P. Maurer
ENS Rennes

Leçon 253. Utilisation de la notion de convexité en analyse.

Devs :

- Théorème du point fixe de Markov-Kakutani
- Théorème de Hahn-Banach en dimension finie

Références :

1. Odoux-Ramis-Deschamps, Topologie et éléments d'analyse
2. Testard, Analyse mathématique
3. Rouvière, Petit guide du calcul différentiel
4. Objectif Agrégation
5. Hirriat, Optimisation et analyse convexe
6. Gourdon, Algèbre
7. Brianes-Pagès, Théorie de l'intégration
8. Stein & Shakarchi, Real Analysis
9. Ouvrard, Probabilités 2
10. Cadre-Vial, Statistique mathématique

1 Fonctions et ensembles convexes

1.1 Fonctions convexes en dimension 1

Théorème 1. *Soit $f: I \to \mathbb{R}$. Les propositions suivantes sont équivalentes :*

1. $\forall (x,y) \in I^2 \quad \forall t \in [0,1] \quad f(tx + (1-t)\,y) \leq t\,f(x) + (1-t)\,f(y)$.
2. $\forall (x,y,z) \in I^3$, $(x<y<z) \Longrightarrow \left(\dfrac{f(y)-f(x)}{y-x} \leq \dfrac{f(z)-f(x)}{z-x} \leq \dfrac{f(z)-f(y)}{z-y}\right)$.
3. $\forall a \in I \quad t \mapsto \dfrac{f(t)-f(a)}{t-a}$ *est croissante sur* I.
4. *L'ensemble* $\mathrm{epi}(f) = \{(x,y) \in \mathbb{R}^2 \ :\ x \in I$ *et* $f(x) \leq y\}$, *appelé épigraphe de* f, *est une partie convexe de* $\mathbb{R}^2$.

Définition 2. *Une application $f: I \to \mathbb{R}$ qui vérifie les propositions équivalentes du théorème 21 est dite convexe sur I.*

Proposition 3. *Si f est convexe sur I, pour toute famille $(x_i)_{1 \leq i \leq n}$ d'éléments de I et toute famille $(\alpha_i)_{1 \leq i \leq n}$ de réels positifs vérifiant $\sum_{i=1}^{n} \alpha_i = 1$, on a*

$$f\left(\sum_{i=1}^{n} \alpha_i\, x_i\right) \leq \sum_{i=1}^{n} \alpha_i\, f(x_i).$$

Définition 4. *On dit qu'une fonction $f: I \to \mathbb{R}$ est concave sur I si $-f$ est convexe sur I.*

Théorème 5. *(Régularité des fonctions convexes)*
Soit f une fonction convexe sur I. Alors f admet une dérivée à droite et à gauche en tout point $a \in \mathrm{Int}(I)$, et $f_g'(a) \leq f_d'(a)$; f est continue sur $\mathrm{Int}(I)$, et d'autre part, si $(a,b) \in \mathrm{Int}(I)^2$ et $a<b$, on a

$$f_d'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f_g'(b).$$

Enfin, f_d' et f_g' sont des fonctions croissantes sur I.

Exemple 6. La fonction f peut être convexe sur I sans être continue sur I.
Par exemple, $f: [0,1] \to \mathbb{R}$ définie par $f(t) = 0$ si $t \in \,]0,1[$ et $f(0) = f(1) = 1$.

Théorème 7. *Soit f une fonction convexe sur I.*

- *Si f est dérivable sur I, alors f est convexe sur I si et seulement si f' est croissante sur I.*
- *Si f est dérivable deux fois sur I, alors f est convexe sur I si et seulement si $f'' \geq 0$ sur I.*

1.2 Ensembles convexes

Définition 8. *Si E est un espace vectoriel et $C \subset E$, on dit que C est une partie convexe de E si pour tout $x_1, \ldots, x_k \in C$ et $\lambda_1, \ldots, \lambda_d \in [0,1]$ tels que $\lambda_1 + \cdots + \lambda_d = 1$, la combinaison convexe $\lambda_1 x_1 + \cdots + \lambda_k x_k$ est encore un élément de C.*

Exemple 9. Soit $(E, \|\cdot\|)$ un espace vectoriel normé réel. Alors la boule unité fermée de E est convexe.

Définition 10. *Une partie $A \subset E$ est dite étoilée s'il existe $x \in A$ tel que pour tout $y \in A$, toute combinaison convexe de x et de y est dans A.*

Remarque 11. Une partie convexe est étoilée.

Définition 12.

Soit $\mathcal{C}$ un ouvert convexe de E contenant zéro. On définit la jauge de $\mathcal{C}$ comme l'application

$$j_{\mathcal{C}}\colon \left\{ \begin{array}{lll} E & \to & \mathbb{R}_+ \\ x & \mapsto & \inf\{\lambda>0 : x\in\lambda\mathcal{C}\} \end{array} \right. .$$

Lemme 13. *La jauge de $\mathcal{C}$ est bien définie sur E. Elle vérifie les propriétés suivantes, pour x et y des vecteurs de E :*

1. $\mathcal{C}=\{x\in E : j_{\mathcal{C}}(x)<1\}$,
2. $\forall\mu>0 \quad j_{\mathcal{C}}(\mu x)=\mu j_{\mathcal{C}}(x)$, *($j_{\mathcal{C}}$ est positivement homogène)*
3. $j_{\mathcal{C}}(x+y)\leq j_{\mathcal{C}}(x)+j_{\mathcal{C}}(y)$. *($j_{\mathcal{C}}$ est sous-additive)*

Définition 14. *Soit $X\neq\varnothing$ une partie de E. L'intersection $\mathrm{Conv}(X)$ de toutes les parties convexes de $\mathcal{E}$ qui contiennent X est appelée l'enveloppe convexe de X.*

Théorème 15. *(Carathéodory).*
Soit $A\subset\mathcal{E}$, avec $\mathcal{A}\neq\varnothing$. Tout élément de l'enveloppe convexe $\mathrm{Conv}(A)$ de A s'écrit comme combinaison convexe de $N+1$ points de $\mathcal{A}$, où $N+1=\dim(E)$.

Développement 1 :

Corollaire 16. *L'enveloppe convexe d'un ensemble compact est compacte.*

Corollaire 17. *L'enveloppe convexe d'un ensemble borné est bornée, et de plus si (E,d) est un espace métrique et $A\subset E$, on a $\delta(\mathrm{Conv}(A))=\delta(A)$, où $\delta(A):=\sup_{x,x'\in A} d(x,x')$ désigne le diamètre de A.*

Théorème 18. *(Point fixe de Markov-Kakutani).*
Soit $(E,\|\cdot\|)$ un espace euclidien de dimension $N\in\mathbb{N}^*$, et G un sous-groupe compact de $\mathrm{GL}(E)$.
On se donne K une partie convexe compacte de E et on suppose que K est stable par tous les éléments de G. Alors il existe $x\in K$ tel que pour tout $g\in G$, on ait $g(x)=x$. Autrement dit, G possède un point fixe dans K.

1.3 Séparation des convexes

Théorème 19. *(Hahn-Banach, forme analytique)*
Soit $p\colon E\to\mathbb{R}$ positivement homogène et sous-additive. Soit G un sous-espace vectoriel de E et g une forme linéaire sur G telle que $g\leq p$.
Alors il existe une forme linéaire f sur E telle que $f_{|G}=g$ et $f\leq p$ sur E.

Définition 20. *Soit φ une forme linéaire sur E. On appelle hyperplan affine de E tout ensemble H de la forme $H=\mathrm{Ker}(\varphi-c)$, où c est un réel quelconque.*
H est alors un espace affine dirigé par l'hyperplan vectoriel $\mathrm{Ker}(\varphi)$.

Définition 21.
Soit H un hyperplan affine de E. On se donne $\varphi\in E^*$ et $c\in\mathbb{R}$ tels que $H=\mathrm{Ker}(\varphi-c)$.

- *On appelle les demi-espaces limités par H les deux ensembles*
$$E_1=\{x\in E : \varphi(x)\leq c\} \quad \text{et} \quad E_2=\{x\in E : \varphi(x)\geq c\}.$$
- *Etant donné A et B deux parties de E, on dit que H sépare A et B si $A\subset E_1$ et $B\subset E_2$ ou $A\subset E_2$ et $B\subset E_1$.*

Développement 2 :

Lemme 22. *Soit C un convexe ouvert de E non vide et $x_0\in E\setminus C$. Alors il existe un hyperplan affine H de E séparant $\{x_0\}$ et C.*

Théorème 23. *(Hahn-Banach, forme géométrique)*
Soit A et B deux convexes de E disjoints et non vides. Si A est ouvert, il existe un hyperplan affine H qui sépare A et B.

2 Inégalités de convexité

2.1 Inégalités classiques

Proposition 24. *Pour tout $x\in\mathbb{R}$, on a $\exp(x)\geq x+1$ et pour tout $x\in]0,+\infty[$, on a $\ln(x)\leq x-1$.*

Proposition 25. *Pour tout $x\in\left[0,\frac{\pi}{2}\right]$, on a $\frac{2}{\pi}x\leq\sin(x)\leq x$.*

Proposition 26. *(Inégalités arithmético-géométriques)*

- *Soit $t\in]0,1[$ et $a,b\in[0,+\infty]$. Alors*
$$a^{1-t}b^t\leq(1-t)a+tb.$$
- *Soit $t\in]0,1[$. Alors, pour $a,b\in[0,+\infty]$ et $\lambda>0$ on a*
$$a^{1-t}b^t\leq(1-t)\lambda^{\frac{1}{1-t}}a+t\frac{b}{\lambda^{\frac{1}{t}}}.$$
- *Soit $p,q>1$ avec $\frac{1}{p}+\frac{1}{q}=1$. Alors pour $a,b\in[0,+\infty]$ et $\lambda>0$ on a*
$$ab\leq\lambda^p a^p/p+\lambda^{-q}b^q/q.$$
- *Soit $n\in\mathbb{N}^*$ et $(x_1,\ldots,x_n)\in\mathbb{R}_+^n$. Alors* $\left(\prod_{i=1}^n x_i\right)^{1/n}\leq\frac{1}{n}\sum_{i=1}^n x_i.$

Théorème 27. *(Inégalité de Brunn-Minkowski, admis).*
Soit $d \geq 1$ et A et B deux sous ensembles Lebesgue-mesurables de $\mathbb{R}^d$, tels que $A+B$ soit Lebesgue-mesurable. En notant par m la mesure de Lebesgue, il vient

$$m(A+B)^{1/d} \geq m(A)^{1/d} + m(B)^{1/d}.$$

2.2 Inégalités dans les espaces de Lebesgue

On se donne $(X, \mathcal{A}, \mu)$ un espace mesuré.

Proposition 28. *(Inégalités de Hölder)*

- *Soit $f, g: X \to [0, +\infty]$ deux fonctions mesurables positives et $t \in]0,1[$. Alors on a*
$$\int_X f^{1-t} g^t \, d\mu \leq \left(\int_X f d\mu\right)^{1-t} \left(\int_X g d\mu\right)^t.$$
- *Soit $f, g: X \to [0, +\infty]$ deux fonctions mesurables positives et $p, q \in]1, +\infty[$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Alors*
$$\int_X f g d\mu \leq \left(\int_X f^p\right)^{1/p} \cdot \left(\int_X g^q\right)^{1/q}.$$
- *Soit $p, q \in]1, +\infty[$ tels que $\frac{1}{p} + \frac{1}{q} = 1$, et $f: X \to [0, +\infty]$ une fonction mesurable positive telle que $\int_X f^p \, d\mu < +\infty$. Alors*
$$\left(\int_X f^p \, d\mu\right)^{1/p} = \sup_{g \geq 0} \frac{\int_X f g d\mu}{\left(\int_X g^q \, d\mu\right)^{1/q}} \leq \sup_{g \geq 0, \ \int_X g^q d\mu \leq 1} \int_X f g d\mu.$$

Proposition 29. *(Inégalité de Minkowski)*
Soit $p \in]1, +\infty[$. Si $f, g: X \to [0, +\infty]$ sont deux fonctions mesurables positives sur X, alors

$$\left(\int_X (f+g)^p \, d\mu\right)^{1/p} \leq \left(\int_X f^p \, d\mu\right)^{1/p} + \left(\int_X g^p \, d\mu\right)^{1/p}.$$

Application 30. (Théorème de Riesz-Fischer).
Pour $p \in [1, +\infty[$, l'espace $(L^p, \|\cdot\|_p)$ est un espace de Banach.

2.3 Inégalités en probabilités

Proposition 31. *(Inégalité de Jensen).*
Soit X une variable aléatoire réelle intégrable sur $(\Omega, \mathcal{F}, \mathbb{P})$, et $\varphi: \mathbb{R} \to \mathbb{R}$ une fonction convexe. Alors

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)].$$

Proposition 32. *Soit X une variable aléatoire réelle intégrable sur $(\Omega, \mathcal{F}, \mathbb{P})$. On suppose que $|X| \leq 1$ $\mathbb{P}$-presque sûrement, et que $\mathbb{E}[X] = 0$.*

Proposition 33. *Pour tout $t \in \mathbb{R}$, $e^{tX} \in L^1$ et on a $\mathbb{E}[\exp(tX)] \leq \cosh(t)$.*

Remarque 34. On déduit du développement en série entière de cosh (proposition 19) l'inégalité

$$\mathbb{E}[\exp(tX)] \leq \exp\left(\frac{t^2}{2}\right).$$

Théorème 35. *(Inégalité de Hoeffding)*
Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes définies sur Ω. On suppoose que pour tout $n \in \mathbb{N}$, X_n est $\mathbb{P}$-presque sûrement bornée par une constante $c_n > 0$, intégrable et d'espérance nulle. On note $S_n = \sum_{i=1}^n X_j$
Pour tout $\varepsilon > 0$, on a

$$\mathbb{P}(|S_n| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right).$$

Application 36. Soit $(\mathcal{H}^n, \{P_\theta\}_{\theta \in \Theta})$ un modèle statistique, avec $\mathcal{H} \subset \mathbb{R}$ et $\Theta \subset \mathbb{R}$. On veut estimer le paramètre $g(\theta)$ avec $g: \Theta \to \mathbb{R}$ continue. Si $(X_1, \ldots, X_n)$ sont indépendantes et identiquement distribuées, de loi P_θ, bornées P_θ-p.s et avec $\mathbb{E}[X_1] = g(\theta)$, alors

$$I_\alpha = \left[\overline{X_n} - c\sqrt{\frac{2}{n} \ln\left(\frac{2}{\alpha}\right)}, \overline{X_n} + c\sqrt{\frac{2}{n} \ln\left(\frac{2}{\alpha}\right)}\right]$$

est un intervalle de confiance pour $g(\theta)$ au niveau $1-\alpha$, où c est une borne p.s de $X_1 - g(\theta)$.

3 Applications en optimisation

3.1 Recherche d'un extremum

Théorème 37. *(Caractérisation des fonctions convexes différentiables)*
Soit U un ouvert connexe de $\mathbb{R}^n$ et $f: U \to \mathbb{R}$ une application différentiable sur U. Alors les propriétés suivantes sont équivalentes :

- *f est convexe.*
- *Le graphe de f est au-dessus de ses tangentes*
$$\forall x, y \in U \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$
- *L'application $\nabla f: U \to \mathbb{R}^n$ vérifie*
$$\forall x, y \in U \quad \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

Si de plus f est deux fois différentiable sur U, elle est convexe si et seulement si sa matrice Hessienne est positive : $\forall x, h \in U \quad \langle H f(x) \cdot h, h \rangle \geq 0$, et elle est strictement convexe si et seulement si se matrice Hessienne est définie positive.

Exemple 38. Soit $A \in \mathcal{S}_n$. Alors $f: x \mapsto \langle x, Ax \rangle$ est convexe sur $\mathbb{R}^n$ si et seulement si A est positive. Elle est strictement convexe si et seulement si $f \in \mathcal{S}_n^{++}$.

Proposition 39. *Si $f: E \to \mathbb{R}$ est convexe, alors tout minimum local est global.*

Théorème 40. *Soit C un convexe non vide de E et $f: C \to \mathbb{R}$ une application strictement convexe. Alors il existe au plus un minimum de f sur C.*

Corollaire 41. *Soit $f: E \to \mathbb{R}$ une fonction convexe continue. Alors f admet un unique minimum sur E.*

3.2 Résolution numérique de systèmes linéaires

On introduit la méthode du gradient, dont l'objectif est de trouver une solution à $Ax = b$ via la recherche d'un minimum local de la fonction $x \mapsto \frac{1}{2}\langle Ax, x \rangle + \langle b, x \rangle$ pour $A \in \mathcal{S}_n^{++}(\mathbb{R})$.

Lemme 42. *(Inégalité de Kantorovitch)*

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$. On note λ_1 et λ_n la plus petite (respectivement la plus grande) valeur propre de A. Alors :

$$\begin{aligned}\forall x \in \mathbb{R}^n \quad \langle Ax, x \rangle \langle A^{-1}x, x \rangle &\leq \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2 \|x\|^4 \\ &= \frac{1}{4}\left(\sqrt{c(A)} + \sqrt{\frac{1}{c(A)}}\right)^2 \|x\|^4,\end{aligned}$$

où $c(A) \stackrel{\text{def}}{=} \frac{\lambda_1}{\lambda_n}$ désigne le contenu de A.

Théorème 43. *Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à minimiser $f: x \mapsto \frac{1}{2}\langle Ax, x \rangle + \langle b, x \rangle$ quand x parcourt $\mathbb{R}^n$.*

Il existe une unique solution $\bar{x}$ à ce problème, et elle est caractérisée par $\nabla f(\bar{x}) = 0$.

De plus, la suite définie par $\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N} \quad x_{k+1} = x_k + t_k d_k \end{cases}$ converge vers $\bar{x}$, où $t_k = -\nabla f(x_k)$ est l'unique réel minimisant la fonction $t \mapsto f(x_k + t d_k)$.

3.3 Convexité dans les espaces de Hilbert

On se donne H un espace de Hilbert, et on note $\|\cdot\|$ la norme dérivant du produit scalaire $\langle \cdot, \cdot \rangle$ de H.

Théorème 44. (de projection sur un convexe fermé)

Soit Γ une partie fermée, convexe, non vide de H. Alors pour tout point $x \in H$, il existe un unique point $y \in \Gamma$ tel que

$$\|x - y\| = \inf_{\gamma \in \Gamma} \|x - \gamma\| = d(x, \Gamma).$$

Ce point est appelé projection de x sur Γ et est noté $p_\Gamma(x)$. Il est caractérisé par la propriété suivante :

$$\begin{cases} p_\Gamma(x) \in \Gamma \\ \forall \gamma \in \Gamma \quad \mathfrak{Re}\langle x - p_\Gamma(x), \gamma - p_\Gamma(x) \rangle \leq 0 \end{cases}.$$

Application 45. Soit F un sous-espace vectoriel de H (non nécessairement fermé). Pour $x \in H$, le projeté orthogonal $p_F(x)$ de x sur F est caractérisé par :

$$\begin{cases} p_F(x) \in F \\ x - p_F(x) \in F^\perp \end{cases}$$

Par ailleurs, on a l'égalité :

$$H = \bar{F} \oplus (\bar{F})^\perp = \bar{F} \oplus F^\perp.$$

On en déduit un critère de densité pour un sous-espace vectoriel d'un espace de Hilbert :

$$F \text{ est dense dans } H \iff F^\perp = \{0\}.$$

Théorème 46. (de représentation de Riesz).

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. Alors l'application δ définie par

$$\mathcal{L}: \begin{cases} H & \to & H' \\ a & \mapsto & \mathcal{L}_a = (x \mapsto \langle x, a \rangle) \end{cases}$$

est une bijection antilinéaire isométrique.

Application 47. (Théorème de Radon-Nikodym).

Soit $(X, \mathcal{A})$ un espace mesurable et μ, ν deux mesures positives σ-finies sur $(X, \mathcal{A})$. Il y a équivalence entre :

1. $\forall A \in \mathcal{A} \quad \mu(A) = 0 \Longrightarrow \nu(A) = 0 \quad$ (on note $\nu \ll \lambda$).
2. $\exists f: (X, \mathcal{A}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ intégrable telle que $\forall A \in \mathcal{A} \quad \nu(A) = \int_A f d\mu$.

En outre, la fonction f est unique (à une égalité μ-presque partout près).

On note $f = \frac{d\nu}{d\mu}$ et on dit que f est la dérivée de Radon-Nikodym, ou la densité de ν par rapport à μ.