253. Utilisation de la notion de convexité en analyse

Ref : . Berger T2 (Géométrie)
. Brézis, La cuisine aux champignons
. objectif Agrég.
. Gourdon (AN)
. Ouvrard T2

I Utilisation de la notion d'ensemble convexe

1) Définition et exemples

[Ber] p.73 **Définition** : Une partie S d'un espace affine X est dite convexe si $\forall x, y \in S : [x;y] \subset S$ où $[x;y] = \{\lambda x + (1-\lambda) y : \lambda \in [0;1]\}$.

[Ber] p.74 **Exemples** : 1) sont convexes ; X, tout sous-espace affine de X, les points, les droites et les hyperplans.

2) les convexes de $\mathbb{R}$ sont les intervalles.

3) Soient X, Y des espaces affines, $S \subset X$ et $T \subset Y$ des convexes, et $f : X \to Y$ une application affine. Alors $f(S) \subset Y$ et $f^{-1}(T) \subset X$ sont convexes.

4) Toute intersection de convexes est un convexe. En particulier les intersections finies de demi-espaces : ce sont les polyèdres convexes.

· Une réunion croissante de convexes est encore convexe.

[OA] p.26 5) Dans un espace de Hilbert, un convexe fermé est l'intersection des demi-espaces qui le contiennent.

2) Enveloppe convexe

[Ber] p.24 **Théorème (Carathéodory)** : Pour toute partie A de X, espace affine de dimension d

$$\mathcal{E}(A) = \left\{ \sum_{i=1}^{d+1} \lambda_i x_i \;;\; \forall i\; x_i \in A, \forall i\; \lambda_i \geq 0, \sum_{i=1}^{d+1} \lambda_i = 1 \right\}$$

[Ber] p.27 **Corollaire** : si A est compact, alors $\mathcal{E}(A)$ l'est aussi.

3) Théorèmes de séparation $|E\ \mathbb{K}\text{-evn}|$

[Bré] p.1 **Théorème (Hahn-Banach, forme analytique)** : E un $\mathbb{R}$-evn
soit $p : E \to \mathbb{R}$ une application vérifiant :
1) $p(\lambda x) = \lambda p(x)$, $\forall x \in E$ et $\forall \lambda > 0$
2) $p(x+y) \leq p(x) + p(y)$, $\forall x, y \in E$
soit $G \subset E$ un sous espace vectoriel et $g : G \to \mathbb{R}$ une application linéaire telle que
3) $g(x) \leq p(x)$, $\forall x \in G$
Alors il existe une forme linéaire f qui prolonge G telle que 4) $f(x) \leq p(x)$, $\forall x \in E$

[Bré] p.4 **Définition** : Un hyperplan (affine) 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 $\alpha \in \mathbb{R}$

[Bré] p.5-6 DEV **Théorème** : 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.

[Bré] p.6 **Lemme** : Soit $C \subset E$ un convexe ouvert non vide et soit $x_0 \in E$ avec $x_0 \notin C$. Alors il existe $f \in E'$ tel que $f(x) < f(x_0)$ $\forall x \in C$. En particulier l'hyperplan d'équation $[f = f(x_0)]$ sépare $\{x_0\}$ et C au sens large.

[Bré] p.7 **Théorème** : Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides, 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 strict.

[Bré] p.8 **Corollaire** : 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 $\langle f, x \rangle = 0$, $\forall x \in F$

4) Applications

[OA] p.95 **Théorème (Projection sur un convexe fermé)** : Soit $(H, \langle\cdot,\cdot\rangle)$ un espace de Hilbert et C un convexe fermé (non vide) de H. Alors, pour tout $x \in H$, il existe un unique élément de C, qui réalise la distance de x à C. Ce point est appelé la projection de x sur C, il est noté $P_C(x)$. On a ainsi $\forall y \in C$, $\|x - P_C(x)\| \leq \|x - y\|$.
L'élément $P_C(x)$ est de plus caractérisé par
i) $P_C(x) \in C$
ii) $\mathrm{Re}(\langle x - P_C(x), y - P_C(x) \rangle) \leq 0$, $\forall y \in C$

[OA] p97 **Application** : Hahn-Banach géométrique : La forme géométrique du théorème de Hahn-Banach exprime que l'on peut « séparer les convexes disjoints » sous certaines hypothèses pour un énoncé général dans un espace de Banach. Dans un espace de Hilbert, la preuve est bien plus simple.

[OA] p98 **Corollaire** : Projection sur un sous-espace fermé : Soient $(H;\langle\cdot,\cdot\rangle)$ un espace de Hilbert et 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 qui vérifie : $p \in F$ et $x - p \in F^\perp$.
De plus l'application $p_F : H \to F$ est linéaire continue et surjective. L'espace H se décompose en la somme directe orthogonale $H = F \oplus F^\perp$, et p_F est le projecteur associé à cette décomposition. Enfin, pour trois éléments x, x_1 et x_2 de H, les propriétés suivantes sont équivalentes : 1) $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$

2) $x_1 = p_F(x)$ et $x = p_{F^\perp}(x)$

[OA] p100 **Applications** : 1) Critère de densité : soit F un sous-espace vectoriel de H. On a $H = \overline{F} \oplus (\overline{F})^\perp = \overline{F} \oplus F^\perp$. D'où un critère de densité dans un espace de Hilbert : F est dense dans $H \Leftrightarrow F^\perp = 0$

2) Espérance conditionnelle : Soient $(X, \mathcal{A}, \mu)$ un espace probabilisé et $\mathcal{F}$ une sous tribu de $\mathcal{A}$. Soit F le sous-espace vectoriel de $L^2(X;\mathcal{A})$ formé des fonctions $\mathcal{F}$-mesurables. Notons ν la restriction de μ à $\mathcal{F}$. Si f est $\mathcal{F}$-mesurable et positive, alors $\int_X f\,d\mu = \int_X f\,d\nu$. On définit alors l'espérance conditionnelle sachant $\mathcal{F}$ d'un élément de $L^2(X;\mathcal{A})$ comme sa projection orthogonale sur F.

[OA] p100 **Application** : soient E et F deux espaces vectoriels normés, et $f : U \to F$ différentiable sur un ouvert U de E. Si U est un ouvert convexe de E et si $\|Df(x)\|_{\mathcal{L}(E;F)} \leq k$ pour tout $x \in U$ (où k est une constante positive), alors l'application f est k-lipschitzienne sur U.

II Utilisation de la notion de fonction convexe

1) Définition, exemples :

Définition : Soit $A \subset X$ un convexe d'un espace affine X ; une application $f : A \to \mathbb{R}$ est dite convexe si $\forall \lambda \in [0,1], \forall x, y \in A$ on a $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$. [Ber] p54
Elle est dite strictement convexe si l'inégalité est stricte.

Prop : Soit $f : A \to \mathbb{R}$, où A est un convexe. Alors « f est convexe » est équivalent à « l'épigraphe de f est un convexe de $X \times \mathbb{R}$ ». [Ber] p54

Exemple : 1) Soit $(E;\langle\cdot,\cdot\rangle)$ un espace euclidien et $\|\cdot\|$ la norme associée. Alors l'application $x \to \|x\|^2$ est une fonction strictement convexe sur E. [OA] p27

2) Soit $(E;\|\cdot\|)$ un espace vectoriel normé. L'application norme $\|\cdot\| : E \to \mathbb{R}$ est convexe (inégalité triangulaire)

3) $x \to \exp(x)$ est convexe sur $\mathbb{R}$.

Application : inégalité de Hoeffding : Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles indépendantes bornées presque partout et centrées ; on suppose que $|X_n| \leq c_n$ $\mathbb{P}$ presque partout, avec $c_n > 0$. On note $\forall n \in \mathbb{N}^*$, $S_n = \sum_{j=1}^{n} X_j$. Alors $\forall \varepsilon > 0$, $\mathbb{P}(|S_n| > \varepsilon) \leq 2 \exp\left(\frac{-\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right)$ [OUV2]

Proposition : Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction convexe sur I. Alors $\forall x \in I$ la fonction "pente issue de x"

$$p_x : \begin{cases} I\setminus\{x\} \longrightarrow \mathbb{R} \\ y \longrightarrow \dfrac{f(y) - f(x)}{y - x} \end{cases}$$

est croissante.

Application : La fonction Γ est entière caractérisée sur $\mathbb{R}^{+*}$ par :
- $\Gamma(1) = 1$
- Pour tout $x > 0$, on a $\Gamma(x+1) = x\Gamma(x)$
- la fonction composée $\ln \circ \Gamma$ est convexe sur $\mathbb{R}^{+*}$.

[GA] p.28 **Propriétés** : Soit C un convexe de E ;

- La composée d'une fonction convexe et d'une fonction convexe d'une variable réelle est encore convexe.
- La borne supérieure de fonctions convexes est encore convexe (lorsqu'elle existe).
- Soit $(f_j)_{j\in J}$ une famille de fonctions convexes sur C. S'il existe une fonction $g : C \to \mathbb{R}$ telle que $f_j(x) \le g(x)$ $\forall x \in C$ et $\forall j \in J$, alors $\sup_{j\in J} f_j$ est une fonction convexe

[GA] p.26 → application : soit $\mathcal{S}_n$ l'espace vectoriel des matrices symétriques réelles. L'application $f : \begin{cases} \mathcal{S}_n \to \mathbb{R} \\ A \mapsto \lambda_{max}(A) \end{cases}$ est convexe.

[FGN] p.93 **Proposition** (Jensen) : Soit une application $f : I \to \mathbb{R}$ convexe (I intervalle de $\mathbb{R}$). Alors $\forall x_1,\dots,x_n \in I, \forall \alpha_1,\dots,\alpha_n > 0$, $f\left(\frac{\alpha_1 x_1 + \dots + \alpha_n x_n}{\alpha_1+\dots+\alpha_n}\right) \le \frac{\alpha_1 f(x_1)+\dots+\alpha_n f(x_n)}{\alpha_1+\dots+\alpha_n}$

② Fonctions convexes différentiables

Cadre : $\mathbb{R}^n$ muni de sa structure euclidienne canonique.

[GA] p.21 **Théorème** : Soient U un ouvert convexe de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}$. Si f est différentiable sur U, alors les propriétés suivantes sont équivalentes

(i) f est convexe

(ii) le graphe de f est "au-dessus de ses tangentes" $\forall x,y \in U$, $f(y) \ge f(x) + \langle \nabla f(x), y-x\rangle$

(iii) l'application $\nabla f : U \to \mathbb{R}^n$ est "monotone" : $\forall x,y \in U$, $\langle \nabla f(y) - \nabla f(x), y-x\rangle \ge 0$.

(iv) si f est 2 fois différentiable sur U, $Hf(x)$ est positive : $\forall x, h \in U$, $\langle Hf(x)h, h\rangle \ge 0$. (hessienne)

application : i) $x \mapsto -\ln x$ est convexe sur $\mathbb{R}_+^*$

conséquence 1 (successives...) : $\forall a,b > 0$ $ab \le \frac{1}{p}a^p + \frac{1}{q}b^q$, $\frac{1}{p}+\frac{1}{q}=1$ (inégalité de Young)

[BRE] p.56-57 $\Rightarrow \forall f \in L^p, g \in L^q$, $1 \le p \le \infty$, alors $\int |fg| \le \|f\|_{L^p}\|g\|_{L^q}$ (Hölder)

$\Rightarrow \|f+g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p}$ $\forall f,g \in L^p$, $1 \le p \le \infty$ (Minkowski)

$\Rightarrow L^p$ est un espace vectoriel et $\|\cdot\|_{L^p}$ est une norme $\forall 1 \le p \le \infty$.

[GOU] p.74 conséquence 2 (inégalité arithmético-géométrique) : $\forall x_1,\dots,x_n \in \mathbb{R}^+$, $(x_1\dots x_n)^{1/n} \le \frac{x_1+\dots+x_n}{n}$.

[GA] p.29 ii) Soit $A \in \mathcal{S}_n$ une matrice symétrique ; alors $x \mapsto \langle Ax,x\rangle$ convexe $\iff A$ est positive.

③ Fonctions convexes et extrema

[GA] p.30 **Remarque** : Soit C un convexe non vide. La convexité des fonctions étant une propriété "globale", elle permet d'assurer des inégalités sur C tout entier.

[GA] p.16 **Proposition** : Si f est une fonction convexe, alors les conditions nécessaires suivantes sont aussi suffisantes :

i) x^* minimum local de f et f différentiable en $x^* \iff Df(x^*)=0$.

[GA] p.18 ii) si $Df(x^*)=0$, alors $\begin{cases} x^* \text{ minimum local de } f \iff D^2f(x^*) \text{ est positive} \\ D^2f(x^*) \text{ définie positive} \iff x^* \text{ minimum local strict} \end{cases}$

iii)

[GA] p.30 **Proposition** : unicité du minimum

Soient C un convexe non vide et $f : C \to \mathbb{R}$ une application strictement convexe sur C. Alors il existe au plus un point $\bar{x} \in C$ minimisant f sur C.

exemple : en statistiques, on utilise la concavité de la fonction log-vraisemblance pour assurer que l'estimateur du maximum de vraisemblance est bien l'unique maximum global.

[FGN3] DEV application : ellipsoïde de John Loewner. Soit K un compact d'intérieur non vide de $\mathbb{R}^n$. Il existe un unique ellipsoïde centré en O de volume minimal contenant K.

④ Norme strictement convexe

[GA] p.32 **Définition** : Une norme $\|\cdot\|$ est strictement convexe si $\forall x \ne y$ et $\|x\|=\|y\|$ alors $\forall t \in]0,1[$, $\|tx+(1-t)y\| < t\|x\| + (1-t)\|y\|$

interprétation géométrique : soit S la sphère associée à $\|\cdot\|$. $\|\cdot\|$ est strictement convexe si le segment reliant deux points sur S ne rencontre S qu'en ses extrémités (figure 2)

[GA] p.30 **Proposition** : Soient $(F, \|\cdot\|)$ un evn, avec $\|\cdot\|$ strictement convexe, E un sous-espace vectoriel de F de dimension finie. Alors $\forall x \in F$, $\exists x^* \in E$ tel que $\|x - x^*\| = \inf_{y \in E} \|x-y\|$