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

1. Notation. Dans toute cette leçon, on considère un $\mathbf{R}$-espace vectoriel E.

1. Les fonctions convexes en analyse

1.1. Quelques rappels et régularité des fonctions convexes dans le cas réel

2. Définition. Soit $C \subset E$ une partie convexe. Une fonction $f\colon C \longrightarrow \mathbf{R}$ est *convexe* si, pour tous vecteurs $x, y \in C$ et tout réel $\lambda \in [0,1]$, on a

$$f(\lambda x + (1-\lambda)x) \leqslant \lambda f(x) + (1-\lambda) f(y). \qquad (*)$$

Elle est *strictement convexe* si l'inégalité $(*)$ est stricte et tient avec $x \neq y$ et $\lambda \in \left]0,1\right[$.

3. Exemple. Si l'espace E est euclidien, la fonction $x \longmapsto \langle x, x\rangle$ est convexe. Les fonctions constantes sont convexes.

4. Remarque. Une fonction $f\colon C \longrightarrow \mathbf{R}$ est convexe si et seulement si son *épigraphe*

$$\operatorname{epi}(f) \coloneqq \{(x,r) \in C \times \mathbf{R} \mid f(x) \leqslant r\}$$

est convexe.

5. Exemple. Une fonction affine est convexe.

6. Proposition *(inégalité des pentes)*. Soient $I \subset \mathbf{R}$ un intervalle et $f\colon I \longrightarrow \mathbf{R}$ une fonction convexe. Soient $a, b, c \in I$ trois réels vérifiant $a < b < c$. Alors

$$\frac{f(a)-f(b)}{a-b} \leqslant \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(b)-f(c)}{b-c}.$$

7. Théorème. Une fonction convexe sur un intervalle I est continue sur l'intervalle $\mathring{I}$.

8. Remarque. Le résultat se généralise à une fonction convexe $f\colon C \subset \mathbf{R}^n \longrightarrow \mathbf{R}$: elle est continue sur l'ouvert $\mathring{C}$. Mais le résultat est faux en dimension infinie : il suffit de prendre une forme linéaire qui n'est pas continue.

9. Théorème. Une fonction dérivable sur un intervalle I est convexe si et seulement si sa dérivée est croissante sur l'intervalle I.

10. Exemple. Les fonctions exponentielles et logarithmes sont respectivement convexe et concave.

11. Proposition. Une fonction convexe sur un intervalle ouvert est la borne supérieure d'une famille de fonctions affines.

1.2. Fonctions convexes différentiables et optimisation sur un convexe

12. Théorème. Soient $\Omega \subset \mathbf{R}^n$ un ouvert convexe et $f\colon \Omega \longrightarrow \mathbf{R}$ une fonction différentiable. Alors les points suivants sont équivalents :

– la fonction f est convexe ;
– pour tous points $x, y \in \Omega$, on a $f(y) \geqslant f(x) + \langle \nabla f(x), y - x\rangle$;
– pour tous points $x, y \in \Omega$, on a $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geqslant 0$.

De plus, si la fonction f est deux fois différentiables, alors on rajoute le point suivant :

– pour tout point $x \in \Omega$, la hessienne $\mathrm{H} f(x)$ est positive.

13. Contre-exemple. Le convexe Ω nécessite d'être ouvert sans quoi le théorème est faux. Par exemple, la fonction $(x,y) \longmapsto x^2 - y^2$ est convexe sur la partie $\mathbf{R} \times \{0\}$, mais sa hessienne en tout point, à savoir la matrice $\operatorname{diag}(1,-1)$, n'est pas positive

14. Application. Soit $A \in \mathscr{S}_n(\mathbf{R})$ une matrice symétrique. Alors elle est positive si et seulement si la fonction $x \longmapsto \langle Ax, x\rangle$ est convexe.

15. Théorème *(inégalité d'Euler)*. Soient $\Omega \subset \mathbf{R}^n$ un ouvert et $C \subset \Omega$ un convexe. Soient $f\colon \Omega \longrightarrow \mathbf{R}$ une fonction différentiable. On suppose que la fonction $f|_C$ admet un minimum en un point $x^\star \in C$ et qua la fonction f est différentiable en ce point. Alors

$$df(x^\star)(y - x^\star) \geqslant 0.$$

16. Proposition. Soient $\Omega \subset \mathbf{R}^n$ un ouvert convexe. Alors tout point critique d'une fonction différentiable $\Omega \longrightarrow \mathbf{R}^n$ est un minimum global pour cette dernière.

17. Contre-exemple. L'hypothèse de convexité est nécessaire : la fonction $x \longmapsto x^3$ admet l'origine comme point critique, mais ce n'en est pas un minimum.

18. Théorème. Une fonction strictement convexe admet au plus un minimum.

19. Application *(point de Fermat)*. Soient A, B et C trois points non alignés du plan euclidien $\mathbf{R}^2$. On suppose que les trois angles du triangle ABC sont strictement inférieurs à $2\pi/3$. Alors la fonction

$$\left|\begin{aligned} &\mathbf{R}^2 \longrightarrow \mathbf{R}, \\ &M \longmapsto MA + MB + MC \end{aligned}\right.$$

admet un unique point minimum qui est dans l'intérieur stricte du triangle ABC.

2. Quelques inégalités de convexité

2.1. Premières inégalités

20. Théorème *(inégalité arithmético-géométrique)*. Soient $x_1, \dots, x_n \geqslant 0$ des réels positifs. Alors

$$(x_1 \cdots x_n)^{1/n} \leqslant \frac{x_1 + \cdots + x_n}{n}.$$

21. Proposition. Soient $p, q > 0$ deux réels positifs vérifiant $1/p + 1/q = 1$. Soient $x, y > 0$ deux réels strictement positifs. Alors

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}.$$

22. Proposition. Soit $x \in \mathbf{R}$ un réel. Alors $e^x \geqslant 1 + x$ et $\ln(1+x) \leqslant x$.

23. Remarque. Cette première inégalité permet de trouver la domination nécessaire au théorème de convergence dominée et d'obtenir la limite

$$\Gamma(z) = \lim_{n \to +\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{z-1} \, dt, \qquad \operatorname{Re} z > 0.$$

24. Proposition. Soit $x \in \left]0, \pi/2\right[$ un réel. Alors $\frac{\pi}{2} x < \sin x < x$.

2.2. Applications aux probabilités et aux espaces de Lebesgue

25. Théorème *(inégalité de Hölder)*. Soit $(X, \mathscr{A}, \mu)$ un espace mesuré. Soient $p, q > 1$ deux réels positifs vérifiant $1/p + 1/q = 1$. Soient $f \in \mathrm{L}^p(X)$ et $g \in \mathrm{L}^q(X)$. Alors la

fonction fg est intégrable et

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q.$$

26. Corollaire. Avec les mêmes notations, on a

$$\left| \int_X fg \, \mathrm{d}\mu \right| \leqslant \|f\|_p \|g\|_q.$$

27. Théorème *(inégalité de Minkowski).* Soit $p \geqslant 1$ un réel. Alors pour toutes fonctions $f, g \in \mathrm{L}^p(X)$, on a

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p.$$

En particulier, l'application $\| \ \|_p$ est une norme sur l'espace vectoriel $\mathrm{L}^p(X)$.

28. Théorème *(inégalité de Jensen).* Soient $(\Omega, \mathscr{F}, \mathbf{P})$ un espace probabilisé et $\varphi \colon \mathbf{R} \longrightarrow \mathbf{R}_+$ une fonction bornée convexe. Soit X une variable aléatoire réelle intégrable. Alors

$$\varphi(\mathbf{E}[X]) \leqslant \mathbf{E}[\varphi(X)].$$

29. Remarque. On retrouve l'inégalité $\mathbf{E}[X]^2 \leqslant \mathbf{E}[X^2]$.

3. Résultats en analyse fonctionnelle

3.1. Le théorème de projection dans un espace de Hilbert

30. Théorème *(de projection).* Soient H un espace de Hilbert et $C \subset H$ un convexe fermé non vide. Pour tout $x \in H$, il existe un unique élément $p_C(x) \in C$ tel que

$$\|p_C(x) - x\| = \mathrm{d}(x, C).$$

De plus, le point $p_C(x)$ est caractérisé par les conditions

$$p_C(x) \in C \qquad \text{et} \qquad \forall y \in C, \quad \mathrm{Re}\langle x - p_C(x), y - p_C(x)\rangle \leqslant 0.$$

Enfin, l'application $p_C \colon H \longrightarrow C$ est 1-lipschitzienne.

31. Contre-exemple. Toutes les hypothèses sont nécessaires.

- L'hypothèse hilbertienne est nécessaire. Dans l'espace $(\mathscr{C}^0([0,1]), \| \ \|_\infty)$, la distance $\mathrm{d}(1, C)$ avec $C \coloneqq \{f \in \mathscr{C}^0([0,1]) \mid 0 \leqslant f \leqslant 1, f(0) = 0\}$ est réalisée par les fonctions $f \in C$.
- L'hypothèse de complétude est nécessaire. En prenant $E \coloneqq \mathscr{C}^0([0,1]) \subset \mathrm{L}^2([0,1])$ avec $C \coloneqq (\mathbf{1}_{[0,1/2]})^\perp$ et $C_1 \coloneqq C \cap E$, la distance $\mathrm{d}(f_1, C_1)$ n'est pas atteinte pour toute fonction $f_1 \in E \setminus C_1$.
- L'hypothèse de convexité est nécessaire. Dans $\mathbf{R}^2$, l'origine admet une infinité de projetés sur la sphère unité $\mathscr{S}^1 \subset \mathbf{R}^2$.

32. Application *(moindres carrés).* Étant donnés n points $(x_i, y_i) \in \mathbf{R}^2$ tels que les réels x_i ne soient pas tous égaux entre eux, il existe $\lambda, \mu \in \mathbf{R}$ qui minimisent la somme

$$\sum_{i=1}^n (\lambda x_i + \mu - y_i)^2.$$

33. Application *(espérance conditonnelle).* Soient $(\Omega, \mathscr{F}, \mathbf{P})$ un espace probabilité et $\mathscr{G} \subset \mathscr{F}$ une sous-tribu. On peut écrire $\mathrm{L}^2(\Omega, \mathscr{G}, \mathbf{P}) \subset \mathrm{L}^2(\Omega, \mathscr{F}, \mathbf{P})$. Pour une variable aléatoire $X \in \mathrm{L}^2(\Omega, \mathscr{F}, \mathbf{P})$, on définit alors

$$\mathbf{E}[X \mid G] \coloneqq p_{\mathrm{L}^2(\Omega, \mathscr{G}, \mathbf{P})}(X).$$

L'application d'espérance conditionnelle

$$\mathbf{E}[\cdot \mid \mathscr{G}] \colon \mathrm{L}^2 \cap \mathrm{L}^1(\Omega, \mathscr{F}, \mathbf{P}) \longrightarrow \mathrm{L}^1(\Omega, \mathscr{F}, \mathbf{P})$$

étant uniformément continue, elle se prolonge donc à l'espace $\mathrm{L}^1(\Omega, \mathscr{F}, \mathbf{P})$.

34. Corollaire *(théorème du supplémentaire orthogonal).* Soient H un espace de Hilbert et $F \subset H$ un sous-espace vectoriel. Alors $H = F \oplus F^\perp$.

35. Théorème *(Riesz).* Soit H un espace de Hilbert. Alors l'application

$$\left| \begin{array}{l} H \longrightarrow H', \\ y \longmapsto \langle \cdot, y \rangle \end{array} \right.$$

est une isométrie surjective.

36. Théorème. Soit H un espace de Hilbert. Alors toute suite bornée $(x_n)_{n \in \mathbf{N}}$ de H admet une sous-suite convergeant faiblement, c'est-à-dire qu'il existe une extraction $\varphi \colon \mathbf{N} \longrightarrow \mathbf{N}$ et un vecteur $x \in H$ tels que

$$\forall y \in H, \qquad \langle x_{\varphi(n)}, y \rangle \longrightarrow \langle x, y \rangle.$$

37. Proposition. Soient H un espace de Hilbert et $C \subset H$ une partie convexe non bornée. Soit $J \colon C \longrightarrow \mathbf{R}$ une fonction convexe, continue et coercive. Alors cette dernière atteint sa borne inférieure.

38. Remarque. Soient $b \in \mathbf{R}^n$ un vecteur et $A \in \mathscr{S}_n^{++}(\mathbf{R})$ une matrice symétrique définie positive. En appliquant la proposition à la fonction $x \longmapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$, on trouve qu'elle admet un minimum qui se trouve être la solution du système $Ax = b$

3.2. Le théorème de séparation de Hahn-Banach

39. Définition. Soit E un espace vectoriel normé. Un *hyperplan* est un ensemble de la forme $\{x \in E \mid f(x) = \alpha\}$ pour une forme linéaire $f \in E^* \setminus \{0\}$ et un réel $\alpha \in \mathbf{R}$. Ce dernier *sépare* deux parties $A, B \subset E$ *au sens large* si

$$\forall x \in A, \qquad f(x) \leqslant \alpha,$$
$$\forall x \in B, \qquad f(x) \geqslant \alpha.$$

Il les *sépare au sens strict* s'il existe un réel $\varepsilon > 0$ tel que

$$\forall x \in A, \qquad f(x) \leqslant \alpha - \varepsilon,$$
$$\forall x \in B, \qquad f(x) \geqslant \alpha + \varepsilon.$$

40. Lemme. Soit $C \subset E$ un ouvert convexe contenant le vecteur nul. La fonction

$$p \colon \left| \begin{array}{l} E \longrightarrow \mathbf{R}, \\ x \longmapsto \inf\{\alpha > 0 \mid \alpha^{-1} x \in C\}. \end{array} \right.$$

est une semi-norme sur E et elle vérifie les points suivants :

- il existe une constante $M > 0$ telle que $0 \leqslant p(x) \leqslant M \|x\|$ pour tout $x \in E$;
- $C = \{x \in E \mid p(x) < 1\}$.

41. Lemme. Soit $C \subset E$ un convexe ouvert non vide avec $C \neq E$. Soit $x_0 \in E \setminus C$ un point. Alors il existe une forme linéaire continue $f \in E^*$ telle que

$$\forall x \in C, \qquad f(x) < f(x_0).$$

42. Application. Munissons l'espace $\mathscr{M}_n(\mathbf{R})$ de la norme $\| \ \|_2$. Alors l'enveloppe convexe du groupe orthogonal $\mathrm{O}(n)$ est la boule unité fermée de $\mathscr{M}_n(\mathbf{R})$.

D1

- D2 -

43. Théorème *(Hahn-Banach géométrique, première forme).* Soient $A, B \subset E$ deux convexes non vides disjoints. On suppose que le premier A est ouvert. Alors il existe un hyperplan $\{f = \alpha\}$ avec $f \in E^* \setminus \{0\}$ qui sépare les parties A et B au sens large.
44. Théorème *(Hahn-Banach géométrique, deuxième forme).* Soient $A, B \subset E$ deux convexes non vides disjoints. On suppose que le premier A est fermé et que le second B est compact. Alors il existe un hyperplan $\{f = \alpha\}$ avec $f \in E^* \setminus \{0\}$ qui sépare les parties A et B au sens strict.
45. Corollaire. Soit $F \subset E$ un sous-espace vectoriel. On suppose que toute forme linéaire continue de E' qui est nulle sur F est nulle sur E. Alors le sous-espace F est dense dans E.

[1] Vincent Beck, Jérôme Malick et Gabriel Peyré. *Objectif Agrégation.* 2e édition. H&K, 2005.
[2] Haïm Brézis. *Analyse fonctionnelle.* 2e tirage. Masson, 1983.
[3] Marc Briane et Gilles Pagès. *Théorie de l'intégration.* Vuibert, 2012.
[4] Xavier Gourdon. *Algèbre.* 2e édition. Ellipses, 2009.