253 – Utilisation de la notion de convexité en analyse.

1 Fonctions convexes à variable réelle

Définition 1. Une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) est dite convexe si $\forall (x,y) \in I^2$, $\forall \lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$. Elle est dite convave si l'inégalité est dans l'autre sens.
Concrètement, cela signifie que f est au dessus de ses cordes.
La convexité est dite stricte si l'inégalité est stricte.

Remarque 2. f est convexe $\iff$ $-f$ est concave.
Exemple 3. Faire un dessin de fonction convexe et concave
Exemple 4. $x \mapsto x^2$ est convexe, $y \mapsto ln(y)$ est concave.

Proposition 5. *Une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) est convexe $\iff$ $\forall x_0 \in I$, $g_{x_0} : \begin{array}{ccc} I & \to & \mathbb{R} \\ x & \mapsto & \frac{f(x)-f(x_0)}{x-x_0} \end{array}$ est croissante.*

Corollaire 6. *Cela fournit l'inégalité suivante pour f convexe et $a,b,c \in I$ tel que $a < b < c$:*

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

Elle représente la croissance des cordes pour une fonction convexe. Faire un shéma à côté.

Corollaire 7. *Soit une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) convexe*

- *f possède en tout point de l'intérieur de I une dérivée à gauche et à droite.*
- *f est continue à l'intérieur de I*
- *Les applications f'_d et f'_g sont croissantes sur l'intérieur de I, et $f'_g(x) \leq f'_d(x)$.*

Théorème 8. *Soit une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) dérivable sur I. Alors : f est convexe $\iff$ f' est croissante sur I $\iff$*

Corollaire 9. *Soit une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) deux-fois dérivables sur I. Alors f est convexe $\iff$ $f''(x) > 0 \forall x \in I$.*

Application 10 (Méthode de Newton). *Soit $f : [c,d] \to \mathbb{R}$ de classe C^2. On suppose que $f(c) < 0 < f(d)$ et $f'(x) > 0$ sur $[c,d]$. On considère la suite définie par réccurence : $x_0 \in [c,d]$, $x_{n+1} = F(x_n)$ où $F : t \mapsto t - \dfrac{f(t)}{f'(t)}$. Alors :*

- *Faire un dessin en **ANNEXE**.*
- *f admet un unique zéro en a sur $[c,d]$. Il existe $\alpha > 0$ tel que $[a-\alpha, a+\alpha] = I$ soit F-stable et que $\forall x_0 \in [a-\alpha, a+\alpha] = I$, la suite $(x_n)_n$ converge quadratiquement vers a.*
- *Si de plus f est convexe sur $[c,d]$, alors le résultat précédent est valable pour $I = [a,d]$, et de plus $(x_n)_n$ est décroissante (ou constante) vers a et on à l'équivalent : $x_{n+1} - a \sim \dfrac{f''(a)}{2f'(a)}(x_n - a)^2$.*

2 Des inégalités classiques

2.1 Des inégalités sur $\mathbb{R}^n$

Proposition 11. *Soit une fonction $f : I \to \mathbb{R}$ (I intervalle de $\mathbb{R}$) convexe. Alors $\forall (x_1,...,x_n) \in I^n$, $\forall (\alpha_1,...,\alpha_n) > 0$, on à $f(\dfrac{\alpha_1 x_1 + ... + \alpha_n x_n}{\alpha_1 + ... + \alpha_n}) \leq \dfrac{\alpha_1 f(x_1) + ... + \alpha_n f(x_n)}{\alpha_1 + ... + \alpha_n}$.*

Corollaire 12 (Inégalité arithmético-géométrique). *Soient $(x_1,...,x_n)$ des nombres réels positifs. Alors $(x_1...x_n)^{\frac{1}{n}} \leq \dfrac{x_1 + ... + x_n}{n}$*

Proposition 13 (Inégalité de Hölder). *Soient deux nombres positifs p et q tel que $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Soient $(a_1,...,a_n)$ et $(b_1,...,b_n)$ des réels positifs. Alors : $\sum_{k=0}^{n} a_k b_k \leq (\sum_{k=0}^{n} a_k^p)^{\frac{1}{p}} (\sum_{k=0}^{n} b_k^q)^{\frac{1}{q}}$*

Proposition 14 (Inégalité de Minkowski). *Soient $1 \neq p$ réel et $(a_1,...,a_n)$ et $(b_1,...,b_n)$ des réels positifs. Alors : $(\sum_{k=0}^{n} (a_k + b_k)^p)^{\frac{1}{p}} \leq (\sum_{k=0}^{n} a_k^p)^{\frac{1}{p}} (\sum_{k=0}^{n} b_k^p)^{\frac{1}{p}}$*

Remarque 15. Cela permet de démontrer que $N(a_1,...,a_n) = (\sum_{k=0}^{n} a_k^p)^{\frac{1}{p}}$ définit une norme sur $\mathbb{R}^n$.

2.2 Des inégalités dans les espaces $L^p(\Omega, A, \mu)$

Cadre : Soit (Ω, A, μ) un espace mesuré.

Théorème 16 (Minkowski). *Pour $f, g \in L^p(\Omega, A, \mu)$, $||f+g||_p \leq ||f||_p + ||g||_p$*

Corollaire 17. *$L^p(\Omega, A, \mu)$ est un espace vectoriel.*

Théorème 18 (Hölder). *Soit $f \in L^p(\Omega, A, \mu)$ et $g \in L^q(\Omega, A, \mu)$ ($\frac{1}{p}+\frac{1}{q}=1$), alors* $\int_\Omega \leq ||f||_p ||g||_q$

Application 19 (Convolution). — *Si f est dans $L^p(\mathbb{R}, B(\mathbb{R}, \mu))$ et g dans $L^1(\mathbb{R}, B(\mathbb{R}, \mu))$,alors le produit de convolution est bien définie PRESQUE-partout et il appartient à $L^p(\mathbb{R}, B(\mathbb{R}, \mu))$ et $||f \star g||_p \leq ||f||_1 ||g||_p$ (dans Lp et L1 respectivement)*

— *Si f est dans $L^p(\mathbb{R}, B(\mathbb{R}, \mu))$ et g dans $L^q(\mathbb{R}, B(\mathbb{R}, \mu))$ (q l'exposant conjugué de $p \in [1, +\infty[$), alors le produit de convolution est bien définie partout et il est bilinéaire, et $||f \star g||_\infty \leq ||f||_p ||g||_q$*

Théorème 20 (Inégalité de Jensen). *Soit $\varphi : \mathbb{R} \to \mathbb{R}^+$ une fonction convexe, alors $\varphi(\int_\Omega f d\mu) \leq \int_\Omega \varphi(f(x)) d\mu(x)$*

Application 21 (Jensen en probabilité). *Soit X une variable aléatoire d'un espace probabilisé dont l'espérance existe. Alors $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$*

3 Optimisation et convexité

3.1 Optimisation sur $\mathbb{R}^n$

Proposition 22. *Soit $f : I \to \mathbb{R}$*

— *Si f est convexe, tout extremum local est global.*

— *Si f est strictement convexe, f admet un unique extremum global.*

Définition 23. Une fonction $U \to \mathbb{R}$ est dite convexe sur U convexe ouvert de $\mathbb{R}^n$ si, pour tout, $x, y \in U$, $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$

Proposition 24. *Soit U ouvert connexe de $\mathbb{R}^n$.On suppose que f est différentiable sur U. On à l'équivalence : f est convexe $\iff$ $\forall (x,y) \in U^2$, $Df(x).(y-x) \leq f(y-x)$.*

Proposition 25. *Si $f : U \to \mathbb{R}$ (U ouvert connexe de $\mathbb{R}^n$) est convexe, alors f admet un extrema global en $a \iff a$ est un point critique de f.*

Proposition 26. *Soit U ouvert connexe de $\mathbb{R}^n$. On suppose que f est 2-fois différentiable sur U. On à l'équivalence : f est convexe $\iff$ $\forall x \in U$, $D^2 f(x)$ est une forme quadratique positive.*

Proposition 27. *Soit $f : U \to \mathbb{R}$ deux-fois différentiable. Alors f admet un extrema relatif en $a \iff a$ est un point critique de f et $\forall x \in U$, $D^2 f(x)$ est une forme quadratique positive.*

Lemme 28 (Inégalité de Kantorovitch). *Soit $A \in S_n^{++}(\mathbb{R})$, alors $\forall x \in \mathbb{R}^n$, $< Ax|x >< A^{-1}x|x > \leq \frac{1}{4}\frac{(\lambda_{max}+\lambda_{min})^2}{\lambda_{max}\lambda_{min}}$*

Théorème 29 (Algorithme du gradient à pas optimal). *Objectif : On veut résoudre le système linéaire de taille n : $Ax = b$, avec $A \in S_n^{++}(\mathbb{R})$.*

— *Ce système admet un unique minimum x^*. x est solution de ce système $\iff$ x^* est l'unique minimum de la fonctionnelle $\Phi \begin{array}{l} \mathbb{R}^n \to \mathbb{R} \\ x \mapsto \frac{1}{2} < Ax|x > - < b|x > \end{array}$*

— *L'algorithme suivant : $x_0 \in \mathbb{R}^n$, $r_0 = Ax_0 - b$, puis $x_{n+1} = x_n + \alpha_k r_k$, $r_{k+1} = r_k - A\alpha_k r_k$, $\alpha_k = \frac{||r_k[||^2}{< Ar_k|r_k >}$ fournit une suite $(x_k)_k$ qui vérifie : $||x_k - x^*|| \leq C(\frac{\lambda_{min}-\lambda_{max}}{\lambda_{min}+\lambda_{max}})^k$, et qui converge donc bien vers x^*.*

3.2 Optimisation sur un Hilbert

Théorème 30 (Projection sur un convexe fermé). *Soit C un partie convexe et fermé non-vide de H espace de Hilbert. Alors pour tout $x \in H$, il existe un unique $y = P_C(x)$ tel que : $dist(x, C) = ||x - y||$. On dit que $y = P_C(x)$ est la projection de x sur C. On à la caractérisation suivante pour y dans H :*

$$y = P_C(x) \iff < x - y, z - y > \leq 0 \quad \forall z \in C$$

(dans le cas réel) On fait un dessin de la caractérisation.

Corollaire 31. *Si F est un sev fermé de H, on a :*

$$y = P_C(x) \iff < x - y, z >= 0 \quad \forall z \in F$$

Corollaire 32. *Si F est un sev fermé de H, on a : $H = F \oplus F^\perp$ et la projection sur F parallèlement à $F^\perp$ est P_F. On dit que P_F est la projection orthogonale sur F.*

Remarque 33. Ces résultats restent valable pour un espace euclidien, mais la démonstration est plus simple.

Application 34 (Problème des moindres carrés). *Soit $A \in \mathbb{M}_{n,p}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche les solutions $x \in \mathbb{R}^p$ qui vérifient $||Ax - b|| = \min_{y \in \mathbb{R}^p} ||Ay_b||$.*

— *Ce problème possède un minimum $\bar{x}$, unique à $ker(A)$-près.*

— *x est solution du problème $\iff$ x vérifie ${}^tAAx = {}^tAb$.*

*C'est le problème rencontré lors du problème de régréssion linéaire (avec p = 2) : on cherche (a_0, a_1) qui minimise la quantité : $\sum_{k=1}^{m} |y_k - (a_0 + a_1 t_i)|^2$. On fait un dessin en **ANNEXE**.*

Théorème 35 (de représentation de Riesz). *Soit $\varphi : H \to \mathbb{R}$, alors il existe un unique $x \in H$ tel que : $\forall y \in H$, $\varphi(y) = < x|y >$.*