Leçon 253 : Utilisation de la convexité en analyse

RM
2022-2023

1 Définition et propriétés de la convexité

On se place dans un $\mathbb{R}$-espace vectoriel E de dimension $n \geq 1$ normée.

1.1 Ensemble convexe

Définition 1 : Soit C une partie de E; on dit que C est convexe si tout segment dont les extrémités sont dans C est inclus dans C, c'est-à-dire si

$$\forall (x,y) \in C^2, \quad \forall \lambda \in [0,1], \quad \lambda x + (1-\lambda)y \in C.$$

Exemple 2 : Les intervalles de $\mathbb{R}$ sont convexes. Les sous espaces vectoriels de E sont convexes.
Remarque 3 : Toutes les parties convexes sont connexes.
Proposition 4 : Soit $(C_i)_{i\in I}$ une famille de convexes de E. Alors l'intersection des C_i est un convexe de E.
Définition 5 : Soit A une partie de E. Il existe une plus petite partie convexe de E contenant A. On l'appelle enveloppe convexe de A et on la note $\mathrm{Conv}(A)$. L'ensemble $\mathrm{Conv}(A)$ est aussi l'ensemble des barycentres des points de A affectés de coefficients positifs, i.e $\mathrm{Conv}(A)$ est l'ensemble des x tels que

$$\exists x_1, ..., x_n \in A, \exists \lambda_1, ..., \lambda_n \in \mathbb{R}^+, \sum_{i=1}^n \lambda_i = 1, \quad x = \sum_{i=1}^n \lambda_i x_i.$$

Remarque 6 : On peut donc plus généralement, définir un convexe dans un espace affine ou le segment $\lambda x + (1-\lambda)y$ est vue comme un barycentre.
Théorème (de Carathéodory) 7 : Pour tout point x de l'enveloppe convexe d'une partie A de E, x est le barycentre de coefficients positifs de au plus $n+1$ points de A.
Application 8 : Si A est une partie compact de E, alors $\mathrm{Conv}(A)$ est compacte.

1.2 Fonction convexe

Définition 9 : Soient C un convexe de E est f une application de C dans $\mathbb{R}$; on dit que f est une application convexe sur C si

$$\forall (x,y) \in C^2, \quad \forall \lambda \in [0,1], \quad f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Elle est dite concave si $-f$ est convexe.
On dit que la fonction f est strictement convexe si l'inégalité précédente est stricte pour tout $x \neq y$ et pour tout $\lambda \in [0,1]$. Elle est strictement concave si $-f$ est strictement convexe.
Remarque 10 : Définir une fonction convexe sur un ensemble qui n'est pas convexe n'a pas de sens.
Proposition 11 : On appelle l'épigraphe de f l'ensemble $\mathrm{epi}(f) = \{(x,r) \in C \times \mathbb{R}, \quad f(x) \leq r\}$.
Alors, une application est convexe si, et seulement si, son épigraphe est une partie convexe de $E \times \mathbb{R}$.
Remarque 12 : Cela permet de faire le lien entre ensemble convexe et fonction convexe.
Exemple 13 : L'application norme $\|.\| : E \mapsto \mathbb{R}^+$ est toujours convexe.
On suppose maintenant que $C = I$ un intervalle de $\mathbb{R}$.
Exemple 14 : La fonction $x \mapsto e^x$ est convexe et la fonction $x \mapsto \ln(x)$ est concave.
Proposition 15 : L'application $f : I \mapsto \mathbb{R}$ est convexe si et seulement si on a l'inégalité des pentes, i.e

$$\forall a < b < c, \quad \frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(a)}{c-a} \leq \frac{f(c)-f(b)}{c-b}.$$

Théorème 16 : Soit $f : I \mapsto \mathbb{R}$ une application dérivable sur I. Les assertions suivantes sont équivalentes : (i) f est convexe.
(ii) f' est croissante.
(iii) La courbe représentative de f est au dessus de ces tangentes.
Corollaire 17 : Une application $f : I \mapsto \mathbb{R}$ deux fois dérivable est convexe si et seulement si $f''(x) \geq 0$ pour tout $x \in I$.
Proposition 18 : Soit une application $f : I \mapsto \mathbb{R}$ convexe. Alors

$$\forall x_1, ...x_n \in I, \forall \alpha_1, ..., \alpha_n > 0, \quad f(\frac{\alpha_1 x_1 + ... + \alpha_n x_n}{\alpha_1 + ... + \alpha_n}) \leq \frac{\alpha_1 f(x_1) + ... + \alpha_n f(x_n)}{\alpha_1 + ... + \alpha_n}.$$

1.3 Séparation des ensembles convexes

Définition 19 : Un hyperplan est un ensemble de la forme

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

où f est une forme linéaire sur E, non identiquement nulle et $\alpha \in \mathbb{R}$. On dit que H est l'hyperplan d'équation $[f = \alpha]$.
Définition 20 : Soient $A \subset E$ et $B \subset E$. On dit que l'hyperplan H d'équation $[f = \alpha]$ sépare A et B au sens large si l'on a

$$f(x) \leq \alpha, \forall x \in A \quad \text{et} \quad f(x) \geq \alpha, \forall x \in B.$$

On dit que H sépare au sens strict s'il existe $\epsilon > 0$ tel que

$$f(x) \leq \alpha - \epsilon, \forall x \in A \quad \text{et} \quad f(x) \geq \alpha + \epsilon, \forall x \in B.$$

Théorème (Hahn-Banach géométrique, forme 1) 21 : Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides et disjoints. On suppose que A est ouvert. Alors il existe un hyperplan fermé qui sépare A et B au sens large.
Théorème (Hahn-Banach géométrique, forme 2) 22 : Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides et disjoints. On suppose que A est fermé et que B est compact. Alors il existe un hyperplan fermé qui sépare A et B au sens stricte.
Corollaire 23 : Soit $F \subset E$ un sous espace vectoriel tel que $\overline{F} \neq E$. Alors il existe $f \in E^*$, $f \neq 0$ tel que

$$f(x) = 0 \quad \forall x \in F.$$

Remarque 24 : Ce corollaire est utile pour montrer qu'un sous espace vectoriel $F \subset E$ est dense.

2 Inégalités de convexité

2.1 Inégalités classiques

Proposition 25 : Pour tout $x \in \mathbb{R}$, $e^x \geq 1+x$ et pour tout $y \in]-1;+\infty[$, $\ln(1+y) \leq y$.
Théorème (Inégalité arithmético-géométrique) 26 : Soient $x_1, ..., x_n$ des nombres réels positifs. On a

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

Proposition (Inégalité de Young) 27 : Soient $x, y \in \mathbb{R}_*^+$ et $p, q \in \mathbb{N}$ tel que $1/p + 1/q = 1$. Alors grâce à la concavité du logarithme, on a

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

2.2 Dans les espaces L^p

On considère un espace mesuré $(X, \mathcal{A}, \mu)$.
Soit $p \in [1, +\infty[$ et $q \in \mathbb{R}$ tel que $1/p + 1/q = 1$.
Définition 28 : On définit

$$L^p_{\mathbb{K}}(X, \mathcal{A}, \mu) := \{f : (X, \mathcal{A}) \mapsto (\mathbb{K}, \mathcal{B}(\mathbb{K})) \text{ mesurable}/ \int_X |f|^p d\mu < +\infty\}.$$

$$L^\infty_{\mathbb{K}}(X, \mathcal{A}, \mu) := \{f : (X, \mathcal{A}) \mapsto (\mathbb{K}, \mathcal{B}(\mathbb{K})) \text{ mesurable}/ \ \exists C > 0, |f(x)| \leq C \ \mu - p.p\}$$

l'espace vectoriel noté L^p par la suite et on pose pour toute fonction $f : (X, \mathcal{A}) \mapsto \mathbb{K}$ la quantité

$$\text{Si } p < +\infty, \|f\|_p := \left(\int_X |f|^p d\mu\right)^{1/p} \leq +\infty \quad \text{et} \quad \|f\|_\infty = \inf\{C > 0, |f(x)| \leq C \ \mu\text{-}p.p\}$$

Théorème (Inégalité de Hölder) 29 : Soient $f, g : X \mapsto \mathbb{K}$ mesurables. Si $f \in L^p$ et $g \in L^q$ alors $fg \in L^1$ et

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Théorème (Inégalité de Minkowski) 30 : Si $p \in [1, +\infty[$, alors

$$\forall f, g \in L^p, \quad \|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Remarque 31 : Cette inégalité permet de donné à L^p une structure d'espace vectoriel semi-normé.
Définition 32 : On pose $L^p := L^p/\{\|.\|_p = 0\}$. On obtient alors que $(L^p, \|.\|_p)$ est un espace vectoriel normé.
Théorème (Riesz-Ficher) 33 : Pour tout $p \in [1; +\infty[$, l'espace vectoriel normé $(L^p, \|.\|_p)$ est complet.

2.3 Dans un espace probabilisé

On considère $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire.
Proposition (Inégalité de Jensen) 34 : Soient g une fonction convexe sur $\mathbb{R}$, alors on a

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

Application 35 : Cela permet de par exemple retrouver la formule $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$.
Proposition (Inégalité de Hoeffding) 36 : Soit une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires réelles indépendantes, bornées $\mathbb{P}$-p.s et centrées; on suppose que $|X_n| \leq c_n$ $\mathbb{P}$-$p.s$, avec $c_n > 0$. On note $S_n = \sum_{j=1}^n X_j$. Alors pour tout $\epsilon > 0$, on a

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

Application 37 : Soit $\alpha > 0$. On suppose qu'il existe $\beta > 0$ tel que $\sum_{j=1}^n c_j^2 \leq n^{2\alpha - \beta}$. Alors on a

$$\frac{S_n}{n^\alpha} \xrightarrow[n \to +\infty]{p.s} 0.$$

3 Application de la Convexité dans certains espaces

3.1 En probabilité

Soit X une variable aléatoire définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$.
Définition 38 : On appelle fonction génératrice de X, notée G_X, la fonction définie pour $s \in \mathbb{R}$ par

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{+\infty} \mathbb{P}(X = k) s^k$$

lorsque la série converge.

Développement (Processus de branchement de Galton-Watson) 39 : Soit $(X_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}^*}$ une famille de variables aléatoires à valeur dans $\mathbb{N}$ indépendantes identiquement distribuées de loi $p_k = \mathbb{P}(X = k)$ pour $k \in \mathbb{N}$ avec $p_0 \in]0,1[$ et d'espérance m. On définit la suite (Z_n) de la manière suivante

$$Z_0 = 1 \quad et \quad \forall n \in \mathbb{N},\ Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i} \quad (Z_{n+1} = 0 \text{ si } Z_n = 0).$$

Enfin on note $\pi_n = \mathbb{P}(Z_n = 0)$ et $P_{ext} = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$.
Si $m \leq 1$, $P_{ext} = 1$ et le processus s'éteint presque surement.
Si $m > 1$, $P_{ext} < 1$ et il y a une probabilité de survie non nul.

Dev 1

Remarque 40 : Ici Z_n modélise le nombre d'individus à la génération n et $X_{n,i}$ le nombre de descendant de l'individu i à la n-ième génération, π_n la probabilité d'extinction à la génération n et P_{ext} la probabilité d'extinction de la population.
On étudie alors la descendance d'un seul individu et donc on pose $Z_0 = 1$.
Le développement nous dit alors que si la moyenne de descendant pour chaque individus est inférieur ou égale à 1, alors la lignée va s'éteindre presque surement, et si la moyenne est plus grande que 1, alors il y a une probabilité non nul que la lignée survive.

3.2 Dans un espace de Hilbert

Soit $(H, \langle .,. \rangle)$ un espace de Hilbert.

Développement (Projection sur un convexe fermé) 41 : Soit $\mathcal{C}$ un convexe fermé (non vide) de H. Alors, pour tout $x \in H$, il existe un unique élément $y \in \mathcal{C}$, tel que $d(x,\mathcal{C}) = \inf_{z\in\mathcal{C}} \|x - z\| = \|x - y\|$ et on appelle le point $y = p_{\mathcal{C}}(x)$ le projeté orthogonale de x sur $\mathcal{C}$. On a ainsi

$$\forall z \in \mathcal{C}, \qquad \|x - p_{\mathcal{C}}(x)\| \leq \|x - z\|.$$

Il vérifie les propriétés suivante :
- $\forall z \in \mathcal{C}$, $Re\langle p_{\mathcal{C}}(x) - x, p_{\mathcal{C}}(x) - z\rangle \leq 0$.
- l'application p est 1-lip.

Application (Théorème de représentation de Riesz) 42 : Soit f une forme linéaire continue sur H. Alors il existe un unique vecteur $a \in H$ tel que, pour tout x de H, $f(x) = \langle a, x\rangle$.

Dev 2

Remarque 43 : On a le même résultat avec les hypothèses ou H est préhilbertien mais $\mathcal{C}$ est un convexe complet.

Remarque 44 : On voit sur la figure en annexe que le projeté $p_{\mathcal{C}}(x)$ est l'élément de $\mathcal{C}$ le plus proche du point de x. On retrouve aussi l'inégalité $\forall z \in \mathcal{C}$, $Re\langle p_{\mathcal{C}}(x) - x, p_{\mathcal{C}}(x) - z\rangle \leq 0$ car l'"angle" formé entre les vecteurs $x - p_{\mathcal{C}}(x)$ et $y - p_{\mathcal{C}}(x)$ est obtus.

Théorème (Projection sur un sous-espace fermé) 45 : Soit F un sous-espace vectoriel fermé de H. Pour $x \in H$, le projeté $p_F(x)$ de x sur F est l'unique élément $p \in H$ qui vérifie

$$p \in F \quad \text{et} \quad x - p \in F^{\perp}$$

De plus, l'application $p_F : H \mapsto F$ est linéaire, continue, surjective. L'espace H se décompose en la somme directe orthogonale

$$H = F \oplus F^{\perp}$$

Théorème (Lax-Milgram) 46 : Soit $a(.,.)$ une forme bilinéaire continue et coercive ($\exists \alpha > 0$ telle que $a(v,v) \geq \alpha|v|^2 \quad \forall v \in H$).
Alors pour tout $\varphi \in H^*$ une forme linéaire, il existe un unique $u \in H$ tel que

$$a(u,v) \geq \varphi(v) \qquad \forall v \in K.$$

De plus, si a est symétrique, alors u est caractérisé par la propriété

$$u \in H \qquad \text{et} \qquad \frac{1}{2}a(u,u) - \varphi(u) = \min_{v\in K}\{\frac{1}{2}a(v,v) - \varphi(v)\}.$$

Remarque 47 : Ce théorème est un outil fondamental dans l'étude de certaines équations aux dérivées partielles.

4 Fonctions convexes différentiables et optimisation

On se place dans $\mathbb{R}^n$ muni du produit scalaire canonique.

Théorème 48 : Soient $\mathcal{U}$ un ouvert de $\mathbb{R}^n$ et f une application de $\mathcal{U}$ dans $\mathbb{R}$. Si f est différentiable sur $\mathcal{U}$, alors les propriétés suivantes sont équivalentes :
(i) f est convexe.
(ii) le graphe de f est "au-dessus de ses tangentes" :

$$\forall x, y \in \mathcal{U}, \qquad f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$$

(iii) l'application $\nabla f : \mathcal{U} \mapsto \mathbb{R}^n$ est "monotone" :

$$\forall x, y \in \mathcal{U}, \qquad \langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0.$$

Ajoutons, si f est deux fois différentiable sur $\mathcal{U}$
$(iiii)$ $\nabla^2 f(x)$ est positive :

$$\forall x \in \mathcal{U}, \quad \forall h \in \mathbb{R}^n, \qquad \langle \nabla^2 f(x)h, h\rangle \geq 0$$

Application 49 : Soit $A \in S_n$ une matrice symétrique; alors on a l'équivalence

$$x \mapsto f(x) = \langle Ax, x \rangle \text{ est convexe} \Longleftrightarrow A \text{ est positive.}$$

Proposition (Inégalité d'Euler) 50 : Soient $\mathcal{U}$ un ouvert de $\mathbb{R}^n$, $\mathcal{C}$ un convexe inclus dans $\mathcal{U}$ et f une application de $\mathcal{U}$ dans $\mathbb{R}$. Si f admet un minimum local en $x^* \in \mathcal{C}$ et si elle est différentiable en x^*, alors

$$df(x^*).(y - x^*) \geq 0 \qquad \forall y \in \mathcal{C}$$

Proposition 51 : Soient $\mathcal{C}$ un ouvert connexe de $\mathbb{R}^n$ et f une application convexe de $\mathcal{C}$ dans $\mathbb{R}$. Alors si f est différentiable en x^* et $\mathrm{d}f(x^*) = 0$, alors x^* est un minimum local de f, et même globale.
Proposition 52 : Soient $\mathcal{C}$ un convexe non vide et $f : \mathcal{C} \mapsto \mathbb{R}$ une application strictement convexe sur $\mathcal{C}$. Alors, il existe au plus un point $\overline{x} \in \mathcal{C}$ minimisant f sur $\mathcal{C}$.

5 Bibliographie

[1] Xavier GOURDON, *Les maths en tête : Algèbre Probabilité 3ème édition*

[2] Xavier GOURDON, *Les maths en tête : Analyse 3ème édition*

[3] Vincent BECK, Jérôme MALICK et Gabriel PEYRE, *Objectif agrégation*

[4] François ROUVIERE, *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*

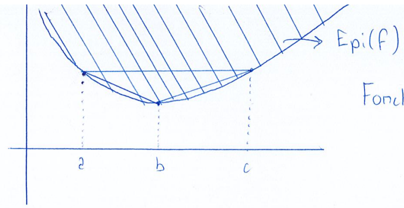
[5] Serge FRANCINOUS, Hervé GIANELLA et Serge NICOLAS, *Oraux X-ENS tome 3 analyse*

[6] Haïm BREZIS, *Analyse fonctionnelle : Théorie et applications*

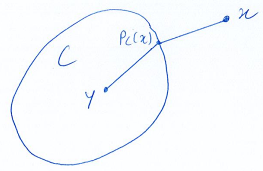
[7] Marc BRIANE, Gilles PAGES, *Analyse : Théorie de l'intégration*

[8] Jean-Yves OUVRARD, *Probabilités 2*

6 Annexe



Fonction convexe.



Projection sur un convexe fermé.