Cadre : I est un intervalle de $\mathbb{R}$ non réduit à un singleton, et C un convexe d'un espace vectoriel E.

I Fonctions monotones

1) Définition et premières propriétés

Définition 1. Une application $f : D \to \mathbb{R}$, où $D \subseteq \mathbb{R}$, est dite croissante (resp. strictement croissante) lorsque :

$$\forall x, y \in D, x < y \Rightarrow f(x) \leqslant f(y) \quad (\text{resp. } x < y \Rightarrow f(x) < f(y))$$

f est dite (strictement) décroissante si $-f$ est (strictement) croissante.
f est dite (strictement) monotone si f est (strictment) croissante ou (strictement) décroissante.

Exemple 2. *$x \mapsto \frac{1}{x}$ est décroissante sur $\mathbb{R}^{-\star}$ et $\mathbb{R}^{+\star}$, mais pas sur $\mathbb{R}^\star$.*

Proposition 3. *Une application monotone est injective si, et seulement si, elle est strictement monotone.*

Proposition 4.
- *(i) Le produit d'une application monotone par un scalaire positif est monotone de même variation.*
- *(ii) La somme de deux applications croissantes est croissante.*
- *(iii) Le produit de deux fonctions croissantes positives est croissant.*
- *(iv) La composée d'applications croissantes ou décroissantes est croissante. La composée d'une application croissante et d'une application décroissante est décroissante.*

Remarque 5. *L'ensemble des fonctions monotones n'est pas un espace vectoriel. L'espace engendré par les fonctions monotones est l'espace des fonctions à variations bornées.*

2) Monotonie, limites et continuité

Théorème 6. *Soient $f : D \to \mathbb{R}$ monotone, où $D \subset \mathbb{R}$, et $a \in \overline{\mathbb{R}}$ un point adhérent à $D \cap]a, +\infty[$ (resp. $D \cap]-\infty, a[$). Alors f admet une limite à droite (resp. à gauche) au point a.*

Corollaire 7. *Soient $f : D \to \mathbb{R}$ croissante et $a \in \overline{\mathbb{R}}$ un point adhérent à $D \cap]a, +\infty[$. Alors f admet une limite finie à droite, au point a si, et seulement si, f est minorée sur $D \cap]a, +\infty[$.*

Théorème 8. *Soit $f : I \to \mathbb{R}$ monotone. L'ensemble des points de discontinuité de f est au plus dénombrable.*

Exemple 9. *Posons $u_n(x) = \frac{1}{2^n}\mathbb{1}_{\{\varphi(n)<x\}}$, où φ est une bijection de $\mathbb{N}$ sur $\mathbb{Q} \cap]0, 1[$. Soit $f(x) = \sum_n u_n(x)$. Alors f est strictement croissante sur $[0, 1]$ et discontinue en tout point de $\mathbb{Q} \cap]0, 1[$.*

Théorème 10. *Soit $f : I \to \mathbb{R}$ monotone. f est continue sur I si, et seulement si, $f(I)$ est un intervalle.*

Corollaire 11. *Soit $f : I \to \mathbb{R}$ continue et strictement monotone. Alors $J = f(I)$ est un intervalle et f induit un homéomorphisme de I sur J. Réciproquement, un homéomorphisme entre deux intervalles est une fonction strictement monotone.*

Exemple 12. *La fonction sinus induit un homéomorphisme de $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ sur $[-1; 1]$ croissant, dont la réciproque est l'arcsinus.*

3) Monotonie et dérivabilité

Théorème 13. *Soit $f : I \to \mathbb{R}$ continue, et dérivable à droite sur $\mathring{I}$.*

- *(i) f est constante si, et seulement si : $\forall t \in \mathring{I}, f'_d(t) = 0$*
- *(ii) f est croissante si, et seulement si : $\forall t \in \mathring{I}, f'_d(t) \geqslant 0$*
- *(iii) f est décroissante si, et seulement si : $\forall t \in \mathring{I}, f'_d(t) \leqslant 0$*

Théorème 14. *Soit $f : I \to \mathbb{R}$ continue sur I et dérivable à droite sur $\mathring{I}$. f est strictement croissante sur I si, et seulement si, $f'_d \geqslant 0$ et l'ensemble $\left\{t \in \mathring{I} \,\middle|\, f'_d(t) = 0\right\}$ est d'intérieur vide.*

Exemple 15. *$t \mapsto t^3$ est strictement croissante sur $\mathbb{R}$, mais sa dérivée $t \mapsto 3t^2$ s'annule sur $\mathbb{R}$.*

Théorème 16. *Une fonction monotone est dérivable presque partout.*

Remarque 17. *L'escalier de Cantor est une fonction monotone dérivable presque partout, de dérivée nulle presque partout, sans être constante.*

II Fonctions convexes

1) Définition et premières propriétés

Définition 18. On dit que la fonction $f : C \to \mathbb{R}$ est convexe lorsque, pour tous $a, b \in C$ et tout $\lambda \in [0,1]$, on a :

$$f((1-\lambda)a + \lambda b) \leqslant (1-\lambda)f(a) + \lambda f(b)$$

On dit que f concave si $-f$ est convexe. Lorsque l'inégalité est stricte pour $a \neq b$ et $0 < \lambda < 1$, f est strictement convexe. Pour $\alpha > 0$, on dit que f est α-convexe si pour tous $a, b \in C$ distincts et tout $\lambda \in]0,1[$, on a :

$$f((1-\lambda)a + \lambda b) \leqslant (1-\lambda)f(a) + \lambda f(b) - \frac{\alpha}{2}\|a-b\|^2 \lambda(1-\lambda)$$

Remarque 19. *L'α-convexité implique la stricte convexité, qui implique la convexité.*

Remarque 20. *Une fonction $f : C \to \mathbb{R}$ est convexe si l'ensemble $\{(x,y) \in C \times \mathbb{R} \mid y \geqslant f(x)\}$ est convexe.*

Exemple 21. *L'application $x \mapsto \|x\|$ est convexe*

Théorème 22. *Une fonction $f : C \to \mathbb{R}$ est convexe si, et seulement si, pour tous $x, y \in C$, $t \mapsto f((1-t)x + ty)$ est convexe sur $[0,1]$.*

Proposition 23.
- *(i) Une combinaison linéaire à coefficients réels positifs de fonctions convexes est convexe.*
- *(ii) La composée d'une fonction convexe croissante avec une fonction convexe est convexe.*
- *(iii) Une limite simple de fonctions convexes est convexe.*
- *(iv) Le maximum de deux fonctions convexes est convexe.*

Remarque 24. *Le produit de deux fonctions convexes n'est pas nécessairement convexe ($-x \cdot x^2 = x^3$), et leur composition non plus ($(x \mapsto -x) \circ (x \mapsto x^2) = (x \mapsto -x^2)$).*

Proposition 25. *Une fonction convexe $f : I \to \mathbb{R}$ possède en tout point de $\mathring{I}$ une dérivée à droite et une dérivée à gauche. De plus, les applications f'_d et f'_g sont croissantes sur $\mathring{I}$, et $f'_g \leqslant f'_q$.*

Corollaire 26. *Une fonction convexe sur I est continue sur $\mathring{I}$.*

2) Caractérisation des fonctions convexes

En dimension 1

Théorème 27. *Pour $f : I \to \mathbb{R}$, il y a équivalence entre :*
- *(i) f est convexe sur I.*
- *(ii) Pour $a < b < c$ dans I, on a : $\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b}$*
- *(iii) Pour $a \in I$, la fonction $x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $I \setminus \{a\}$.*

Corollaire 28. *Une fonction de $\mathbb{R}$ dans $\mathbb{R}$ est affine si, et seulement si, elle est convexe et concave.*

Théorème 29. *Soit $f : I \to \mathbb{R}$ dérivable. Il y a équivalence entre :*
- *(i) f est (strictement) convexe sur I.*
- *(ii) La fonction dérivée f' est (strictement) croissante.*
- *(iii) La courbe représentative de f est située (strictement) au-dessus de sa tangente en tout point de I.*

Proposition 30. *Si f est deux fois dérivable sur I, elle est alors convexe si, et seulement si, $f'' \geqslant 0$.*

En dimension $n \geqslant 1$

Théorème 31. *Soit $J : C \to \mathbb{R}$ différentiable. Il y a équivalence entre :*
- *(i) J est convexe sur C.*
- *(ii) $\forall x, y \in C$, $\langle \nabla J(x) - \nabla J(y), x - y \rangle \geqslant 0$.*
- *(iii) $\forall x, y \in C$, $J(x) \geqslant J(y) + \langle \nabla J(y), x - y \rangle$.*

Si J est deux fois différentiable, on a aussi : $\left\langle d^2 J(x) \cdot y, y \right\rangle \geqslant 0$.

Théorème 32. *Soit $J : C \to \mathbb{R}$ différentiable. Il y a équivalence entre :*
- *(i) J est α-convexe sur C.*
- *(ii) $\forall x, y \in C$, $\langle \nabla J(x) - \nabla J(y), x - y \rangle \geqslant \alpha \|x-y\|^2$.*
- *(iii) $\forall x, y \in C$, $J(x) \geqslant J(y) + \langle \nabla J(y), x - y \rangle + \frac{\alpha}{2}\|x-y\|^2$.*

Si J est deux fois différentiable, on a aussi : $\left\langle d^2 J(x) \cdot y, y \right\rangle \geqslant \alpha \|y\|^2$.

Exemple 33. *Si A est une matrice symétrique définie positive, alors la fonctionnelle quadratique $J : X \mapsto \langle AX, X \rangle - \langle B, X \rangle$ est λ_1-convexe, où λ_1 est la plus petite valeur propre de A.*

III Applications

1) Inégalité de convexité

Proposition 34. *Soient $n \in \mathbb{N}^\star$ et $a_1, \dots, a_n \in \mathbb{R}^+$, alors :*

$$\sqrt[n]{\prod_{i=1}^{n} a_i} \leqslant \frac{1}{n} \sum_{i=1}^{n} a_i$$

Il y a égalité si, et seulement si, tous les a_i sont égaux.

Proposition 35 (Young)**.** *Soient $p, q > 0$ tels que $\frac{1}{p} + \frac{1}{q} = 1$ et $a, b \in \mathbb{R}^+$:*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

Il y a égalité si, et seulement si, $a^p = b^q$.

Corollaire 36 (Hölder et Minkowski)**.** *Soient $p, q > 0$ tels que $\frac{1}{p} + \frac{1}{q} = 1$ et $a, b \in \mathbb{R}^+$ et $f, g : E \to \mathbb{K}$ mesurables, alors :*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q \quad \text{et} \quad \|f+g\|_p \leqslant \|f\|_p + \|g\|_p$$

Lemme 37. *Soient $A, B \in M_n(\mathbb{R})$ symétriques définies positives, et $\alpha, \beta \geqslant 0$ tels que $\alpha + \beta = 1$, alors :*

$$\det(\alpha A + \beta B) \geqslant \det(A)^\alpha \det(B)^\beta$$

Application 38 (Ellipsoïde de John-Loewner)**.** *Soit K un compact d'intérieur non vide de $\mathbb{R}^n$, alors il existe un unique ellipsoïde de centre 0 et de volume minimal contenant K.*

2) Optimisation

Théorème 39. *Si $J : C \to \mathbb{R}$ est différentiable en $u \in C$ et admet un minimum local en u, alors :*

$$\forall v \in C,\ \langle \nabla J(u), v - u \rangle \geqslant 0$$

Théorème 40. *On considère $J : C \to \mathbb{R}$.*

(i) Si J est convexe, tout minimum local est global.

(ii) Si J est strictement convexe, J admet au plus un minimum global.

(iii) Si J est α-convexe, J admet un unique minimum global.

(iv) Si J est définie sur un ouvert contenant C et différentiable en $u \in C$, alors le théorème précédent donne en fait une équivalence.

(v) Si C est ouvert, le théorème précédent équivaut à $\nabla J(u) = 0$.

3) Méthodes de gradient

Soit $J : \mathbb{R}^n \to \mathbb{R}$. On suppose J différentiable. On cherche, s'il existe, un élément $u \in \mathbb{R}^n$ tel que :

$$J(u) = \inf_{v \in \mathbb{R}^n} J(v)$$

Pour cela, on utilise les méthodes de gradient. On considère la suite :

$$u_0 \in \mathbb{R}^n \quad \text{et} \quad \forall k \in \mathbb{N},\ u^{k+1} = u^k - \rho^k \nabla J(u^k)$$

Il existe plusieurs possibilités pour choisir les ρ^k, par exemple :

— Gradient à pas fixe : $\rho^k = \rho$ une constante positive fixée.

— Gradient à pas optimal : ρ^k minimise $\rho \mapsto J\left(u^k - \rho \nabla J(u^k)\right)$.

Théorème 41. *Si J est α-convexe et différentiable, et que ∇J est L-lipschitzienne, alors la méthode de gradient à pas optimal converge vers l'unique minimum de J.*

Développements

— Ellipsoïde de John-Loewner (37,38) [FGN13c]

— Algorithme de gradient à pas optimal (41) [Cia88]

Références

[RDO91] Edmond Ramis, Claude Deschamps, and Jacques Odoux. *Cours de Mathématiques, Topologie et éléments d'analyse.* Masson, 1991

[Rom19] Jean-Étienne Rombaldi. *Éléments d'analyse réelle.* EDP Sciences, 2019

[Hau07] Bertrand Hauchecorne. *Les Contre-exemples en Mathématiques.* Ellipses, 2007

[FGN13c] Serge Francinou, Hervé Gianella, and Serge Nicolas. *Oraux X-ENS Algèbre 3.* Cassini, 2013

[Cia88] Philippe Ciarlet. *Introduction à l'analyse numérique et à l'optimisation.* Masson, 1988

Annexes

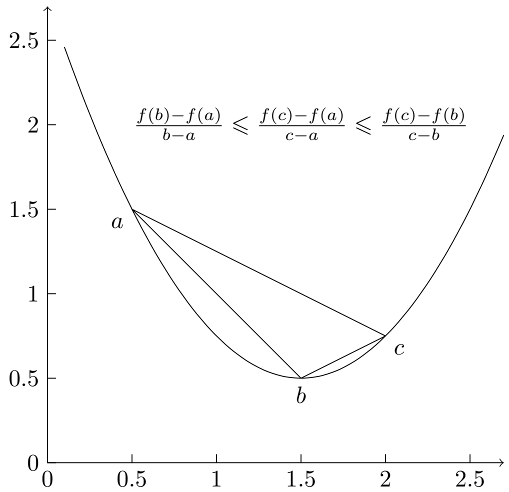

Figure 1 – Inégalité des trois pentes

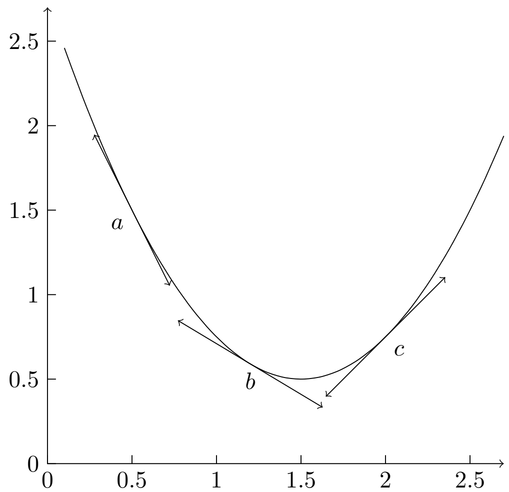

Figure 2 – Tangentes d'une fonction convexe