai en dimension finie. En effet, $C \cap H$ est un convexe compact non vide, donc il admet un point extrêmal $x \in Ext(C \cap H) = Ext(C) \cap H$. (Voir le lemme 2.6.4.)

3 Fonctions convexes

Les fonctions f considérées dans cette partie seront définies sur une partie convexe C d'un espace vectoriel normé E (le plus souvent $\mathbb{R}^N$), ou sur un ouvert contenant C, et à valeurs réelles.

3.1 Définitions et propriétés élémentaires

3.1.1 Différentes notions de convexité

Définition 3.1.1. Convexité, convexité stricte, convexité forte

Soient C un convexe non vide de E espace vectoriel normé, et $f : C \to \mathbb{R}$.

- f est dite convexe sur C si

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

- f est dite concave si $-f$ est convexe.
- f est dite strictement convexe sur C si

$$\forall x, y \in C, x \neq y, \forall t \in]0,1[, \; f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

- f est dite fortement convexe sur C s'il existe $\alpha > 0$ tel que

$$\forall x, y \in C, \forall t \in]0,1[, \; f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2}t(1-t)\|x-y\|^2.$$

On dira dans ce cas que f est α-convexe.

Remarque 3.1.2 Forte convexité

- Il apparaît clairement que f fortement convexe $\Rightarrow$ f strictement convexe $\Rightarrow$ f convexe.
- Dans le cas euclidien ou préhilbertien, on montre que f est α-convexe si et seulement si $f - \alpha/2\| \cdot \|^2$ est convexe. (Exercice) La forte convexité est une notion qui peut sembler peu naturelle. On verra que c'est un cadre très agréable pour l'optimisation, car elle donne non seulement l'existence et l'unicité d'un point de minimum, mais aussi des algorithmes de minimisation performants. Notons que ce cadre inclut en particulier les fonctions quadratiques, voir section 3.1.3.

Figure 11: Epigraphe d'une fonction convexe

Proposition 3.1.3. Caractérisation dans le cas continu

Soit f une fonction *continue* sur un convexe C. Alors on a les caractérisations suivantes.

i f est convexe sur C si et seulement si

$$\forall x, y \in C,\ f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}.$$

ii f est α-convexe sur C si et seulement si

$$\forall x, y \in C,\ f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} - \frac{\alpha}{8}\|x-y\|^2.$$

Remarquons que dans la littérature, c'est parfois cette caractérisation (ii) qui est prise comme définition de la forte convexité.

Preuve. Montrons le point 2. On obtiendra 1 en prenant $\alpha = 0$. Soit $\alpha \geq 0$, f une fonction satisfaisant la condition, et soient $t \in]0,1[, x, y \in C$. En utilisant la décomposition dyadique, on peut trouver une suite (t_n) avec $t_n = 2^{-n}p_n$, $p_n \in \{0, \cdots, n\}$, $n \geq 1$, qui converge vers t quand n tend vers $+\infty$. En décomposant t_{n+1} sous la forme

$$t_{n+1} = \frac{\tau_n + \theta_n}{2}, \theta_n = 2^{-n}q_n, \tau_n = 2^{-n}r_n,$$

on montre par récurrence que

$$f(t_n x + (1-t_n)y) \leq t_n f(x) + (1-t_n)f(y) - \frac{\alpha}{2}t_n(1-t_n)\|x-y\|^2.$$

La continuité de f permet de passer à la limite et d'obtenir le résultat annoncé.

3.1.2 Critère de l'épigraphe

Ce critère permet de faire un parallèle avec la partie géométrique du cours.

Proposition 3.1.4. Critère de l'épigraphe

La fonction f est convexe si et seulement si son épigraphe $epi(f)$ est convexe, où

$$epi(f) = \{(x,r) \in C \times \mathbb{R},\ f(x) \leq r\}.$$

La preuve est laissée en exercice.

Ce critère permet de transposer facilement en analyse des résultats de géométrie. En particulier, en utilisant la stabilité des convexes par somme et par intersection, on obtient le résultat suivant (qui peut aussi très facilement se montrer à partir de la définition).

Proposition 3.1.5. Propriétés élémentaires

Une somme de fonctions convexes est convexe. Un suprêmum de fonctions convexes est convexe.

Exemple 3.1.6 Somme des plus grandes valeurs propres

Soit E l'ensemble des matrices symétriques réelles de taille N. On associe à une matrice $A \in E$ ses valeurs propres, rangées par ordre décroissant :

$$\lambda_1(A) \geq \ldots \geq \lambda_N(A).$$

Pour tout entier $m, 1 \leq m \leq N$, on note

$$f_m(A) = \sum_{j=1}^{m} \lambda_j(A).$$

On peut montrer que f_m vérifie

$$\forall A \in E, f_m(A) = \sup\{tr(Q^T AQ), Q \in \Omega_m\},$$

où $\Omega_m = \{Q \in M_{N,m}(\mathbb{R}), Q^T Q = I_m\}$. (Voir [8], p. 202.) Ainsi, f_m est une fonction convexe, comme sup de fonctions linéaires.

Remarque 3.1.7 Fonctions convexes semi-continues inférieurement

On a vu dans la première partie du cours que dans bien des cas (théorèmes de projection, de séparation, etc.) la bonne notion est celle de convexe fermé non vide. On vient de voir que les fonctions dont l'épigraphe est convexe sont les fonctions convexes. Qu'en est-il des fonctions dont l'épigraphe est fermé ? On les appelle fonctions semi-continues inférieurement (en abrégé, s.c.i). Cette terminologie vient de la caractérisation suivante :

f est s.c.i. sur un ouvert U d'un espace métrique si et seulement si pour tout $x \in U$, pour toute suite (x_n) dans U, tendant vers x,

$$f(x) \leq \liminf_{n \to +\infty} f(x_n).$$

Une théorie très riche peut être développée pour les fonctions convexes s.c.i. Nous ne la détaillerons pas ici.

3.1.3 Exemples

A ce stade où on n'a pas d'autre outil que la définition et des caractérisations très simples, on se limitera à des exemples élémentaires pour éviter les calculs fastidieux. D'autres exemples usuels de fonctions convexes seront présentés ultérieurement.

i Une fonction affine est convexe, non strictement convexe.

ii Sur $\mathbb{R}$, la fonction valeur absolue est convexe, non strictement.

iii Toute norme est convexe, grâce à l'inégalité triangulaire. Elle n'est jamais strictement convexe, puisqu'elle est positivement homogène.

Figure 12: Boules unités associées aux normes a) $\|\cdot\|_\infty$ (cas non strictement convexe) et b) euclidienne (cas fortement convexe).

iv Si f est convexe et ϕ convexe croissante, alors $\phi \circ f$ est convexe. (Exercice)

v En particulier, les fonctions $x \mapsto |x|^\alpha$ sont convexes sur $\mathbb{R}$, pour $\alpha > 1$.

vi De iv, on déduit aussi que l'application $\|\cdot\|^2$ est convexe. Elle n'est pas en général strictement convexe, par exemple la norme $\|\cdot\|_\infty$ dans $\mathbb{R}^2$ vérifie

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

(voir figure 12, a).

vii Si une norme dérive d'un produit scalaire, alors l'application $\|\cdot\|^2$ est fortement convexe. En effet, elle est continue et vérifie l'identité du parallélogramme

$$\forall x, y \in E, \quad \left\|\frac{x+y}{2}\right\|^2 = \frac{1}{2}(\|x\|^2 + \|y\|^2) - \frac{1}{4}\|x-y\|^2$$

(voir figure 12, b).

viii Dernier exemple à ce stade : si $\phi : [a,b] \to \mathbb{R}$ est croissante, alors la fonction Φ définie sur $[a,b]$ par

$$\Phi(x) = \int_a^x \phi(u)du$$

est convexe. Nous pouvons démontrer ce résultat sans dériver Φ (si ϕ n'est pas continue, on ne peut pas affirmer que Φ est dérivable). Pour cela, on prend $x, y \in [a,b], x < y$ et $t \in]0,1[$. Notons, pour alléger, $x_t = tx + (1-t)y$. En utilisant la relation de Chasles, on a :

$$\begin{array}{rcl}\Delta & = & \Phi(x_t) - t\Phi(x) - (1-t)\Phi(y) \\ & = & -(1-t)\displaystyle\int_x^y \phi(u)du + \int_x^{x_t} \phi(u)du.\end{array}$$

Pour pouvoir comparer ces deux intégrales, on fait un changement de variables pour les ramener à l'intervalle $[0, y-x]$, en remarquant que $x_t = x + (1-t)(y-x)$. Il vient

$$\Delta = -(1-t)\int_0^{x-y} \phi(x+u)du + (1-t)\int_0^{x-y} \phi(x+(1-t)u)du \leq 0$$

grâce à la croissance de ϕ.

3.1.4 Critère des pentes croissantes

Proposition 3.1.8. Critère des pentes croissantes

On suppose ici que $E = \mathbb{R}$, et que C est un intervalle de $\mathbb{R}$. Alors f est convexe sur C si et seulement si, pour tout $x_0 \in C$, la fonction

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$

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

La démonstration ne pose pas de difficulté et est laissée en exercice.

Corollaire 3.1.9.

Soit $f :]0, +\infty[\to \mathbb{R}$, et soit g la fonction définie par

$$\forall x > 0,\ g(x) = xf\left(\frac{1}{x}\right).$$

Alors f est convexe si et seulement si g est convexe.

Preuve. Fixons $x_0 > 0$ et considérons la fonction pente s_g définie par

$$s_g(x) = \frac{g(x) - g(x_0)}{x - x_0}.$$

On définit s_f de la même façon. On a alors

$$\begin{aligned} s_g(x) &= \frac{xf(1/x) - x_0 f(1/x_0)}{x - x_0} \\ &= f(1/x_0) + \frac{x}{x - x_0}(f(1/x) - f(1/x_0)) \\ &= f(1/x_0) - \frac{1}{x_0} s_f(1/x). \end{aligned}$$

On remarque que s_g est croissante si et seulement si s_f l'est.

Exemple d'application : on verra que la fonction $-\log$ est convexe sur $]0, +\infty[$. On en déduit que $x \mapsto x \log x$ est convexe.

3.1.5 Vrai ou faux ?

i VRAI OU FAUX ? La composée de deux fonctions convexes est convexe.

ii VRAI OU FAUX ? Le produit de deux fonctions convexes est convexe.

iii VRAI OU FAUX ? Une fonction de deux variables x, y, convexe par rapport à x pour tout y, et convexe par rapport à y pour tout x, est convexe par rapport au couple (x, y).

Les réponses.

i C'est FAUX. Contre-exemple : sur $\mathbb{R}$, $f(x) = |e^x - 1|$. Il faut ajouter l'hypothèse "croissante" pour la fonction "à gauche".

ii C'est FAUX. Contre-exemple sur $\mathbb{R}$: $f=-1$, $g(x)=x^2$. Rem: c'est encore FAUX si on suppose que les deux fonctions sont positives ou nulles. Ex : $f(x)=x, g(x)=e^{-x}$ sur $[0,+\infty[$.

iii C'est encore FAUX. Contre-exemple sur $(\mathbb{R}^+)^2$: $f(x,y)=xy$. On a par exemple $f(1,1)=1>\frac{1}{2}(f((0,2)+f(2,0))$.

3.2 Caractérisations dans le cas différentiable

On considère dans cette section des fonctions définies sur un voisinage Ω d'un convexe C de $\mathbb{R}^N$. La majorité des résultats reste valables dans un Hilbert.

Soit donc $f:\Omega\to\mathbb{R}$ une fonction différentiable. On notera $\langle Df(x),h\rangle$ la différentielle de f au point x, appliquée à un vecteur $h\in\mathbb{R}^N$. On utilisera aussi la notation gradient,

$$\langle Df(x),h\rangle=\langle\nabla f(x),h\rangle.$$

On va généraliser les critères bien connus en dimension 1 :

- Une fonction dérivable est convexe si et seulement si sa dérivée est croissante.
- Une fonction dérivable est convexe si et seulement si son graphe est situé au-dessus des tangentes.

3.2.1 Caractérisation géométrique

Théorème 3.2.1. Caractérisation par l'hyperplan tangent

Soit Ω un ouvert de $\mathbb{R}^N$, $C\subset\Omega$ un convexe et $f:\Omega\to\mathbb{R}$, différentiable. On a les équivalences suivantes.

i La fonction f est convexe sur C si et seulement si

$$\forall x,x_0\in C, f(x)\geq f(x_0)+\langle Df(x_0),x-x_0\rangle,$$

i.e. le graphe de f est au-desssus de tous les hyperplans tangents.

ii La fonction f est strictement convexe sur C si et seulement si

$$\forall x,x_0\in C, x\neq x_0,\ f(x)>f(x_0)+\langle Df(x_0),x-x_0\rangle.$$

iii La fonction f est α-convexe sur C si et seulement si

$$\forall x,x_0\in C, f(x)\geq f(x_0)+\langle Df(x_0),x-x_0\rangle+\frac{\alpha}{2}\|x-x_0\|^2,$$

i.e. le graphe de f est au-dessus d'un paraboloïde tangent.

Preuve.

i Supposons d'abord f convexe. Soient $x_0,x\in C, t\in]0,1[$ et notons $x_t=tx+(1-t)x_0$. On a par hypothèse

$$f(x_t)\leq tf(x)+(1-t)f(x_0).$$

$$\frac{f(x_t)-f(x_0)}{t} \leq f(x)-f(x_0).$$

En passant à la limite $t \to 0$, on obtient le résultat.

Réciproquement, supposons la condition vérifiée et montrons que f est convexe. Soient donc $x, y \in C, t \in]0,1[$, et posons $x_t = tx + (1-t)y$. On a par hypothèse

$$\begin{array}{rcl} f(x) & \geq & f(x_t) + \langle Df(x_t), x - x_t \rangle \\ f(y) & \geq & f(x_t) + \langle Df(x_t), y - x_t \rangle \end{array}$$

En multipliant la première inégalité par t, la seconde par $1-t$ et en les additionnant, on obtient le résultat voulu.

ii Pour le point 2, la réciproque se démontre comme au point 1, en remplaçant les inégalités larges par des strictes. Par contre, le sens direct ne peut être calqué sur la preuve du point 1, car le passage à la limite $t \to 0$ ne permet pas de conserver les inégalités strictes. Cependant, si f est strictement convexe, et $x, x_0 \in C$, $x \neq x_0$, $t \in]0,1[$, on a

$$f(x_t) < tf(x) + (1-t)f(x_0).$$

On a aussi, d'après le point 1,

$$f(x_t) \geq f(x_0) + \langle Df(x_0), x_t - x_0 \rangle.$$

En combinant ces deux inégalités, et en simplifiant, on obtient l'inégalité stricte attendue.

iii Le point 3 se démontre comme le point 1, ou se déduit de celui-ci grâce à l'équivalence : f α-convexe si et seulement si $f - \frac{\alpha}{2}\|\cdot\|^2$ convexe.

3.2.2 Fonctions monotones

Définition 3.2.2. Fonctions monotones

Soient C un convexe de $\mathbb{R}^N$ (ou d'un Hilbert), $\alpha \geq 0$ et $F : C \to \mathbb{R}^N$. Alors

i F est dite monotone sur C si

$$\forall x, y \in C,\ \langle F(x) - F(y), x - y) \rangle \geq 0.$$

ii F est dite strictement monotone sur C si

$$\forall x, y \in C, x \neq y,\ \langle F(x) - F(y), x - y) \rangle > 0.$$

iii F est dite α-monotone sur C si

$$\forall x, y \in C,\ \langle F(x) - F(y), x - y) \rangle \geq \alpha \|x - y\|^2.$$

Attention à la terminologie : il faut comprendre ici “monotone” au sens de “croissante” si on pense à la dimension 1.

Théorème 3.2.3. Critère du gradient monotone

Soit f différentiable sur Ω. Soit $\alpha > 0$. Alors on a les équivalences suivantes.

i f est convexe sur $C \iff \nabla f$ est monotone sur C.

ii f est strictement convexe sur $C \iff \nabla f$ est strictement monotone sur C.

iii f est α-convexe sur $C \iff \nabla f$ est α-monotone sur C.

Preuve. On va montrer le point 3. Le 1 s'en déduira en prenant $\alpha = 0$ et le 2 se montre de manière analogue avec des inégalités strictes.

Supposons d'abord f α-convexe. Soient $x, y \in C$, alors on a d'après le théorème 3.2.1

$$\begin{aligned} f(y) &\geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|x - y\|^2 \\ f(x) &\geq f(y) + \langle \nabla f(y), x - y\rangle + \frac{\alpha}{2}\|x - y\|^2. \end{aligned}$$

En sommant ces deux inégalités, on obtient exactement l'α-monotonie de ∇f.

Réciproquement, supposons maintenant que ∇f est α-monotone. Fixons $x, y \in C$ et posons, pour $t \in [0,1]$,

$$x_t = x + t(y - x), \quad \phi(t) = f(x_t).$$

Alors ϕ est dérivable sur $[0,1]$ et on a

$$\phi'(t) = \langle \nabla f(x_t), y - x\rangle.$$

Nous allons minorer $\phi'(t)$. Grâce à notre hypothèse, on a

$$\begin{aligned} \phi'(t) - \phi'(0) &= \frac{1}{t}\langle \nabla f(x_t) - \nabla f(x), x_t - x\rangle \\ &\geq \frac{\alpha}{t}\|x_t - x\|^2 \\ &= \alpha t\|y - x\|^2. \end{aligned}$$

On veut maintenant en déduire une minoration de $\phi(1) - \phi(0)$.

Attention ! En général, on ne peut pas écrire

$$\phi(1) - \phi(0) = \int_0^1 \phi'(t)dt$$

car on ne sait pas si ϕ' est intégrable... On va contourner cette difficulté en utilisant l'inégalité des accroissements finis, sous la forme

Lemme 3.2.4. Inégalité des accroissements finis

Soient f, g deux fonctions dérivables sur $[0,1]$, vérifiant $f' \geq g'$. Alors on a :

$$f(1) - f(0) \geq g(1) - g(0).$$

Cette inégalité donne alors

$$\begin{aligned}\phi(1)-\phi(0) &\geq \phi'(0)+\int_0^1 \alpha t\|y-x\|^2 dt\\ &= \phi'(0)+\frac{\alpha}{2}\|y-x\|^2\end{aligned}$$

ce qui se réécrit

$$f(y)-f(x)\geq \langle \nabla f(x), y-x\rangle+\frac{\alpha}{2}\|y-x\|^2.$$

Le théorème 3.2.1 permet de conclure que f est α-convexe.

3.2.3 Nouveaux exemples

Le critère de la dérivée permet de montrer la convexité des fonctions ci-dessous.

Exemple 3.2.5 Fonctions d'une variable réelle

Sur $\mathbb{R}$, les fonctions $x \mapsto e^{\beta x}$, $\beta \in \mathbb{R}$, sont convexes. Sur $]0,+\infty[$, les fonctions définies par $x \mapsto x^{-\gamma}$, pour $\gamma > 0$, sont convexes. La fonction $\ln$ est concave.

Exemple 3.2.6 Fonctions quadratiques

Soitn A une matrice symétrique de taille N, $b \in \mathbb{R}^N$. On dira que A est semi-définie positive si elle vérifie

$$\langle Ax, x\rangle \geq 0, \quad \forall x \in \mathbb{R}^N,$$

autrement dit si les valeurs propres de A sont positives ou nulles. On dira que A est définie positive si elle vérifie

$$\langle Ax, x\rangle > 0, \quad \forall x \in \mathbb{R}^N, x \neq 0,$$

autrement dit si les valeurs propres de A sont positives strictement.

Soit f la fonction définie sur $\mathbb{R}^N$ par

$$f(x)=\frac{1}{2}\langle Ax, x\rangle.$$

Alors f est différentiable et vérifie

$$\forall x \in \mathbb{R}^N, \quad \nabla f(x)=Ax-b.$$

La matrice A étant symétrique, elle est diagonalisable sur $\mathbb{R}$ dans une base orthonormée. En diagonalisant, on montre que le gradient de f est monotone (resp. strictement monotone) si et seulement si les valeurs propres de A sont positives ou nulles (resp. strictement positives).

On a démontré le résultat ci-dessous :

Proposition 3.2.7. Fonctions quadratiques

Soit f la fonction définie sur $\mathbb{R}^N$ par

$$f(x) = \frac{1}{2}\langle Ax, x\rangle.$$

Alors on a les équivalences suivantes.

i la fonction f est convexe si et seulement si A est semi-définie positive,

ii la fonction f est strictement convexe si et seulement si A est définie positive. Dans ce cas, f est même α-convexe, où α est la plus petite valeur propre de A.

Une application importante de ce résultat concerne les systèmes linéaires $Ax = b$, avec A symétrique réelle définie positive. Dans ce cas, résoudre le système $Ax = b$ est équivalent à minimiser la fonction f. Cette remarque justifie les algorithmes de type gradient ou gradient conjugué pour les systèmes linéaires.

3.3 Inégalités de convexité

La définition-même de la convexité repose sur une inégalité, que l'on peut interpréter ainsi : l'image du barycentre de deux points, affectés de coefficients positifs, est inférieure au barycentre des images, affectées des mêmes coefficients. Cette formulation se généralise aisément à un nombre arbitraire de points, et même à une moyenne continue pour une mesure de probabilité : c'est le théorème de Jensen.

3.3.1 Inégalités discrètes

Le théorème ci-dessous découle immédiatement de la définition par récurrence sur k.

Théorème 3.3.1. Jensen discret

Soit f une fonction convexe définie sur C convexe de $\mathbb{R}^N$. Alors, pour tout entier $k \geq 2$, on a

$$\forall x_1, \ldots, x_k \in C,\ \forall \alpha \in \Delta_k,\ f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i).$$

Exemple 3.3.2 Inégalités classiques

$$\ln \text{ concave} \quad \Rightarrow \quad \prod_{i=1}^{k} {x_i}^{\alpha_i} \leq \sum_{i=1}^{k} \alpha_i x_i. \tag{10}$$

$$x \mapsto x\ln(x) \text{ convexe} \quad \Rightarrow \quad \left(\sum_{i=1}^{k} \alpha_i x_i\right)^{\sum_{i=1}^{k} \alpha_i x_i} \leq \prod_{i=1}^{k} {x_i}^{\alpha_i x_i}. \tag{11}$$

$$x \mapsto \frac{-1}{1+e^{-x}} \text{ convexe} \quad \Rightarrow \quad \sum_{i=1}^{k} \frac{\alpha_i}{1+x_i} \leq \left(1 + \prod_{i=1}^{k} {x_i}^{\alpha_i}\right)^{-1}. \tag{12}$$

La première inégalité, connue sous le nom d'inégalité arithmético-géométrique, permet notamment de démontrer l'inégalité de Hölder, dont découle également l'inégalité de Minkowski.

Application

Le théorème de Jensen discret intervient dans la démonstration de l'inégalité de Kantorovitch ci-dessous. (Ref: [6]). Cette inégalité intervient notamment dans l'estimation d'erreur de la méthode du gradient à pas optimal pour les fonctions quadratiques.

Exemple 3.3.3 Inégalité de Kantorovitch

Soit A une matrice symétrique définie positive. On note $\lambda_1 \geq \ldots \geq \lambda_N$ ses valeurs propres. Alors on a, pour tout $x \in \mathbb{R}^N$,

$$\|x\|^4 \leq \langle Ax, x\rangle \langle A^{-1}x, x\rangle \leq \left(\sqrt{\frac{\lambda_1}{\lambda_N}} + \sqrt{\frac{\lambda_1}{\lambda_N}}\right)^2 \|x\|^4.$$

3.3.2 Inégalité de Jensen

Théorème 3.3.4. Jensen

Soient $(X, \mathcal{M}, \mu)$ un espace de probabilité, I un intervalle non vide de $\mathbb{R}$, et soit $g : X \to I$ une fonction intégrable. Alors pour toute fonction convexe $f : I \to \mathbb{R}$, telle que $f \circ g$ soit intégrable ou positive, on a

$$f\left(\int_X g d\mu\right) \leq \int_X f \circ g d\mu.$$

L'image d'une moyenne est inférieure à la moyenne des images.

Remarque : on retrouve le théorème discret en prenant

$$X = R, \mathcal{M} = \mathcal{P}(\mathbb{R}), \mu = \frac{1}{k}\sum_{i=1}^{k} \delta_{x_i}.$$

Les applications de l'inégalité de Jensen concernent notamment la physique statistique et la théorie de l'information. Citons l'application suivante, utile en physique statistique.

Exemple 3.3.5

Soit X une variable aléatoire sur un espace de probabilité. Alors on a

$$e^{E(X)} \leq E(e^X).$$

3.4 Régularité des fonctions convexes

3.4.1 Continuité

Les fonctions convexes sont-elles nécessairement continues ? En dimension infinie, la réponse est NON, puisque même des fonctions linéaires peuvent être discontinues. En fait, la continuité des fonctions linéaires est équivalente au fait qu'elles soient bornées sur un voisinage de 0. On va voir que pour les fonctions convexes, on a également un lien fort entre la continuité et le fait d'être bornées au voisinage de tout point.

Dans cette partie, on notera pour simplifier B_ϵ la boule de centre 0 et de rayon ϵ, pour $\epsilon > 0$.

Cas des fonctions bornées

Théorème 3.4.1. Propriété de Lipschitz

Soit Ω un ouvert convexe de $\mathbb{R}^N$ et soit f convexe bornée sur l'ouvert Ω. Alors pour toute partie C incluse dans Ω, telle qu'il existe $\epsilon > 0$ avec $C + B_\epsilon \subset \Omega$, f est lipschitzienne sur C.

Preuve : soit $M > 0$ avec $\forall x \in \Omega, |f(x)| \leq M$. On va montrer que

$$\forall x, y \in C,\ |f(x) - f(y)| \leq \frac{2M}{\epsilon}\|x - y\|.$$

Soient donc $x, y \in C, x \neq y$. On pose

$$z_1 = y + \epsilon\frac{y - x}{\|y - x\|}, \quad z_2 = x - \epsilon\frac{y - x}{\|y - x\|}.$$

Alors z_1 et z_2 sont dans Ω. De plus, y s'écrit sous la forme

$$y = \theta x + (1 - \theta)y,\ \text{avec}\ \theta = \frac{\epsilon}{\|x - y\| + \epsilon}.$$

Alors, la convexité de f donne

$$\begin{aligned}
f(y) &\leq \theta f(x) + (1 - \theta)f(z_1) && (13)\\
f(y) - f(x) &\leq (1 - \theta)(f(z_1) - f(x)) && (14)\\
&\leq 2M(1 - \theta) = 2M\frac{\|y - x\|}{\|y - x\| + \epsilon} && (15)\\
&\leq \frac{2M}{\epsilon}\|y - x\|. && (16)
\end{aligned}$$

Le même raisonnement, appliqué au point z_2, donne

$$f(x) - f(y) \leq \frac{2M}{\epsilon}\|y - x\|,$$

ce qui prouve le résultat.

Remarquons que ce résultat et la preuve présentée sont valables également en dimension infinie.

Cas général

En dimension infinie, on l'a vu, on ne peut se passer de l'hypothèse "f bornée" et obtenir la continuité de toute fonction convexe sur un ouvert. En dimension finie, pourtant, on a le résultat suivant.

Théorème 3.4.2. Continuité des fonctions convexes (dimension finie)

Soit Ω un ouvert convexe de $\mathbb{R}^N$ et soit f convexe sur l'ouvert Ω. Alors f est continue sur Ω, et lipschitzienne sur tout compact.

La preuve s'appuie sur le théorème précédent, et suppose de montrer qu'une fonction convexe est localement bornée. Majorer f au voisinage d'un point de Ω va être relativement aisé, grâce à la définition de la convexité. La minoration de f est plus subtile, elle passe par le lemme suivant.

Lemme 3.4.3. Minoration des fonctions convexes

Soit Ω un ouvert convexe de $\mathbb{R}^N$ et soit f convexe sur l'ouvert Ω. Alors il existe une fonction g affine qui minore f sur Ω. De plus, pour tout $x_0 \in \Omega$, on peut choisir g telle que $g(x_0) = f(x_0)$.

Preuve : Soit $C = epi(f)$, alors C est convexe, non vide, et pour $x_0 \in \Omega$ on a $(x_0, f(x_0)) \in C$. Alors il existe un hyperplan d'appui à C en $(x_0, f(x_0))$, i.e. il existe une forme linéaire l et un réel α tels que

$$(l,\alpha) \neq (0,0), \quad \beta = l(x_0) + \alpha f(x_0) \leq l(x) + \alpha r, \forall (x,r) \in C. \tag{17}$$

On va montrer que $\alpha > 0$. En effet, supposons d'abord $\alpha < 0$. En considérant les points $(x, f(x)+n) \in C$, $n \in \mathbb{N}$, et en faisant tendre n vers $+\infty$, on contredit la relation (17). Supposons maintenant $\alpha = 0$. La relation (17) implique alors que l'application linéaire l, restreinte à Ω, atteint un minimum en x_0, et donc l est identiquement nulle. Mais ceci contredit l'affirmation $(l,\alpha) \neq (0,0)$. On a donc montré que $\alpha > 0$. En prenant $r = f(x)$ dans (17) et en divisant par α, il vient

$$f(x) \geq g(x) = -\frac{1}{\alpha}(l(x) - \beta).$$

L'application g ainsi définie vérifie les propriétés attendues.

Il reste à prouver le théorème. Soit donc Ω un ouvert convexe, et f une fonction convexe sur Ω. Soit $x_0 \in \Omega$, montrons que f est bornée au voisinage de x_0. On choisit un voisinage de la forme

$$\Delta = co(s_0, \ldots, s_N) \text{ tel que } x_0 \in \Delta \text{ et } \exists \epsilon > 0, \Delta + \bar{B}_\epsilon \subset \Omega.$$

D'après le lemme, il existe g affine qui minore f sur Δ; g est continue sur le compact Δ donc y est bornée, il en découle que f est minorée sur Δ. De plus, tous les éléments de Δ sont combinaisons convexes des sommets $s_0, \ldots, s_N$ et vérifient donc

$$\forall x \in \Delta,\ f(x) \leq \sum_{i=0}^{N} |f(s_i)|.$$

Finalement, f est bien bornée sur Δ. Du théorème 3.4.1 découle donc la continuité de f en x_0.

Par ailleurs, si K est un compact inclus dans Ω, alors f est continue donc bornée sur K. D'après le théorème 3.4.1, f est donc lipschitzienne sur K.

Remarque 3.4.4 Continuité au bord

Attention, le théorème donne la continuité sur un ouvert. Au bord, une fonction convexe n'est pas nécessairement continue.

Remarque 3.4.5 Cas de la dimension infinie

On a vu que le théorème n'est pas valable en dimension infinie. Notons que le lemme permettant de minorer f peut se généraliser aux espaces de Banach, à condition que l'épigraphe de f soit non dense (ce qui exclut par exemple les formes linéaires non continues). Par contre, la méthode de majoration de f proposée ici ne peut s'étendre à la dimension infinie.

3.4.2 Application

Corollaire 3.4.6. Suite de fonctions convexes

Soit (f_k) une suite de fonctions convexes sur $\mathbb{R}^N$, qui converge simplement vers une limite f. Alors f est convexe, et la convergence est uniforme sur tout compact.

Preuve (voir par exemple [7]). La convexité de f découle d'un simple passage à la limite dans les inégalités larges. On va montrer la convergence uniforme sur tout compact de la forme $S = \bar{B}(0, r)$, $r > 0$. La preuve se fait en trois étapes.

i On montre que la suite (f_k) est bornée indépendamment de k sur S. Pour cela, on introduit $g = \sup f_k$. Alors g est convexe sur S à valeurs réelles. On en déduit qu'elle est continue sur S, et donc
$$\exists M > 0, \forall x \in S, \forall k,\ f_k(x) \geq g(x) \geq M.$$
Pour minorer les f_k, on remarque que pour tout $x \in S$,
$$f_k(x) + f_k(-x) \geq 2f_k(0).$$
La suite $f_k(0)$ est convergente donc minorée par un réel μ. On a donc, pour tout $x \in S$,
$$f_k(x) \geq -f_k(-x) + 2f_k(0) \geq -M + 2\mu.$$
Ainsi les fonctions f_k sont bien minorées uniformément sur K.

ii Par le théorème 3.4.1, on en déduit que les f_k et f sont uniformément lipschitziennes sur K (i.e. avec une constante de Lipschitz indépendante de k).

iii Soit $\epsilon > 0$. On recouvre K par un nombre fini de boules de rayon ϵ, centrées en des points $x_i, 1 \leq i \leq m$. Pour $x \in S$, il existe i tel que $x \in B(x_i, \epsilon)$. On a
$$|f(x) - f_k(x)| \leq |f(x) - f(x_i)| + |f(x_i) - f_k(x_i)| + |f_k(x_i) - f_k(x)|.$$

En majorant chaque terme indépendamment de x, pour k assez grand, il vient

$$\exists k_o \in \mathbb{N}, \forall x \in S, \forall k \geq k_0,\ |f_k(x) - f(x)| \leq \epsilon$$

ce qui donne la convergence uniforme sur S.

Des variantes de ce résultat existent dans la littérature : suite de fonctions définies sur un intervalle de $\mathbb{R}$...

3.4.3 Vrai ou faux ?

i VRAI OU FAUX ? Si $f : \mathbb{R}^N \to \mathbb{R}$ est convexe et majorée sur $\mathbb{R}^N$, alors elle est constante.

ii VRAI OU FAUX ? Si $f : \mathbb{R}^N \to \mathbb{R}$ est convexe et minorée sur $\mathbb{R}^N$, alors elle est constante.

iii VRAI OU FAUX ? Si $f : \mathbb{R} \to \mathbb{R}$ est convexe et majorée au voisinage de $+\infty$, alors $\lim_{x\to-\infty} f(x) = +\infty$.

Les solutions :

i VRAI. En effet, sur un compact K fixé, f est lipschitzienne de constante $2M/\epsilon$, où M est une borne sur f, et $\epsilon > 0$. En faisant tendre ϵ vers $+\infty$, on obtient f constante.

ii FAUX. Contre-exemple : la fonction exponentielle.

iii FAUX. Contre-exemple : une fonction constante.

3.4.4 Différentiabilité

Remarque : le théorème de Rademacher (dur) affirme que toute fonction lipschitzienne est différentiable presque partout (au sens de la mesure de Lebesgue). C'est donc le cas pour les fonctions convexes. Notons que cela peut permettre, par exemple, des changements de variable convexes dans les calculs d'intégrale. Ici, on va démontrer ces résultats de dérivabilité presque partout, de façon élémentaire.

Cas de la dimension 1

Lemme 3.4.7. Dérivabilité à gauche et à droite

Soit f une fonction convexe sur un intervalle ouvert I. Alors f admet une dérivée à gauche et à droite en tout point. De plus, on a pour tout $x \in I$, $f'_d(x) \geq f'_g(x)$.

Preuve : ce résultat découle immédiatement du critère des pentes croissantes. Pour $x \in I$, l'application définie sur $\mathbb{R} \setminus \{0\}$ par

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

est croissante, donc admet une limite finie à gauche et à droite en 0, et ces limites vérifient l'inégalité annoncée.

Remarque : si la fonction est définie jusqu'aux bornes de l'intervalle, les limites des taux d'accroissement existent mais peuvent être infinies.

Théorème 3.4.8. Dérivabilité presque partout

Soit f une fonction convexe sur un intervalle I. Alors f est dérivable sur I privé d'un ensemble au plus dénombrable.

Preuve : remarquons d'abord que pour $x_1, x_2 \in I, x_1 < x_2$, on a, toujours par le critère des pentes croissantes, $f'_d(x_1) \leq f'_g(x_2)$. Soit x un point où f n'est pas dérivable. Deux cas sont possibles. Premier cas : la dérivée à gauche ou à droite de f en x est infinie. Ces cas se produisent au plus aux deux bornes de l'intervalle. Deuxième cas : les dérivées à droite et à gauche sont finies mais ne coïncident pas. Notons

$$\Delta = \{x \in I, -\infty < f'_g(x) < f'_d(x) < +\infty\}$$

et considérons la famille

$$\{]f'_g(x), f'_d(x)[, x \in \Delta\}.$$

C'est une famille d'intervalles non vides, ouverts, et disjoints, elle est donc au plus dénombrable. En effet, on peut construire une injection

$$\begin{array}{ccl} \Delta & \to & \mathbb{N} \\ x & \mapsto & \min\{n \in \mathbb{N}, q_n \in]f'_g(x), f'_d(x)[\} \end{array}$$

où les q_n sont les rationnels. Ainsi Δ est au plus dénombrable et le théorème est démontré.

Cas de la dimension N

Lemme 3.4.9. Dérivées partielles et différentiabilité

Soit f une fonction convexe sur un ouvert convexe C de $\mathbb{R}^N$. Soit $x \in C$. Si f admet des dérivées partielles en x, alors f est différentiable en x.

Rappelons que ce résultat est faux sans l'hypothèse de convexité !

Preuve : supposons que f admette des dérivées partielles en x, fixé. Posons, pour h assez petit,

$$g(h) = f(x+h) - f(x) - \langle \nabla f(x), h \rangle.$$

Remarquons que g est convexe. Il nous faut montrer que $g(h) = o(\|h\|)$. Puisqu'on a des informations sur les dérivées partielles, on va décomposer h sous la forme $h = \sum_{i=1}^N h_i e_i$, où les e_i sont les vecteurs de base. L'existence de dérivées partielles nous donne, pour tout i,

$$\frac{g(h_i e_i)}{|h_i|} \to 0.$$

On va majorer $g(h)$ en utilisant la convexité de g. On a

$$\begin{aligned} g(h) = g\left(\frac{1}{N}\sum_{i=1}^N N h_i e_i\right) &\leq \frac{1}{N}\sum_{i=1}^N g(N h_i e_i) \\ &\leq \sum_{h_i \neq 0} h_i \frac{g(N h_i e_i)}{N h_i} \\ &\leq \|h\| \sum_{h_i \neq 0} \left|\frac{g(N h_i e_i)}{N h_i}\right|. \end{aligned}$$

Ainsi,

$$\frac{g(h)}{\|h\|} \leq \sum_{h_i \neq 0} \left| \frac{g(Nh_i e_i)}{Nh_i} \right|.$$

D'autre part, on a, d'après la convexité de g,

$$g(h) + g(-h) \geq 2g(0) = 0$$

donc

$$\frac{-g(h)}{\|h\|} \leq \sum_{h_i \neq 0} \left| \frac{g(-Nh_i e_i)}{Nh_i} \right|.$$

Finalement,

$$\frac{|g(h)|}{\|h\|} \leq \max \left(\sum_{h_i \neq 0} \left| \frac{g(Nh_i e_i)}{Nh_i} \right|, \sum_{h_i \neq 0} \left| \frac{g(-Nh_i e_i)}{Nh_i} \right| \right).$$

Le second membre tend vers 0 quand h tend vers 0, ce qui démontre le lemme.

Théorème 3.4.10. Différentiabilité presque partout

Soit C un ouvert convexe de $\mathbb{R}^N$ et soit f une fonction convexe sur C. Alors f est différentiable presque partout sur C.

Preuve : d'après le lemme, il suffit de montrer que pour tout $1 \leq k \leq N$, l'ensemble

$$E_k = \left\{ x \in C, \frac{\partial f}{\partial x_k} \text{ n'existe pas} \right\} = \{x \in C, f'_+(x, e_k) - f'_-(x, e_k) > 0\}$$

est négligeable. Les fonctions $f'_\pm(\cdot, e_k)$ sont mesurables comme limites simples de fonctions mesurables, donc E_k est bien mesurable. On se ramène au cas d'un ouvert borné en considérant des intersections $E_k \cap B_n$. On a alors, par le théorème de Fubini,

$$\mu(E_k \cap B_n) = \int_{\mathbb{R}^N} \chi_{E_k \cap B_n} dx = \int_{\mathbb{R}} \cdots \left(\int_{\mathbb{R}} \chi_{E_k \cap B_n} dx_1 \right) \cdots dx_N.$$

Or on a, pour tout $(x_2, \cdots x_N)$, l'application $x_1 \mapsto f(x_1, x_2, ..., x_N)$ est convexe et d'après le cas 1D, elle est dérivable presque partout. Donc

$$\int_{\mathbb{R}} \chi_{E_k \cap B_n} dx_1 = 0.$$

On en déduit $\mu(E_k \cap B_n) = 0$, $\forall n$ et donc en passant aux unions dénombrables, $\mu(E_k) = 0$.

References

[1] G. Allaire, Analyse numérique et optimisation, éditions de l'école polytechnique, 2005.

[2] D. Azé, Éléments d'analyse convexe et variationnelle

[3] M. Berger, Géométrie, T.2

[4] H. Brézis, Analyse fonctionnelle, théorie et applications.

[5] A. Chambert-Loir, Fermigier, Maillot, Analyse T.1

[6] J.-B. Hiriart-Urruty, Optimisation et Analyse convexe, Exercices, Presses universitaires de France, 1998.

[7] J.-B. Hiriart-Urruty, C. Lemaréchal, Convex analysis and minimization algorithms I, Springer-Verlag, 1996

[8] A.W. Roberts, Convex functions, Academic Press, New York, London, 1973.

[9] W. Rudin, Analyse fonctionnelle.

[10] R. Schneider, Convex bodies: the Brunn-Minkowski theory, Cambridge University Press, 1993.

[11] P. Tauvel, Géométrie.