Chapitre 11 : Fonctions convexes

Rappelons qu'une partie $\mathcal{C}$ d'un espace affine est *convexe* si et seulement si pour tous points $A\,,\,B \in \mathcal{C}$ le segment $[AB]$ est inclus dans $\mathcal{C}$, c'est à dire que pour tout réel $t \in [0\,,\,1]$ on a : $t\,A + (1-t)\,B \in \mathcal{C}$, et que les parties convexes de $\mathbb{R}$ sont les intervalles.

1 Définitions et premières propriétés

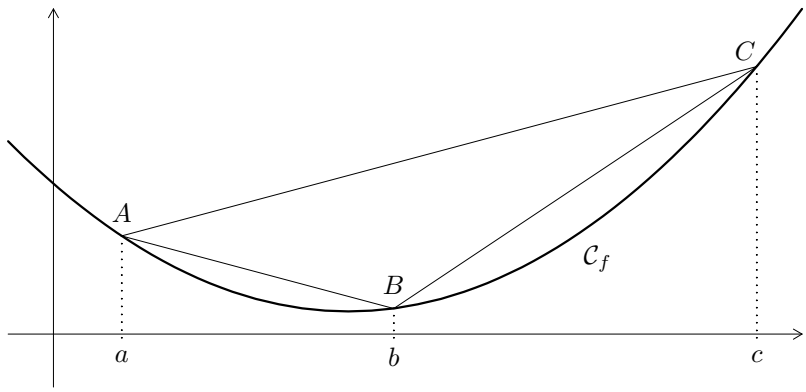
Définition 1 *Soient I un intervalle de $\mathbb{R}$ et $f\,:\,I \to \mathbb{R}$ une application. On dit que f est convexe si et seulement si pour tous points $a\,,\,b \in I$ et pour tout réel $t \in [0\,,\,1]$ on a :*

$$f\big(\,t\,a + (1-t)\,b\,\big) \leqslant t\,f(a) + (1-t)\,f(b) \quad .$$

Interprétation géométrique : on peut échanger a et b dans cette inégalité (en posant $t' = 1-t$) et elle est toujours vérifiée si $a = b$: il suffit donc de la vérifier si $a < b \in I$, et elle signifie que sur l'intervalle $[a\,,\,b]$ le graphe $\mathcal{C}_f$ de f est situé **en dessous de sa corde** entre a et b, c'est à dire du segment reliant $\big(a\,,\,f(a)\big)$ à $\big(b\,,\,f(b)\big)$, ce qui s'écrit aussi :

$$(*) \qquad f(x) \leqslant f(a) + \frac{x-a}{b-a}\,\big(f(b) - f(a)\big) \text{ pour tout } x \in\, [a\,,\,b] \quad .$$

D'autre part, en notant $\mathcal{E}_f = \{(x\,,\,y) \in \mathbb{R}^2 \,/\, x \in I \text{ et } y \geqslant f(x)\,\}$ l'*épigraphe* de f, la fonction f est convexe si et seulement si $\mathcal{E}_f$ est une partie convexe du plan.



Proposition 1 Inégalité des pentes. *Si $f\,:\,I \to \mathbb{R}$ est convexe, on a :*

$$\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b} \qquad \textit{pour tous } a < b < c \in I \quad .$$

Démonstration : en posant $A = \big(a\,,\,f(a)\big)$, $B = \big(c\,,\,f(b)\big)$ et $C = \big(c\,,\,f(c)\big)$, le point B est situé en dessous de la corde $[AC]$, et les inégalités ci-dessus signifient que la pente de la droite (AC) est comprise entre celle de (AB) et celle de (BC), ce qui est clair géométriquement. D'un point de vue analytique, la première inégalité résulte de $(*)$ (appliqué à $x = b \in [a\,,\,c]$), et la seconde s'en déduit puisque :

$$f(c) - \frac{c-b}{c-a}\,\big(f(c)-f(a)\big) \;=\; f(a) + \frac{b-a}{c-a}\,\big(f(c)-f(a)\big) \;\geqslant f(b) \quad .$$

On aurait aussi pu poser : $b = t\,a + (1-t)\,c$ pour se ramener à la définition 1...

Proposition 2 *Si $f : I \to \mathbb{R}$ est convexe, son graphe est situé* **au dessus de ses tangentes**, *c'est à dire que si f est dérivable en un point $a \in I$, on a :*

$$f(x) \geqslant f(a) + (x-a)\, f'(a) \qquad \textit{pour tout } x \in I \quad .$$

Démonstration : si $x = a$ l'inégalité est vraie. Si $x > a$, pour tout $y \in]\, a\, ,\, x\, [$ on obtient :

$$\frac{f(y)-f(a)}{y-a} \leqslant \frac{f(x)-f(a)}{x-a}$$

en appliquant l'inégalité des pentes à $a < y < x$, et en faisant tendre y vers a on en déduit :

$$f'(a) \leqslant \frac{f(x)-f(a)}{x-a} \qquad \text{donc :} \qquad f(x) \geqslant f(a) + (x-a)\, f'(a) \quad .$$

Enfin, si $x < a$, pour tout réel y tel que $x < y < a$ l'inégalité des pentes montre que :

$$\frac{f(a)-f(x)}{a-x} \leqslant \frac{f(a)-f(y)}{a-y} = \frac{f(y)-f(a)}{y-a}$$

et en faisant tendre y vers a on obtient : $\dfrac{f(a)-f(x)}{a-x} \leqslant f'(a)$, d'où le résultat.

Proposition 3 *Soient I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction dérivable. La fonction f est convexe si et seulement si la fonction f' est croissante.*

Démonstration : supposons f convexe, et soient $a < b \in I$. Pour tout réel $a < x < b$, l'inégalité des pentes montre que :

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} = p(a\, ,\, b) \leqslant \frac{f(b)-f(x)}{b-x} = \frac{f(x)-f(b)}{x-b}$$

donc en faisant tendre x vers a on en déduit que : $f'(a) \leqslant p(a\, ,\, b)$, et en faisant tendre x vers b on en déduit que : $p(a\, ,\, b) \leqslant f'(b)$, d'où $f'(a) \leqslant f'(b)$ et la fonction f' est donc croissante. Réciproquement, si f' est croissante, pour tous $a < b \in I$ et tout $x \in]\, a\, ,\, b\, [$ il existe d'après le théorème des accroissements fini des réels $\alpha \in]\, a\, ,\, x\, [$ et $\beta \in]\, x\, ,\, b\, [$ tels que :

$$\frac{f(x)-f(a)}{x-a} = f'(\alpha) \qquad \text{et} \qquad \frac{f(b)-f(x)}{b-x} = f'(\beta)$$

et on a : $\alpha < \beta$ donc $f'(\alpha) \leqslant f'(\beta)$, et on obtient : $\dfrac{f(x)-f(a)}{x-a} \leqslant \dfrac{f(b)-f(x)}{b-x}$, d'où :

$$f(x) \leqslant \frac{b-x}{b-a}\, f(a) + \frac{x-a}{b-a}\, f(b) = f(a) + \frac{x-a}{b-a}\big(f(b)-f(a)\big)$$

et cette inégalité reste vraie si $x = a$ ou $x = b$. D'après $(*)$, on en conclut que f est convexe.

Corollaire 1 *Soient I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction deux fois dérivable. La fonction f est convexe si et seulement si la fonction f'' est à valeurs positives ou nulles.*

Définition 2 *Une application $f : I \to \mathbb{R}$ est dite concave si la fonction $-f$ est convexe.*

Le graphe d'une fonction concave est donc situé au dessus de ses cordes et en dessous de ses tangentes, et une fonction deux fois dérivable f est concave si et seulement si la fonction f'' est à valeurs négatives ou nulles.

Exercice 1 Montrer les inégalités suivantes en utilisant les résultats ci-dessus.

a) $e^x \geqslant 1 + x$ pour tout réel x.

b) $\ln(1+x) \leqslant x$ pour tout réel $x > -1$.

c) $\dfrac{2}{\pi}\, x \leqslant \sin(x) \leqslant x$ pour tout $x \in \left[0\, ,\, \dfrac{\pi}{2}\right]$.

d) $\sqrt{1+x} \geqslant 1 + \dfrac{x}{2}$ pour tout réel $x \geqslant -1$.

2 Inégalités de convexité

Théorème 1 Inégalité de Jensen. *Si $f : I \to \mathbb{R}$ est convexe, si $x_1, \dots, x_n \in I$, et si les réels $\alpha_1, \dots, \alpha_n$ vérifient : $\alpha_k \geqslant 0$ pour tout entier $1 \leqslant k \leqslant n$ et $\displaystyle\sum_{k=1}^{n} \alpha_k = 1$, on a :*

$$f\Big(\sum_{k=1}^{n} \alpha_k\, x_k\Big) \leqslant \sum_{k=1}^{n} \alpha_k\, f(x_k) \quad .$$

Démonstration : l'épigraphe $\mathcal{E}_f$ étant convexe, si $A, B \in \mathcal{E}_f$ tout barycentre à coefficients positifs ou nuls de A et B appartient à $\mathcal{E}_f$ puisqu'il s'écrit $t\,A + (1-t)\,B$ avec $t \in [0, 1]$. Par associativité du barycentre, on en déduit que si $A_1, \dots, A_n \in \mathcal{E}_f$, tout barycentre à coefficients positifs de $A_1, \dots, A_n$ appartient à $\mathcal{E}_f$, d'où le résultat.
Remarque : comme dans la première composition du CAPES 2013, on peut aussi prouver ce résultat par récurrence : le cas $n = 1$ et l'hérédité si $\alpha_{n+1} = 0$ sont évidents, et si $\alpha_{n+1} \neq 0$ l'hérédité découle de la définition 1 et de l'associativité du barycentre dans $\mathbb{R}$ qui s'écrit :

$$\sum_{k=1}^{n+1} \alpha_k\, x_k \;=\; \alpha_{n+1}\; x_{n+1} + (1-\alpha_{n+1}) \sum_{k=1}^{n} \frac{\alpha_k}{1-\alpha_{n+1}}\, x_k \quad .$$

Si f est concave, on obtient bien sûr la même inégalité *inversée*, qu'on emploie la plupart du temps sous la forme :

$$f\Big(\frac{1}{n}\sum_{k=1}^{n} x_k\Big) \geqslant \frac{1}{n}\sum_{k=1}^{n} f(x_k) \quad .$$

Exercice 2 En utilisant la concavité du logarithme, montrer l'inégalité :

$$\Big(\prod_{k=1}^{n} x_k\Big)^{\frac{1}{n}} \leqslant \frac{1}{n}\sum_{k=1}^{n} x_k$$

pour tous réels $x_1, \dots, x_n \geqslant 0$, appelée *l'inégalité arithmético-géométrique.*

Exercice 3 Pour tous réels $x_1, \dots, x_n > 0$ et $p \neq 0$, on pose :

$$M_p(x_1, \dots, x_n) = \Big(\frac{1}{n}\sum_{k=1}^{n} (x_k)^p\Big)^{\frac{1}{p}} \quad .$$

a) Montrer que si $x_1, \dots, x_n > 0$ et $r \geqslant 1$, on a : $M_r(x_1, \dots, x_n) \geqslant M_1(x_1, \dots, x_n)$.
b) En déduire que si $x_1, \dots, x_n > 0$ et $p \geqslant q > 1$, on a :

$$M_p(x_1, \dots, x_n) \geqslant M_q(x_1, \dots, x_n) \quad .$$

On pourra poser $r = \dfrac{p}{q}$ et $y_k = (x_k)^{\frac{1}{q}}$ si $1 \leqslant k \leqslant n$.
c) Montrer que pour tous réels $x_1, \dots, x_n > 0$ et $p \neq 0$, on a :

$$M_{-p}(x_1, \dots, x_n) = \Big(M_p\Big(\frac{1}{x_1}, \dots, \frac{1}{x_n}\Big)\Big)^{-1} \quad .$$

On fixe désormais des réels $x_1, \dots, x_n > 0$ et pour tout réel $p \neq 0$ on pose :

$$f(p) = M_p(x_1, \dots, x_n) \quad .$$

d) Déterminer les limites de f en $+\infty$ et en $-\infty$.
e) Déterminer la limite de f en 0 et en déduire qu'elle se prolonge en une fonction $\widetilde{f}$ continue et croissante sur $\mathbb{R}$. On pose donc : $M_0(x_1, \dots, x_n) = \widetilde{f}(0)$: calculer $M_0(x_1, \dots, x_n)$.
f) En déduire des inégalités entre M_{-1} (*moyenne harmonique*), M_0 (*moyenne géométrique*), M_1 (*moyenne arithmétique*) , M_2 (*moyenne quadratique*) et le plus petit et le plus grand des réels x_k pour $1 \leqslant k \leqslant n$.

3 Régularité des fonctions convexes

Si A est une partie de $\mathbb{R}$, un réel a est un *point intérieur* à A s'il existe un réel $r > 0$ tel que : $]\,a - r\,,\,a + r\,[\,\subset A$, et on appelle *intérieur* de A (parfois noté $Int(A)$ l'ensemble des points intérieurs à A. Si A est un intervalle quelconque d'extrémités a et b, l'intérieur de A est tout simplement l'intervalle *ouvert* $]\,a\,,\,b\,[$.

Théorème 2 *Si $f\,:\,I \to \mathbb{R}$ est convexe, elle est dérivable à droite et à gauche en tout point intérieur à I, et pour tous réels $a < b$ intérieurs à I on a :*

$$f'_g(a) \leqslant f'_d(a) \leqslant \frac{f(b) - f(a)}{b - a} \leqslant f'_g(b) \leqslant f'_d(b) \quad .$$

En particulier, les fonctions f'_g et f'_d sont croissantes.

Corollaire 2 *Si $f\,:\,I \to \mathbb{R}$ est convexe, elle est continue sur l'intérieur de I.*

Aux extrémités de I, elle n'est pas forcément continue (donc *a fortiori* pas dérivable à droite ou à gauche) comme le montre l'exemple de la fonction $f\,:\,[0\,,\,1] \to \mathbb{R}$ définie par : $f(x) = 0$ si $0 < x < 1$ et $f(0) = f(1) = 1$, qui est convexe.

Démonstration : si $a \in Int(I)$, il existe des points $b\,,\,c \in Int(I)$ tels que $c < a < b$, et nous les fixons. Pour tous réels $x\,,\,y$ vérifiant $c < x < a < y < b$, l'inégalité des pentes implique :

$$p_1 = \frac{f(a) - f(c)}{a - c} \leqslant \frac{f(a) - f(x)}{a - x} \leqslant \frac{f(y) - f(a)}{y - a} \leqslant \frac{f(b) - f(a)}{b - a} = p_2 \qquad (1)$$

et en posant $C = \max(\,|\,p_1\,|\,,\,|\,p_2\,|\,)$ on obtient : $|\,f(z) - f(a)\,| \leqslant C\ \,|\,z - a\,|$ pour tout $z \in]\,c\,,\,b\,[$ ce qui prouve le corollaire (qu'on peut aussi déduire aisément du théorème). De plus, pour tous réels $x\,,\,x'$ vérifiant $c < x < x' < a < b$, on obtient de même :

$$p_-(x) = \frac{f(a) - f(x)}{a - x} \leqslant \frac{f(a) - f(x')}{a - x'} \leqslant \frac{f(b) - f(a)}{b - a}$$

donc la fonction $p_-\,:\,]\,c\,,\,a\,[\to \mathbb{R}$ est croissante et majorée, donc elle admet une limite $f'_g(a)$ au point a. De même, pour tous réels $y\,,\,y'$ vérifiant $c < a < y < y' < b$, on a :

$$\frac{f(a) - f(c)}{a - c} \leqslant \frac{f(y) - f(a)}{y - a} \leqslant \frac{f(y') - f(a)}{y' - a} = p_+(y')$$

donc la fonction $p_+\,:\,]\,a\,,\,b\,[\to \mathbb{R}$ est croissante et minorée, donc elle admet une limite $f'_d(a)$ au point a. Enfin, en faisant tendre x et y vers a dans les inégalités (1) on obtient :

$$\frac{f(a) - f(c)}{a - c} \leqslant f'_g(a) \leqslant f'_d(a) \leqslant \frac{f(b) - f(a)}{b - a}$$

ce qui prouve le théorème en échangeant les rôles des points a, b et c.

La fonction "valeur absolue" est convexe, de même que la fonction affine par morceaux $h\,:\,\mathbb{R} \to \mathbb{R}$ définie par : $h(p) = p^2$ si $p \in \mathbb{Z}$, donc si f est convexe il peut y avoir une infinité de points où elle n'est pas dérivable. Par contre, on a le théorème suivant.

Théorème 3 *Si $f\,:\,I \to \mathbb{R}$ est convexe, elle est dérivable en dehors d'un ensemble au plus dénombrable.*

Démonstration : en posant $\mathcal{D} = \{a \in Int(I)\,/\,f'_g(a) < f'_d(a)\}$, la fonction f est dérivable en tout point de $Int(I) \setminus \mathcal{D}$, et comme $I \setminus Int(I)$ contient au plus deux points il suffit de montrer que $\mathcal{D}$ est au plus dénombrable pour conclure. Mais pour tout point $a \in \mathcal{D}$ il existe un nombre rationnel $\varphi(a)$ tel que : $f'_g(a) < \varphi(a) < f'_d(a)$ puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$, et si $a < b \in \mathcal{D}$ le théorème 2 montre qu'on a : $f'_g(a) < \varphi(a) < f'_d(a) \leqslant f'_g(b) < \varphi(b) < f'_d(b)$ donc $\varphi(a) \neq \varphi(b)$. L'application $\varphi\,:\,\mathcal{D} \to \mathbb{Q}$ est donc injective, et comme $\mathbb{Q}$ est dénombrable l'ensemble $\mathcal{D}$ est au plus dénombrable.