253

Utilisation de la notion de convexité en analyse

Cadre : Soit E un $\mathbb{R}$-espace vectoriel.

I) Ensembles convexes, fonctions convexes

A] Ensembles convexes

Def 1 : Une partie C de E est dite convexe si $\forall A,B \in E$, $\forall \lambda \in [0;1]$, $\lambda A + (1-\lambda) B \in C$. (Voir Fig 1)

Ex 2 : Les (sous-)espaces vectoriels sont convexes.

Rem 3 : On peut généraliser cette définition dans le cadre des espaces affines : $\lambda A + (1-\lambda) B$ s'interprète alors comme barycentre.

Prop 4 : Toute partie convexe est connexe par arcs (et donc connexe).

Def 5 : Soit $A \subset E$. On appelle enveloppe convexe de A le plus petit convexe de E contenant A, noté $\mathrm{Conv}(A)$.

Prop 6 : $\mathrm{Conv}(A)$ est l'ensemble des barycentres positifs des points de A.

Thm 7 : (de Carathéodory) On suppose E de dimension $n \in \mathbb{N}$. Alors tout point de $\mathrm{Conv}(A)$ est barycentre positif de au plus $n+1$ points de A.

App 8 : Si $K \subset E$ est compact, $\mathrm{Conv}(K)$ est compact.

B] Fonctions convexes

Def 9 : Soit $f : E \to \mathbb{R}$. f est dite convexe si : $\forall x,y \in E$, $\forall t \in [0;1]$, $f(tx+(1-t)y) \leq tf(x) + (1-t)f(y)$. f est dite strictement convexe si l'inégalité est stricte dès que $x \neq y$ et $t \in]0;1[$. f est concave si $(-f)$ est convexe.

Ex 10 : Toute norme sur E est convexe.

Rem 11 : Si E est un intervalle de $\mathbb{R}$, l'inégalité traduit le fait que f soit en dessous de ses cordes (Voir Fig 2)

Prop-def 12 : Soit $\mathrm{Epi}(f) = \{(x,y) \in E \times \mathbb{R} \mid y \geq f(x)\}$ l'épigraphe de f. f convexe $\Leftrightarrow$ $\mathrm{Epi}(f)$ convexe (Voir Fig 2).

On suppose maintenant que $E = I$ est un intervalle de $\mathbb{R}$.

Thm 13 : f est convexe si et seulement si l'inégalité des trois pentes est satisfaite : $\forall a<b<c \in \mathbb{R}$,

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

Cor 14 : f admet une dérivée à droite et à gauche en tout point de $\mathring{I}$. En particulier, f est continue sur $\mathring{I}$.

Thm 15 : f est convexe si et seulement si f' est croissante.

Thm 16 : (de Bohr-Mollerup) Soit $f : \mathbb{R}_+^* \to \mathbb{R}_+^*$ tel que :
① $\forall x \in \mathbb{R}_+^*$, $f(x+1) = x f(x)$
② $f(1) = 1$
③ $\ln(f)$ est convexe.
Alors $f = \Gamma$ où $\forall x \in \mathbb{R}_+^*$, $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} \, dt$

C] Séparation de convexes dans $\mathbb{R}^n$

On suppose que E est un $\mathbb{R}$-espace vectoriel normé de dimension $n \in \mathbb{N}$.

Def 17 : Soient $A \subset E$ et $B \subset E$. Soit H un hyperplan d'équation $[f = \alpha]$, c'est-à-dire $H = \{x \in E \mid f(x) = \alpha\}$ où f est une forme linéaire continue sur E et $\alpha \in \mathbb{R}$. On dit que H sépare A et B au sens large si $\forall x \in A$, $f(x) \leq \alpha$ et $\forall x \in B$, $f(x) \geq \alpha$. La séparation est au sens strict si $\exists \varepsilon > 0 \,/\, \forall x \in A$, $f(x) \leq \alpha - \varepsilon$ et $\forall x \in B$, $f(x) \geq \alpha + \varepsilon$.

Def 18 : Soit C un convexe de E contenant 0, avec $0 \in \mathring{C}$. On définit la jauge de C par : $\forall x \in E$, $p(x) = \inf\{\alpha > 0 \mid \frac{x}{\alpha} \in C\}$.

Lem 19 : Il existe $M > 0$ tel que $\forall x \in E$, $0 \leq p(x) \leq M\|x\|$. De plus, $C = \{x \in E \mid p(x) < 1\}$.

Thm 20 : (de Hahn-Banach analytique) Soit $p : E \to \mathbb{R}$

Vérifiant : $\forall x,y \in E, \forall \lambda > 0$, $p(\lambda x) = \lambda p(x)$ et $p(x+y) \leq p(x)+p(y)$.
Soient $G \subset E$ et $g : G \to \mathbb{R}$ vérifiant $g \leq p_{|G}$ et linéaire.
Alors il existe $f : E \to \mathbb{R}$ linéaire prolongeant g tel que $f \leq p$.

Thm 21 : (de Hahn-Banach géométrique, forme 1) Soient A et B deux convexes de E non vides et disjoints. Si A est ouvert, il existe un hyperplan fermé séparant A et B au sens large.

[2] **Thm 22 :** (de Hahn-Banach géométrique, forme 2) Si A est fermé et B compact, alors il existe un hyperplan fermé les séparant au sens strict.

App 23 : Soit $F \subset E$, $F \neq E$. Alors il existe $f \in E'$, $f \neq 0$ tel que $f_{|F} = 0$.

Rem 24 : Ceci est pratique pour démontrer des résultats de densité.

II) Inégalités de convexité

A] Inégalités classiques

Thm 25 : (inégalité arithmético-géométrique) Soient $x_1, \dots, x_n$ des réels positifs. Alors $\left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{\sum_{i=1}^{n} x_i}{n}$

[7] **Prop 26 :** (inégalité d'Young) Soient $a, b \in \mathbb{R}_+^*$ et soient $p, q \in \mathbb{N}^*$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Alors $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$

Thm 27 : (inégalité de Kantorovitch) Soit $A \in S_n^{++}(\mathbb{R})$. Alors $\forall x \in \mathbb{R}^n$, $x \neq 0$, pour $\lambda_{max} = \max_{\lambda \in Sp(A)} \lambda$, $\lambda_{min} = \min_{\lambda \in Sp(A)} \lambda$,

[5]
$$\frac{\|x\|^4}{\|x\|_A^2 \|x\|_{A^{-1}}^2} \geq 4 \frac{\lambda_{max} \lambda_{min}}{(\lambda_{max} + \lambda_{min})^2}$$

Prop 28 : (inégalité de Jensen) Soit $g : [a,b] \to \mathbb{R}$ continue et soit $f : \mathbb{R} \to \mathbb{R}$ convexe. Alors $f\left(\frac{1}{b-a}\int_a^b g(t)dt\right) \leq \frac{1}{b-a}\int_a^b f(g(t))dt$

B] Inégalités dans les espaces L^p

[4] **Thm 29 :** (inégalité de Hölder) Soient $p \geq 1$ et $q \geq 1$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soient $f \in L^p$ et $g \in L^q$. Alors $fg \in L^1$ et
$\|fg\|_1 \leq \|f\|_p \|g\|_q$

Cor 30 : (inégalité de Hölder généralisée) Soient $p, q, r \geq 1$ tels que $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Soient $f \in L^p$ et $g \in L^q$. Alors $fg \in L^r$ et $\|fg\|_r \leq \|f\|_p \|g\|_q$

Cor 31 : (inégalité de Minkowski) Soient $p \geq 1$, et $f, g \in L^p$. [1]
Alors $\|f+g\|_p \leq \|f\|_p + \|g\|_p$.

Rem 32 : Ceci permet de donner une structure d'espace vectoriel normé à L^p.

Prop 33 : Soient $p, q \geq 1$ et $f \in L^p$ et $g \in L^q$, avec $\frac{1}{p} + \frac{1}{q} = 1$. Alors $f * g$ est définie partout et $f * g \in L^\infty$ avec :
$$\|f * g\|_\infty \leq \|f\|_p \|g\|_q$$

III) Application à l'optimisation

A] Dans un espace de Hilbert

Soit H un espace de Hilbert, c'est-à-dire un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ qui le rend complet.

DEV 2

Thm 34 : (de projection sur un convexe fermé) Soit C un convexe non vide de H, fermé. Alors $\forall x \in H$, $\exists ! P_C(x) \in C$, $\|x - P_C(x)\| = d(x, C)$. De plus, P_C est caractérisé par : $\forall y \in C$, $Re(\langle x - P_C(x) | y - P_C(x) \rangle) \leq 0$.

Prop 35 : L'application P_C est 1-lipschitzienne.

Prop 36 : Soit $F \subset H$ un sous-espace vectoriel fermé. [3]
Alors $P_F \in \mathcal{L}(H, F)$. De plus, pour tout x de H, $P_F(x)$ est l'unique élément de H tel que $P_F(x) \in F$ et $x - P_F(x) \in F^\perp$.

Cor 37 : Dans ce cas, $H = F \oplus F^\perp$ et $\forall x \in H$, $d(x, F) = \|x - P_F(x)\|$.

Thm 38 : (de représentation de Riesz) $\forall \phi \in E'$,

$\exists!\, y \in H$, $\varphi = \langle \cdot | y \rangle$. De plus, $\|\varphi\| = \|y\|$.

[3] **App 39**: Soit $f : H \to \mathbb{R}$ C^1 en $a \in H$. Alors $\exists!\, \nabla f(a) \in H$, $df_a = \langle \cdot | \nabla f(a) \rangle$. $\nabla f(a)$ est appelé gradient de f en a.

Def 40: Soit $(x_n) \in H^{\mathbb{N}}$. On dit que (x_n) converge faiblement vers $x \in H$, et on note $x_n \underset{n \to +\infty}{\rightharpoonup} x$, si $\forall y \in H$, $\langle x_n, y \rangle \underset{n \to +\infty}{\longrightarrow} \langle x, y \rangle$

[7] **Thm 41**: (DEV 2) (Optimisation dans un Hilbert) Soit $J : H \to \mathbb{R}$ continue, convexe et coercive c'est-à-dire : $J(x) \underset{\|x\| \to \infty}{\longrightarrow} +\infty$. Alors $\exists x^* \in H$, $J(x^*) = \inf_{x \in H} J(x)$

Rem 42: Ce théorème ne donne rien sur l'unicité d'un tel x^*

B] Optimisation sur un convexe

Soit K un convexe de H.

Prop 43: Soit $J : K \to \mathbb{R}$ convexe. Tout minimum local est global et l'ensemble de tels points est convexe (éventuellement vide). Si J est strictement convexe, J admet au plus un minimum.

Cherchons des pistes de recherche d'un minimum :

Def 44: $J : K \to \mathbb{R}$ est dit α-convexe (ou fortement convexe) si, pour un certain $\alpha > 0$, on a : $\forall x, y \in K$, $\forall \theta \in [0;1]$, $J(\theta x + (1-\theta) y) \leq \theta J(x) + (1-\theta) J(y) - \frac{\alpha \theta (1-\theta)}{2} \|x - y\|^2$

[6] **Rem 45**: Une fonction fortement convexe est en particulier convexe.

Prop 46: On suppose ici $J : H \to \mathbb{R}$ C^1. Alors les conditions suivantes sont équivalentes :

① J est α-convexe

② $\forall u, v \in V$, $J(v) \geq J(u) + \langle \nabla J(u), v - u \rangle + \frac{\alpha}{2} \|v - u\|^2$

③ $\forall u, v \in V$, $\langle \nabla J(v) - \nabla J(u), v - u \rangle \geq \alpha \|v - u\|^2$

Rem 47: On a ainsi une caractérisation de la convexité en prenant $\alpha = 0$.

Rem 48: La condition 2 signifie que J est au-dessus de ses plans tangents.

Ex 49: Si $A \in S_n^{++}(\mathbb{R})$, la fonctionnelle quadratique $J : x \mapsto \frac{1}{2} \|x\|_A^2 - \langle x, b \rangle$ est α-convexe où $\alpha = \min_{\lambda \in Sp(A)} \lambda$.

[6] **Thm 50**: (inéquation d'Euler) Soit $J : K \to \mathbb{R}$, C^1 en $u \in K$. Si u est un minimum local de J, alors J satisfait l'inéquation d'Euler : $\forall v \in K$, $\langle \nabla J(u), v - u \rangle \geq 0$.

Rem 51: Si $u \in \mathring{K}$, cette inéquation devient $\nabla J(u) = 0$

Rem 52: Cette inéquation est aussi une condition suffisante pour que u soit minimum local si J est convexe.

Thm 53: Si $J : \mathbb{R}^p \to \mathbb{R}$ est C^1 et α-convexe, alors J admet un unique minimum global.

Thm 54: (Algorithme du gradient à pas optimal) Soit $J : \mathbb{R}^p \to \mathbb{R}$ C^1, α-convexe. Soit $u \in \mathbb{R}^p$. On définit
[7] $(x_n) \in (\mathbb{R}^p)^{\mathbb{N}}$ par :

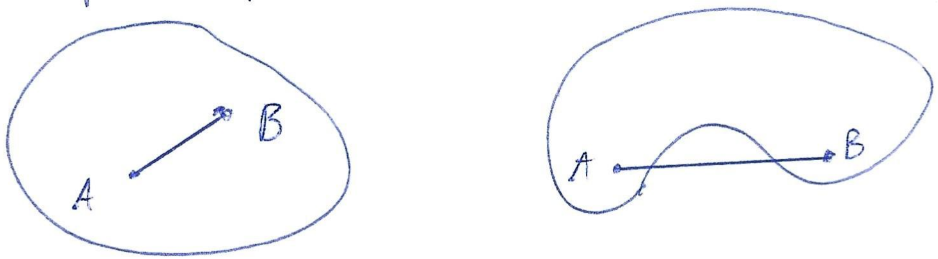
- $x_0 = u$
- $x_{n+1} = x_n$ si $\nabla J(x_n) = 0$
- Sinon, $x_{n+1} = x_n - \alpha_n \nabla J(x_n)$ où $\alpha_n = \underset{\theta > 0}{\operatorname{argmin}} \left(J(x_n - \theta \nabla J(x_n)) \right)$

Alors (x_n) converge vers l'unique minimum global de J.

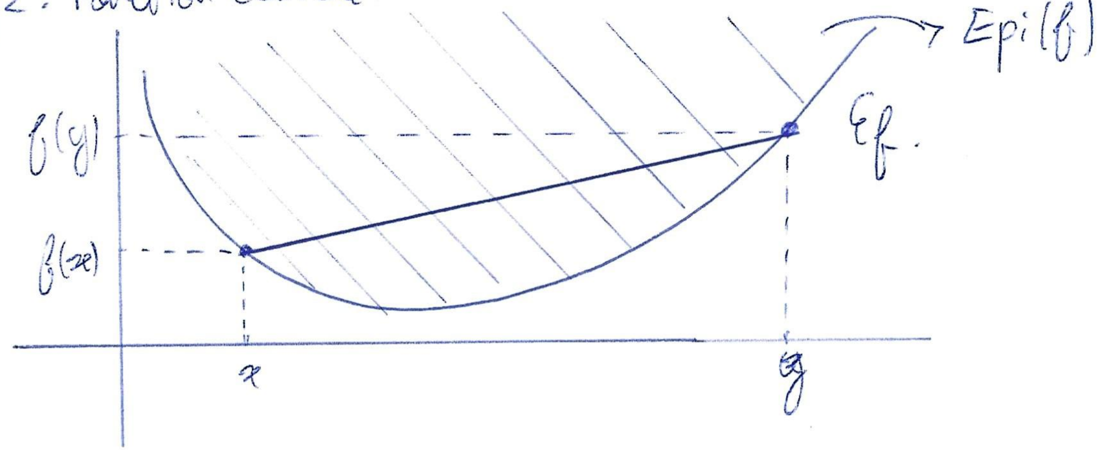
Rem 55: On peut appliquer ceci à la fonctionnelle quadratique afin d'approcher la solution de $Ax = b$.
[5] L'inégalité de Kantorovitch nous permet de plus de donner la vitesse de convergence de l'algorithme : si $\bar{x} = A^{-1} b$ alors :

$$\forall n \in \mathbb{N}, \|x_{n+1} - \bar{x}\| \leq \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \right)^{n+1} \|x_0 - \bar{x}\|$$

* Fig 1 : Exemple d'une partie convexe, et d'une partie non convexe



* Fig 2 : Fonction convexe :



Références :

* Analyse, Gourdon [1]
* Analyse fonctionnelle, Brezis [2]
* Eléments d'analyse fonctionnelle, Hirsch - Lacombe [3]
* Théorie de l'intégration, Briane - Pagès [4]
* Analyse pour l'agreg, Bernis [5]
* Analyse numérique et optimisation, Allaire [6]
* Introduction à l'analyse numérique, Ciarlet [7]