253 Utilisation de la notion de convexité en analyse.

Refs : [Gou2] Gourdon Analyse [All] Allaire, Analyse numérique et optimisation
[Brez] Brezis. Analyse Fonctionnelle. [Pom] Pommellet. Cours d'Analyse
[B-P] Briane Pagès : Théorie de l'intégration. [Cia] Ciarlet. Analyse numérique et optimisation
[FGN3] Oraux X-ens AG3 [H-L] Hirsch-Lacombe, analyse fonctionnelle.

Dvpt : 51, 52 (John Loewner)
57 (gradient optimal)
61, 62, 64, 66 (PCF + Riesz)

Par défaut, on se place dans E un $\mathbb{R}$ espace vectoriel normé.

I. Convexité et premières conséquences.

1) Ensembles.

[Gou2] 51-54

Def 1 : Soient $x, y \in E$, on appelle segment d'extrémités x et y, noté $[x;y]$ l'ensemble $\{\theta x + (1-\theta)y \mid \theta \in [0,1]\}$. On note qu'il s'agit d'un ensemble fermé.

Def 2 : Soit $C \subseteq E$, on dit que C est convexe si pour tout $(x,y) \in C^2$, on a $[x;y] \subseteq C$. On dit que C est étoilé par rapport à $x \in C$ si on a $\forall y \in C,\ [x;y] \subseteq C$.

Rq 3 : Un ensemble est convexe si il est étoilé par rapport à tous ses points.

Ex 4 : Tout convexe est connexe par arc donc connexe. La réciproque est fausse (ex : S^1 dans $\mathbb{R}^2$). Dans $\mathbb{R}$, les convexes sont exactement les connexes, et les intervalles. Les boules de E sont aussi convexes.

Prop 5 : L'intersection de convexes de E est un convexe de E.

Def 6 : Soit $A \subseteq E$. L'intersection des convexes contenant A est le plus petit convexe de E contenant A. Il s'agit aussi de l'ensemble des barycentres des points de A à coefficients positifs. On le note $\mathrm{conv}(A)$.

Ex 7 : Pour $n \geq 3$, en notant $\mathbb{U}_n = \{\xi \in \mathbb{C} \mid \xi^n = 1\}$, $\mathrm{conv}(\mathbb{U}_n)$ dans $\mathbb{R}^2$ est un n-gone régulier.

Théo 8 (Carathéodory) Si $\dim E = n < \infty$ non nul. Pour $x \in E$ barycentre de p vecteurs $x_1 \ldots x_p$ de E avec coefficients positifs. Il existe $I \subseteq [1,p]$ de cardinal $\leq n+1$ tel que x soit un barycentre des $(x_i)_{i \in I}$ à coefficients positifs.

App 9 : L'enveloppe convexe d'un compact est un compact si E est de dimension finie.

2) Fonctions convexes, définitions et régularité en dimension 1.

[All] 294

Def 10 : Soit $K \subseteq E$ un ensemble convexe non vide, on dit que $J : K \to \mathbb{R}$ est convexe (resp strictement convexe) si

$$\forall u, v \in K,\ \theta \in [0,1],\quad J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) \qquad (1)$$

(resp $\forall u \neq v \in K$, $\theta \in]0,1[$, $J(\theta u + (1-\theta)v) < \theta J(u) + (1-\theta)J(v)$ et (1))

On dira que J est (strictement) concave si $-J$ est (strictement) convexe.

Rq 11 : La fonctionnelle J est convexe si et seulement si son épigraphe

$$\mathrm{Epi}(J) = \{(x,\lambda) \in K \times \mathbb{R} \mid \lambda \geq J(x)\} \subseteq K \times \mathbb{R}$$

est convexe.

Ex 12 : L'application $\|.\| : E \to \mathbb{R}$ est convexe. L'application $x \mapsto x^3$ sur $\mathbb{R}$ n'est ni convexe ni concave.

[Gou2] 94-95

Pour le reste de cette partie on se place dans le cas où $K = I \subseteq \mathbb{R}$ est un intervalle.

Prop 13 : L'application $f : I \to \mathbb{R}$ est convexe si et seulement si, pour tout $x_0 \in I$, l'application $g_{x_0} : I \setminus \{x_0\} \to \mathbb{R}$ envoyant x sur $\frac{f(x) - f(x_0)}{x - x_0}$ est croissante.

Cor 14 : Si $f : I \to \mathbb{R}$ est convexe et si $a < b < c$ sont des éléments de I, alors

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(a)}{c-a} \leq \frac{f(c)-f(b)}{c-b} \qquad \text{(cf Fig 1).}$$

Prop 15 : Une fonction convexe $f : I \to \mathbb{R}$ possède en tout point de $\mathring{I}$ une dérivée à droite et à gauche (en particulier, f est continue sur $\mathring{I}$). Les applications f'_d et f'_g sont croissantes avec $f'_g \leq f'_d$ sur $\mathring{I}$.

Ex 16 : Tout ceci est perdu sur I, exemple $\mathbb{1}_{\{0,1\}}$ sur $[0,1]$.

Théo 17 : Soit $f : I \to \mathbb{R}$ dérivable sur I, on a équivalence entre
i) f est convexe ii) f' est croissante iii) Le graphe de f est au-dessus de ses tangentes.

Cor 18 : Une application $f : I \to \mathbb{R}$ deux fois dérivable est convexe si et seulement si f'' est positive sur I.

Prop 19 : Soit $f : I \to \mathbb{R}$ convexe. Alors $\forall x_1 \ldots x_n \in I$, $\alpha_1 \ldots \alpha_n > 0$,

$$f\left(\frac{\alpha_1 x_1 + \ldots + \alpha_n x_n}{\alpha_1 + \ldots + \alpha_n}\right) \leq \frac{\alpha_1 f(x_1) + \ldots + \alpha_n f(x_n)}{\alpha_1 + \ldots + \alpha_n}$$

App 21 : Une application convexe $f : \mathbb{R} \to \mathbb{R}$ majorée est constante.

Ex 22 : Faux sur $\mathbb{R}_+$ pour $f : x \mapsto (1+x)^{-1}$.

3) Séparation des convexes.

[Brez] 4-7

Def 23 : Soit $f \in E^*$, on note $[f = \alpha]$ l'hyperplan $\{x \in E \mid f(x) = \alpha\}$ (affine) (pour $f \neq 0$). Il s'agit d'un fermé si et seulement si $f \in E'$.
Soient $A, B \subseteq E$, on dit que $H = [f = \alpha]$ sépare A et B — cf Fig 2.
- au sens large si $f_{|A} \leq \alpha$ et $f_{|B} \geq \alpha$.
- au sens strict si $\exists \varepsilon > 0 \mid f_{|A} \leq \alpha - \varepsilon$ et $f_{|B} \geq \alpha + \varepsilon$.

Théo 24 : (Hahn Banach 1ère forme géométrique) Soient $A, B \subseteq E$ deux ensembles convexes non vides et disjoints, avec A ouvert. Alors il existe un hyperplan fermé qui sépare A et B au sens large.

Théo 25 (Hahn Banach 2e forme géométrique) Soient $A, B \subseteq E$ convexes non vides disjoints, A fermé et B compact. Alors il existe un hyperplan fermé qui sépare A et B au sens strict.

Cor 26 : Soit $F \subseteq E$ un sous espace vectoriel tel que $F \neq E$. Alors il existe $f \in E'$ non nulle telle que $f_{|F} \equiv 0$.

Rq 27 : Si E est de dimension finie, on peut toujours séparer au sens large deux convexes disjoints non vides.

Ex 28 : Pas toujours strictement : $\{(x,y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y \geq \frac{1}{x}\}$
$\{(x,y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y \leq \frac{-1}{x}\}$

II. Inégalités de convexité.

1) Inégalités classiques

[Rom] 134 136

Prop29: Pour tout $x,y\in\mathbb{R}$, on a $e^x \geq x+1$ et $e^{\frac{x+y}{2}} \leq \frac{1}{2}(e^x+e^y)$.

Prop30: Si $p,q\in\mathbb{R}_+^*$ sont tels que $p^{-1}+q^{-1}=1$, alors $\forall u,v\in\mathbb{R}_+^*$, $uv \leq \frac{1}{p}u^p+\frac{1}{q}v^q$.

Prop31: Sur $]0,\frac{\pi}{2}[$, on a $\frac{2}{\pi}x < \sin x < x$.

Théo32 (Jensen). Soit $f:\mathbb{R}\to\mathbb{R}$ convexe et $\varphi\in[a,b]\to\mathbb{R}$ réglée, on a
$$f\left(\frac{1}{b-a}\int_a^b \varphi(t)dt\right) \leq \frac{1}{b-a}\int_a^b f\circ\varphi(t)dt.$$

Prop33 (Inégalité arithmético-géométrique). Pour toute famille finie $(x_i)_{i\in[\![1,n]\!]}$ de $\mathbb{R}_+^*$, on a $\left(\prod_{i=1}^n x_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n x_i$; l'inégalité est stricte si et seulement si les x_i ne sont pas tous égaux.

2) Inégalités dans les espaces de Lebesgue.

[BP] 153 158

Prop34 (Hölder) Soient $r,p,q>0$ tels que $p^{-1}+q^{-1}=r^{-1}$, $\Omega\subseteq\mathbb{R}^n$ un ouvert, $f,g:\Omega\to\mathbb{R}$ deux applications mesurables, on a $\|fg\|_r \leq \|f\|_p\|g\|_q$ (vrai aussi pour $p=r=1$, $q=\infty$).

Prop35 (Minkowski) Si $\Omega\subseteq\mathbb{R}^n$ un ouvert et $f,g:\Omega\to\mathbb{R}$ mesurables, $p\in[1,+\infty]$, on a $\|f+g\|_p \leq \|f\|_p+\|g\|_p$. Les espaces de Lebesgue sont des espaces vectoriels normés.

Prop36 (Inégalité de Young pour la convolution) Si $1+r^{-1}=p^{-1}+q^{-1}$, on a $\|f\star g\|_r \leq \|f\|_p\|g\|_q$. Donc L^1 muni de $\star$ est une $\mathbb{R}$-algèbre.

Prop37 Si $f\in L^p\cap L^q$, alors $f\in L^r$ pour $r\in[p,q]$ $(p,q\in[1,\infty])$.

Appl38: Si X est de mesure finie, alors $p>q \Rightarrow L^p\subset L^q$.

+ de détails

3) Inégalités en probabilité.

Prop39 (Jensen) Soit $(\Omega,\mathcal{A},\mathbb{P})$ probabilisé, X une v.a.r intégrable, $f:\mathbb{R}\to\mathbb{R}$ convexe alors $f(E(X)) \leq E(f(X))$.

Théo40: (Hoeffding). Soient $Z_1\dots Z_n$ des v.a.r i.i.d sur $(\Omega,\mathcal{A},\mathbb{P})$. Si $\exists a<b$ tq $a\leq Z_1\leq b$ $\mathbb{P}$-presque sûrement. Alors $\forall t>0$ $\mathbb{P}\left(\left|\sum_{i=1}^n Z_i - E(Z_i)\right|\geq t\right) \leq 2\exp\left(-\frac{2t^2}{n(b-a)^2}\right)$.

Appl41 Construction d'intervalles de confiance.

III Applications à l'optimisation.

1) Fonctions convexes et extrema.

[APe] 294 295

Def42: On dit que $f:K\to\mathbb{R}$ est α-convexe (ou α-elliptique) pour $\alpha>0$, $K\subseteq E$ convexe si $\forall x,y\in K$, $\theta\in[0,1]$ $J(\theta x+(1-\theta)y) \leq \theta J(x)+(1-\theta)J(y) - \alpha\frac{\theta(1-\theta)}{2}\|x-y\|^2$.

Rq43: L'α-convexité est plus forte que la convexité stricte : La fonction $x\mapsto\frac{1}{x}$ sur $\mathbb{R}_+^*$ est strictement convexe et non α-convexe.

Ex44 $x\mapsto\frac{1}{2}(Ax,x)-(b,x)$ où A est sdp est α-convexe, où $\alpha=\min|\mathrm{Spec}\,A|$.

[All] 305 307

Prop44: Soit $f:K\to\mathbb{R}$ différentiable sur K (que l'on suppose ouvert), on a équivalence entre
i) f convexe sur K ii) $\forall x,y\in K$ $f(y)\geq f(x)+Df_x(y-x)$ iii) $\forall x,y\in K$, $(Df(x)-Df(y),(x-y))\geq 0$
on a équivalence en rajoutant strictement convexe dans i) et des inégalités strictes dans ii) et iii). Pour $\alpha>0$ on a l'équivalence entre
i) f α-convexe sur K ii) $\forall x,y\in K$ $f(y)\geq f(x)+Df(x)(y-x)+\frac{\alpha}{2}\|y-x\|^2$ iii) $\forall x,y\in K$ $(Df(x)-Df(y),x-y)\geq\alpha\|y-x\|^2$

Prop45: Avec les notations précédentes, si f est deux fois différentiable, la première série d'équivalence est équivalente à $D^2f(x)(z)(z)\geq 0$ pour $x,z\in K$, et la seconde est équivalente à $D^2f(x)(z)(z)\geq\alpha\|z\|^2$ pour $x,z\in K$.

294

Prop46. Si $f:K\to\mathbb{R}$ est convexe, tout minimum global de f est en fait global, l'ensemble des points réalisant ce minimum est convexe (éventuellement vide). Si de plus f est strictement convexe elle admet au plus un minimum global.

291

Ex47: On n'a pas forcément existence : soit $J:\ell^2(\mathbb{R})\to\mathbb{R}$ définie par $J(x)=(\|x\|^2-1)^2+\sum_{i=0}^{\infty}\frac{x_i^2}{i+1}$ alors J n'admet pas de minimum global.

298.

Théo47: Si K est un convexe fermé non vide de E un Hilbert, $f:K\to\mathbb{R}$ continue coercive et convexe. Alors f admet un minimum global sur K.

307.

Théo48: Soit $u\in K$ un convexe, et $f:K\to\mathbb{R}$ différentiable en u. Si u est un minimum global de f sur K alors $Df(u)(v-u)\geq 0$ pour $v\in K$. Si de plus f est convexe, la réciproque est vraie (u est alors un minimum global de f sur K).

Appl49 Résoudre $\frac{1}{2}(Ax,x)-(b,x)$ minimum équivaut à trouver la solution de $Ax=b$ pour A sdp.

310.

Prop50: Si $f:E\to\mathbb{R}$ est deux fois différentiable en $x\in E$. Si x est un minimum local de f, alors $Df(x)=0$, et $D^2f(x)(z,z)\geq 0$ pour tout $z\in E$. Réciproquement si $Df(x)=0$ et $D^2f(y)(z,z)\geq 0$ pour tout $z\in E$ sur un voisinage de x, alors x est minimum local de f.

[FGN3] 229

Lemme51: Soient $A,B\in\mathcal{S}_n(\mathbb{R})$ des matrices symétriques définies positives, et $\alpha,\beta\geq 0$ tels que $\alpha+\beta=1$ on a
$$\det(\alpha A+\beta B)\geq(\det A)^{\alpha}(\det B)^{\beta}$$

DVP

Théo52 (Ellipsoïde de John Loewner) Soit $K\subseteq\mathbb{R}^n$ un compact d'intérieur non vide, alors il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.

2) Algorithmes de descente en dimension finie sans contrainte.

[Gia] 184 188

On fixe $J:\mathbb{R}^n\to\mathbb{R}$, on approche $I=\inf_{x\in\mathbb{R}^n}J(x)$. On met en œuvre des méthodes itératives de la forme $u^0\in\mathbb{R}^n$, $u^{n+1}=u^n+\rho^n d^n$. On appelle d^n la direction de descente et ρ^n le pas de descente. Une première idée est de prendre pour d^n successivement les vecteurs de base canonique : c'est la méthode de relaxation. Si u^k est construit, on pose $u_1^{k+1}\in\mathbb{R}$ tel que $J(u_1^{k+1},u_2^k,\dots,u_n^k)=\inf J(x,u_2^k,\dots,u_n^k)$, et ainsi de suite $u_i^{k+1}\in\mathbb{R}$ réalisant $\inf J(u_1^{k+1}\dots u_{i-1}^{k+1},x,\dots u_n^k)$.

Théo53: Si J est α-convexe et différentiable, la méthode de relaxation converge.

[Ci2] 188 189

Ex54 : Dans le cas d'une fonctionnelle quadratique donnée par A et b, la méthode de relaxation est en fait la méthode de Gauss-Seidel pour résoudre $Ax=b$.
On peut essayer d'être plus astucieux sur le choix de la direction d_k.

Déf55 : Avec les notations précédentes, pour $x \in \mathbb{R}^m$. On dit que $d \in \mathbb{R}^m$ est une direction de descente (pour J en x) si $\exists \eta > 0 \,(\forall \rho \in [0,\eta] \; J(x+\rho d) \leq J(x)$.
On se place dans le cas où J est différentiable, alors $-\nabla J(x)$ donne toujours une direction de descente. Il reste à fixer le pas pour obtenir les méthodes de gradient. Si ρ_k ne dépend pas de k, c'est le gradient à pas fixe.

[ll] 34 – 68

Théo56 : Si J est α-convexe différentiable sur E et ∇J est L-lipschitzien sur E. Alors pour $0<\rho<2\alpha/L^2$ l'algorithme du gradient à pas fixe converge.
On peut vouloir accélérer la convergence en optimisant le pas. C'est le gradient à pas optimal. Le pas ρ_k est choisi comme réalisant le minimum $J(u^k + x\nabla J(u^k))$, potentiellement plus simple à résoudre car à une seule variable.

Théo57 : Si J est α-convexe différentiable, et si ∇J est lipschitzien sur tout ensemble borné de E. Alors l'algorithme du gradient à pas optimal converge. En particulier le problème intermédiaire est toujours bien posé. DVP

Un autre choix de direction peut être donné quand J est deux fois différentiable par $-(D^2 J(u^k))^{-1}$. On retrouve l'algorithme de Newton pour calculer les points d'annulation de ∇J.

Théo58 : Soit J de classe C^2 sur $\mathbb{R}^m$, et u tel que $\nabla J(u)=0$ et $\nabla^2 J(u)$ est inversible. Il existe B une boule centrée en u telle que si $u^0 \in B$, la méthode de Newton converge. Plus précisément il existe $C>0$ telle que $\|u^{n+1}-u\| \leq C\|u^n - u\|^2$.
Cette méthode a deux défauts, il faut connaître B en pratique et inverser un système linéaire à chaque itération.

Ex59 Pour la fonctionnelle quadratique, le calcul d'une itération correspond déjà à résoudre le système linéaire concerné.

Appl60 : En dimension 1, on retrouve la méthode de Newton classique. En particulier la méthode de Syracuse pour la recherche des racines carrées.

3) Cas des Espaces de Hilbert.

On suppose ici que $E=H$ est un espace de Hilbert sur $\mathbb{R}$, et $\langle .,. \rangle$ son produit scalaire.

Théo61 : Soit $\emptyset \neq C \subseteq H$ un convexe fermé. Alors pour tout $x \in H$, il existe un unique $y \in C$ réalisant la distance $d(x,C)$. On note $p_C(x)$ ce point, le projeté de x sur C. Il est caractérisé par la propriété

$$\forall z \in C, \; (y-z, y-x) \leq 0.$$

cf (Fig 2) DVP

Cor62 : Si $F \leq H$ est un sous espace vectoriel fermé. Alors P_F est linéaire et la caractérisation ci-dessus devient $y \in F$ et $x-y \in F^\perp$.

[Gou2] ... cette partie on se place dans le cas où $K = I \subset \mathbb{R}$ et un ...

Cor64 : Si $F \in H$ est un sous espace vectoriel fermé, alors $H = F \oplus F^\perp$.

Cor65 : Un sous espace F de H est dense si et seulement si $F^\perp = \{0\}$.

[H-L] 93, 96

Théo66 (Riesz) : Pour $f \in H'$, il existe un unique $y \in H$ telle que $f(x) = (x,y)$ pour tout $x \in H$. On a de plus $\|f\| = \|y\|$.

Théo67 (Lax-Milgram) Soit $L \in H'$, $a : H \times H \to \mathbb{R}$ une forme bilinéaire continue coercive sur H ($\exists \lambda > 0 \mid a(x,x) \geq \lambda\|x\|^2$ pour $x \in H$). Alors il existe un unique $x \in H$ tel que $a(x,y) = L(y)$ pour $y \in H$.

[All] 65 78

Rq68 : C'est une généralisation de Riesz pour $a = \langle .,. \rangle$ et $L = f$.

Cor69 : Dans les notations et hypothèses du théorème 69. Si a est symétrique, l'élément $x \in H$ est l'unique minimum globale de la fonctionnelle quadratique définie par

$$J(x) = \frac{1}{2} a(x,x) - L(x).$$

Application 70. Équation de Laplace.

On considère $\Omega \subseteq \mathbb{R}^n$ un ouvert convexe. On considère les solutions de l'équation suivante

$\Delta u = f$ sur Ω
$u = 0$ sur $\partial\Omega$ (E)

Posons $V = \{v \in C^1(\Omega) \mid v_{|\partial\Omega} = 0\}$. Si u est solution de E, alors on a

$$\forall v \in V, \int_\Omega \nabla u(x) \nabla v(x) dx = \int_\Omega f(x) v(x) dx.$$

C'est la formulation variationnelle du problème E à laquelle nous voulons appliquer le théorème de Lax-Milgram. Or, V muni du produit scalaire $(u,v) = \int \nabla u \nabla v \, dx$ n'est pas un espace de Hilbert, nous devons nous placer dans l'espace de Sobolev $H^1_0(\Omega)$.
Dans cet espace, on obtient l'existence et l'unicité de la solution u du problème variationnel.

[?] 7.

En travaillant successivement sur des sous espaces vectoriels de dimension finie de $H^1_0(\Omega)$, on construit des méthodes de résolutions aux éléments finis.

Fig 1:

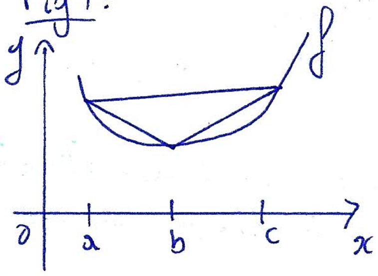

Fig 2: