

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 $\forall A, B \in C$, $\forall \lambda \in [0, 1]$, $\lambda A + (1 - \lambda)B \in C$. (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 convexe 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 $\text{conv}(A)$.

Prop 6: $\text{conv}(A)$ est l'ensemble des barycentres convexes des points de A .

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

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

B) Fonctions convexes

Def 9: Soit $f: E \rightarrow \mathbb{R}$. 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 $\text{Epi}(f) = \{(x, y) \in E \times \mathbb{R} \mid y \geq f(x)\} \subset \mathbb{R}^n$ l'épigraph de f . f convexe $\Leftrightarrow \text{Epi}(f)$ convexe (voir Fig 2).

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

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

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

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

Thm 16: (de Bohr-Mollerup) Soit $f: \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ tel que:

① $\forall x \in \mathbb{R}_+^*$, $f(x+1) = x f(x)$

② $f(1) = 1$

③ f est convexe.

Alors $f = \Gamma$ car $\forall x \in \mathbb{R}_+^*$, $\Gamma(x) = \int_0^{x-1} t^{x-1} e^{-t} dt$

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

On suppose que E est un $\mathbb{R}$ -espace vectoriel muni 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 $\exists \varepsilon > 0 \mid \forall x \in A$, $f(x) \leq \alpha - \varepsilon$ et $\forall x \in B$, $f(x) \geq \alpha + \varepsilon$.

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

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

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

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 $G \subseteq E$ et $g: G \rightarrow \mathbb{R}$ vérifiant $g \leq P|_G$ et linéaire.
 Alors il existe $f: E \rightarrow \mathbb{R}$ 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.

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 $F \subseteq E, F \neq E$. Alors il existe $f \in E', f \neq 0$ tel que $f|_F = 0$.

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

II) Inégalités de convexité

A) Inégalités classiques

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

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 $\lambda_{\max} = \max_{\lambda \in \text{Sp}(A)} \lambda, \lambda_{\min} = \min_{\lambda \in \text{Sp}(A)} \lambda$,

$$\frac{\|Ax\|_2}{\|x\|_2} \geq \lambda_{\min} \text{ et } \frac{\|Ax\|_2}{\|x\|_2} \leq \lambda_{\max}$$

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

B) Inégalités dans les espaces L^p

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 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 $f, g \in L^p$. 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 $f \otimes g$ est définie partout et $f \otimes g \in L^\infty$ avec:
 $\|f \otimes g\|_\infty \leq \|f\|_p \|g\|_q$

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. $\text{Re } \mathbb{Z}$

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 caractérisée par: $\forall g \in C, \text{Re}(\langle x - P_C(x), g - P_C(x) \rangle) \leq 0$.

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

Prop 36: Soit $F \subseteq H$ un sous-espace vectoriel fermé. Alors $P_F \in \mathcal{L}(H, F)$. De plus, pour tout x de $H, P_F(x)$ est l'unique élément de F 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) $\forall \phi \in E'$,

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

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

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

[7] Thm 41: (DEV2) (Optimisation dans un Hilbert) Soit $J: H \rightarrow \mathbb{R}$ continue, convexe et coercive (c'est-à-dire: $J(x) \xrightarrow{\|x\| \rightarrow \infty} +\infty$). 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 $J: K \rightarrow \mathbb{R}$ convexe. Tout minimum local est global et l'ensemble de tels points est convexe (factuellement vide). Si J est strictement convexe, J admet au plus un minimum.

Cherchons des pistes de recherche d'un minimum:

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

Rem 45: Une fonction faiblement convexe est en particulier convexe.

Prop 46: On suppose ici $J: H \rightarrow \mathbb{R} C^1$. 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 $J: x \mapsto \frac{1}{2} \|Ax\|^2 - \langle a, x \rangle$ est α -convexe où $\alpha = \min_{\lambda \in Sp(A)} \lambda$.

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

Rem 51: Si $u \in K$, cette équation devient $\nabla J(u) = 0$.

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

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

Thm 54: (Algorithme de gradient à pas optimal) Soit $J: \mathbb{R}^p \rightarrow \mathbb{R}, C^1, \alpha$ -convexe. Soit $u \in \mathbb{R}^p$. On définit $(x_n) \in (\mathbb{R}^p)^N$ par:

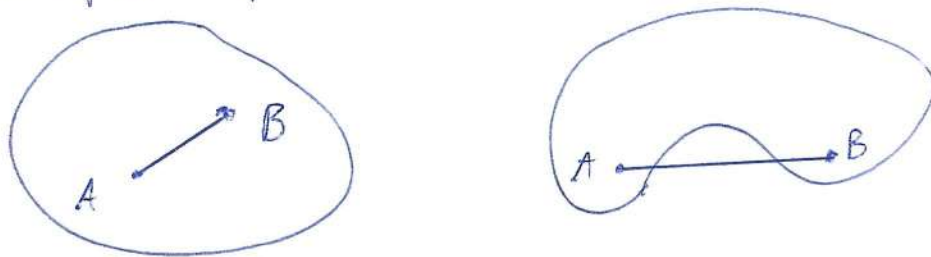
- $x_0 = u$
- $x_{n+1} = x_n$ si $\nabla J(x_n) = 0$
- Sinon, $x_{n+1} = x_n - \alpha_n \nabla J(x_n)$ où $\alpha_n = \arg \min_{\alpha > 0} (J(x_n - \alpha \nabla J(x_n)))$

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

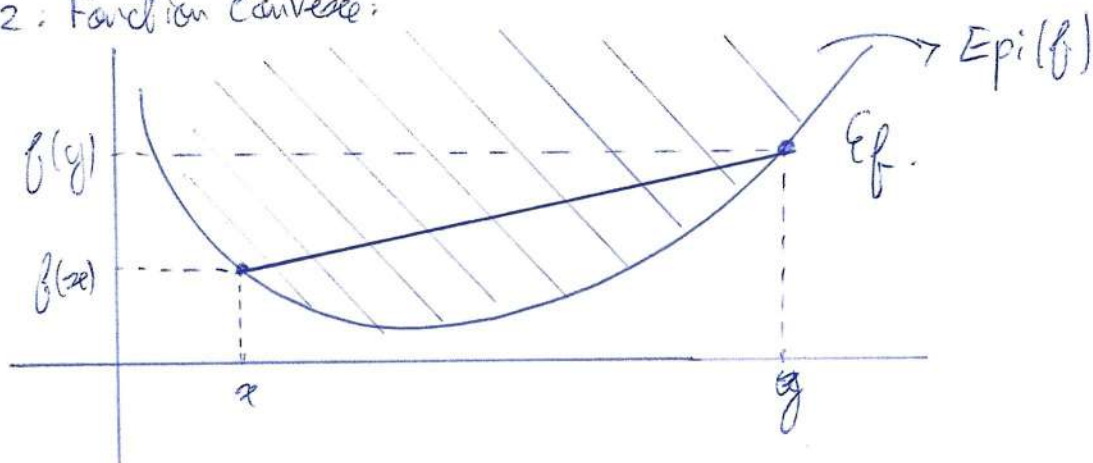
Rem 55: On peut appliquer ceci à la fonctionnelle quadratique afin d'approcher la solution de $Ax=b$. L'inégalité de Kantorovitch nous permet de plus de donner la vitesse de convergence de l'algorithme:

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

④ Fig 1: Exemple d'une partie convexe, et d'une partie non convexe



④ Fig 2: Fonction convexe:



Références:

- ① Analyse, Gourdan [1]
- ② Analyse fonctionnelle, Brezis [2]
- ③ Eléments d'analyse fonctionnelle, Kirsch - Lacombe [3]
- ④ Théorie de l'intégration, Bierme - Pages [4]
- ⑤ Analyse pour l'acqeq, Bernis [5]
- ⑥ Analyse numérique et optimisation, Allaire [6]
- ⑦ Introduction à l'analyse numérique, Ciarlet [7]