Convexité

Geoffrey Deperle

Table des matières

Introduction

Nous présenterons dans ce document quelques résultats liés à la notion de convexité des fonctions de la variable réelle et nous généraliserons dans les cas des fonctions défini sur un espace vectoriel de dimension fini.

1 Définition : Barycentre, convexité

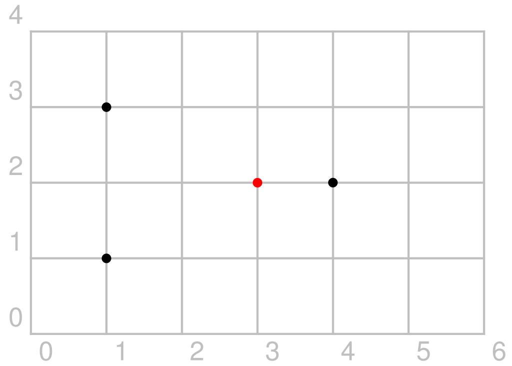
1.1 Barycentre

1.1.1 Définition

Définition 1. Soit $x_1, \ldots, x_n \in \mathbb{R}^d$ et $\lambda_1, \ldots, \lambda_n$ n réels tels que $\lambda_1 + \cdots + \lambda_n = 1$. On appelle **barycentre** de $x_1, \ldots, x_n$ associé aux poids $\lambda_1, \ldots, \lambda_n$, noté $\mathrm{Bar}((x_1, \lambda_1), \ldots, (x_n, \lambda_n))$ le point $\lambda_1 x_1 + \cdots + \lambda_n x_n$.

Remarque 2. Le barycentre de $x_1, \ldots, x_n$ associé à $\lambda_1, \ldots, \lambda_n$ s'interprète comme la moyenne des points $x_1, \ldots, x_n$ pondérée par les poids $\lambda_1, \ldots, \lambda_n$.

Lorsque les réels $\lambda_1, \ldots, \lambda_n$ sont égaux, on parle d'**isobarycentre**.



Exemple 3. Le point $D(3, 2)$ est le barycentre de $A(1, 1)$, $B(1, 3)$, $C(4, 2)$ associé aux poids respectifs $\frac{1}{6}, \frac{1}{6}, \frac{2}{3}$.

Dans toute la suite, nous considérerons plus forcément des réels de somme 1 mais des réels de somme non nul. Le barycentre du système de points est défini comme le barycentre avec les mêmes réels divisé par la somme des poids pour avoir une somme égal à 1.
Le barycentre a une caractérisation vectorielle :

Proposition 4. *Soit $A_1, \ldots, A_n$ n points et $\lambda_1, \ldots, \lambda_n$ n réels de somme non nuls. Le barycentre du système $(A_1, \lambda_1), \ldots, (A_n, \lambda_n)$ est l'unique point G vérifiant $\lambda_1 \overrightarrow{GA_1} + \cdots + \lambda_n \overrightarrow{GA_n} = 0$.*

Preuve : Quitte à diviser tous les réels par $\lambda_1 + \cdots + \lambda_n$ non nul par hypothèse, on peut supposer que $\lambda_1 + \cdots + \lambda_n = 1$.

On a pour $G \in \mathbb{R}^d$,

$$\begin{aligned}
&\lambda_1 \overrightarrow{GA_1} + \cdots + \lambda_n \overrightarrow{GA_n} = 0 \\
&\iff \lambda_1(\overrightarrow{GO} + \overrightarrow{OA_1}) + \ldots \lambda_n(\overrightarrow{GO} + \overrightarrow{OA_n}) = 0 \qquad \text{par relation de Chasles} \\
&\iff \lambda_1 \overrightarrow{OA_1} + \cdots + \lambda_n \overrightarrow{OA_n} = \lambda_1 \overrightarrow{OG} + \cdots + \lambda_n \overrightarrow{OG} \\
&\iff \lambda_1 \overrightarrow{OA_1} + \cdots + \lambda_n \overrightarrow{OA_n} = \overrightarrow{OG} \qquad \text{car } \lambda_1 + \cdots + \lambda_n = 1 \\
&\iff G = \mathrm{Bar}((A_1, \lambda_1), \ldots, (A_n, \lambda_n))
\end{aligned}$$

$\square$

Ainsi, le barycentre est à pondération près, le point à équidistance de tous les points du système de points.
La propriété précédente peut se réecrire autrement en utilisant la notion de fonction de Leibnitz.

Définition 5. Soit $A_1, \ldots, A_n$ n points et $\lambda_1, \ldots, \lambda_n$ n réels de somme non nuls. On appelle **fonction de Leibnitz** associé au système $\{(A_i, \lambda_i)_{i \in [\![1,n]\!]}\}$ la fonction $f : M \mapsto \lambda_1 \overrightarrow{MA_1} + \ldots \lambda_n \overrightarrow{MA_n}$.

La proposition précédente donne donc l'équivalence suivante : un point M est barycentre du système $\{(A_i, \lambda_i)_{i \in [\![1,n]\!]}\}$ (en re-normalisant $\lambda_1, \ldots, \lambda_n$) si et seulement si $f(M) = 0$. Plus généralement, on a

$$\begin{aligned}
f(M) &= f(M) - f(G) \qquad \text{avec } G \text{ le barycentre de } \{(A_i, \lambda_i)_{i \in [\![1,n]\!]}\} \\
&= \lambda_1 \overrightarrow{MA_1} + \ldots \lambda_n \overrightarrow{GA_n} - \lambda_1 \overrightarrow{GA_1} - \ldots \lambda_n \overrightarrow{GA_n} \\
&= \lambda_1 \overrightarrow{MG} + \lambda_n \overrightarrow{MG} \\
&= (\lambda_1 + \cdots + \lambda_n) \overrightarrow{MG}
\end{aligned}$$

Le théorème suivant, bien que simple a de nombreuses conséquences :

Théorème 6 (Associativité du barycentre)**.** *Soit x_1, x_2, x_3 trois points de R^d, $\lambda_1, \lambda_2, \lambda_3$ trois réels. On a*

$$\mathrm{Bar}\left((\mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2)), \lambda_1 + \lambda_2), (x_3, \lambda_3)\right) = \mathrm{Bar}\left((x_1, \lambda_1), (\mathrm{Bar}((x_2, \lambda_1), (x_3, \lambda_2)), \lambda_2 + \lambda_3)\right)$$

et ces deux expressions sont égales à

$$\mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2), (x_3, \lambda_3))$$

Preuve : Il s'agit de vérifier que les deux expressions coïncident :

$$\begin{aligned}
\mathrm{Bar}\left((\mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2)), \lambda_1 + \lambda_2), (x_3, \lambda_3)\right) &= \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2)) + \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} x_3 \\
&= \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2 \right) + \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} x_3 \\
&= \frac{\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3}{\lambda_1 + \lambda_2 + \lambda_3} \\
&= \mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2), (x_3, \lambda_3))
\end{aligned}$$

Par symétrie des deux expressions (en permutant le rôle de x_1, x_2t et x_3), on en déduit par le même raisonnement l'égalité $\mathrm{Bar}\left((x_1, \lambda_1), (\mathrm{Bar}((x_2, \lambda_1), (x_3, \lambda_2)), \lambda_2 + \lambda_3)\right) = \mathrm{Bar}((x_1, \lambda_1), (x_2, \lambda_2), (x_3, \lambda_3))$
□

1.1.2 Aspect géométrique

Avec la notion de barycentre, on peut réecrire les définitions d'objets géométriques usuels.

Définition 7. Soit x et y deux points de $\mathbb{R}^d$, on appelle **segment** liant x à y noté $[x, y]$ l'ensemble des barycentres de x et y.

L'ensemble des barycentres de x et y est l'ensemble des

$$\{tx + sy \,|\, (t, s) \in [0, 1]^2 \text{ et } s + t = 1\} = \{tx + (1 - t)y \,|\, t \in [0, 1]\}$$

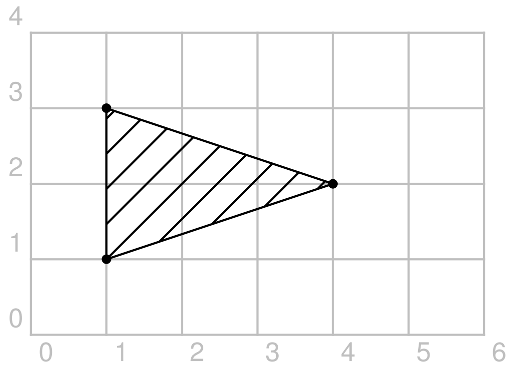
On a donc
$$[x, y] = \{tx + (1 - t)y \mid t \in [0, 1]\}$$

Définition 8. Soit x, y et z trois points de $\mathbb{R}^d$, on appelle **triangle plein** liant x, y et z l'ensemble des barycentres de x, y et z.

Pour n points quelconques, on parle d'enveloppe convexe :

Définition 9. Soit $x_1, \ldots, x_n$ n points de $\mathbb{R}^d$, on appelle **enveloppe convexe** de $x_1, \ldots, x_n$ noté $\mathrm{Conv}(x_1, \ldots, x_n)$ l'ensemble des barycentres de $x_1, \ldots, x_n$.

Définition 10. Soit ABC un triangle de $\mathbb{R}^d$, on appelle médiane de A et B la droite passant par C et par le milieu de $[AB]$.



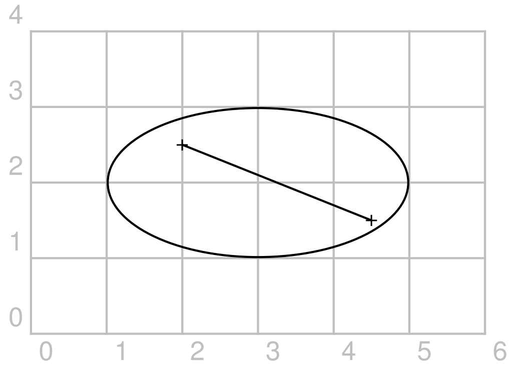
Exemple 11. Le triangle plein reliant $A(1, 1)$, $B(1, 3)$, $C(4, 2)$.

Proposition 12. *Les médianes d'un triangle sont concourantes et se coupent en l'isobarycentre de x, y et z.*

Preuve : Notons M_x (resp. M_y) les médianes partant de x joignant le milieu de $[yz]$ (resp. partant de y joignant le milieu de $[xz]$). Notons t le point d'intersection de M_x et M_y. $t \in M_x$ donc est barycentre de $\mathrm{Bar}((y, 1), (z, 1))$ et (x, λ) avec λ à déterminer. De plus, $t \in M_y$ donc est barycentre de $\mathrm{Bar}((x, 1), (z, 1))$ et (y, μ) avec μ à déterminer. D'où t vérifie
$$\lambda x + y + z = x + \mu y + z.$$
Comme $d \geq 2$, c'est un système à deux inconnues de plus de deux équations. Donc il y a au plus une solution. Comme $\lambda = \mu = 1$ est solution, il s'agit de l'unique solution. Donc t est l'isobarycentre de x, y et z.
Par symétrie, l'intersection de M_x et M_z est également l'isobarycentre de x, y et z. $\square$

1.2 Partie convexe

Définition 13. Soit $A \subset \mathbb{R}^d$ une partie de $\mathbb{R}^d$. On dit que A est **convexe**, si tout segment rejoignant deux points de A est contenu dans A. C'est-à-dire :
$$\forall x, y \in A, \forall t \in [0, 1], tx + (1 - t)y \in A$$



Remarque 14. Visuellement, une partie convexe est "enveloppé" et ne possède pas de creux.

Exemple 15. Soit $z \in \mathbb{R}^d$, $r > 0$. La boule ouverte de centre x de rayon $r > 0$ $B(z, r)$ est convexe. En effet, soit $x, y \in B(z, r)$, $t \in [0, 1]$:

$$\begin{aligned} ||tx + (1-t)y - z|| &= ||tx + (1-t)y - tz - (1-t)z|| \\ &= ||t(x-z) + (1-t)(y-z)|| \\ &< |t| \underbrace{||x-z||}_{<r} + |1-t| \underbrace{||y-z||}_{<r} \qquad \text{par inégalité triangulaire} \\ &< tr + (1-t)r \\ &< r \end{aligned}$$

Donc $tx + (1-t)y \in B(z, r)$.

Proposition 16. *Soit $x_1, \ldots, x_n$ n points de $\mathbb{R}^d$. $\mathrm{Conv}(x_1, \ldots, x_n)$ est la plus petite partie convexe contenant $x_1, \ldots, x_n$.*

Preuve : Il s'agit de montrer que $\mathrm{Conv}(x_1, \ldots, x_n)$ est une partie convexe de $\mathbb{R}^d$ et de montrer que toute partie convexe contenant $x_1, \ldots, x_n$ contient nécessairement $\mathrm{Conv}(x_1, \ldots, x_n)$.

Soit $x, y \in \mathrm{Conv}(x_1, \ldots, x_n), t \in [0, 1]$. Montrons que $tx + (1-t)y \in \mathrm{Conv}(x_1, \ldots, x_n)$.

Comme $x \in \mathrm{Conv}(x_1, \ldots, x_n)$, il existe $\lambda_1, \ldots, \lambda_n$ des réels positifs de somme 1 tel que $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$. De même, il existe $\mu_1, \ldots, \mu_n$ des réels positifs de somme 1 tel que $y = \mu_1 x_1 + \cdots + \mu_n x_n$. D'où

$$tx + (1-t)y = (t\lambda_1 + (1-t)\mu_1)x_1 + \cdots + (t\lambda_n + (1-t)\mu_n)x_n$$

Or, les $(t\lambda_i + (1-t)\mu_i)$ sont des réels positifs de somme égal à 1 donc $tx + (1-t)y \in \mathrm{Conv}(x_1, \ldots, x_n)$.
Soit A une partie convexe contenant $x_1, \ldots, x_n$. Montrons que $\mathrm{Conv}(x_1, \ldots, x_n) \subset A$.
Par récurrence, pour $n = 2$ le résultat est vrai car l'enveloppe convexe de deux points est le segment joignant ces deux points.
Supposons que pour un $n \in \mathbb{N}^*$ fixé, la combinaison convexe de n points de A est inclus dans A. Soit $x_1, \ldots, x_{n+1}$ $n+1$ points de A.
Soit $A \in \mathrm{Conv}(x_1, \ldots, x_{n+1})$, alors il existe $\lambda_1, \ldots, \lambda_{n+1}$ tel que $A = \mathrm{Bar}((x_1, \lambda_1), \ldots, (x_{n+1}, \lambda_{n+1}))$.
Or, par associativité du barycentre $\mathrm{Bar}((x_1, \lambda_1), \ldots, (x_{n+1}, \lambda_{n+1})) = \mathrm{Bar}(\mathrm{Bar}((x_1, \lambda_1), \ldots, (x_n, \lambda_n))(x_{n+1}, \lambda$
Par hypothèse de récurrence $\mathrm{Bar}((x_1, \lambda_1), \ldots, (x_n, \lambda_n)) \in A$. D'où $\mathrm{Bar}(\mathrm{Bar}((x_1, \lambda_1), \ldots, (x_n, \lambda_n))(x_{n+1}, \lambda_{n+1}$
A car A est convexe. $\square$

Remarque 17. On peut montrer qu'une intersection quelconque de parties convexe est convexe. Ainsi, la plus petite partie convexe contenant une partie A est $\bigcap_{\substack{C \subset A \\ C \text{ convexe}}} C$. Ainsi, si A est un ensemble fini $A = \{x_1, \ldots, x_n\}$,, on a $\mathrm{Conv}(x_1, \ldots, x_n) = \bigcap_{\substack{C \subset A \\ C \text{ convexe}}} C$

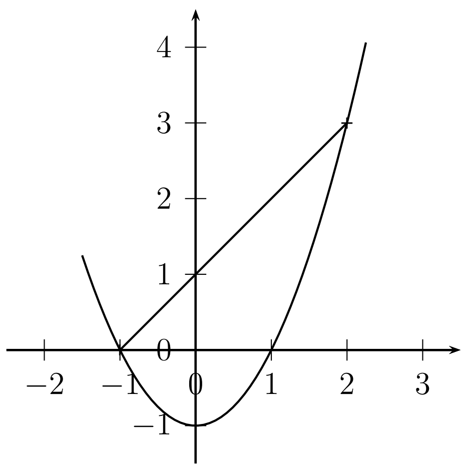
2 Fonction de la variable réelle convexe

Dans toute la suite, on s'intéressera aux fonctions définies sur $\mathbb{R}$ ou une partie de $\mathbb{R}$ seulement. Intuitivement, une fonction convexe est une fonction dont le représentation graphique se "en dessous" de n'importe quel segment reliant deux points de la courbe représentant la fonction. L'image de n'importe quel segment de $\mathbb{R}$ se situe au dessus du segment, ceci motive la définition suivante :

2.1 Définition

Définition 18. Soit $f : \mathbb{R} \to \mathbb{R}$. On dit que f est convexe sur $I \subset \mathbb{R}$ si

$$\forall x, y \in I, t \in [0,1], f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$



Exemple 19. La fonction $x \mapsto x^2 - 1$ est convexe est sur $[-2,3]$

Remarque 20. Par récurrence, on peut aisément montrer qu'une fonction f est convexe si et seulement si pour toute famille finie de réel $x_1, \ldots, x_n$ et toute famille de réel positif $t_1, \dot{,} t_n$ de somme 1,

$$f(t_1x_1 + \cdots + t_nx_n) \leq t_1f(x_1) + \ldots t_nf(x_n).$$

Il y a un lien entre les fonctions convexes et les parties convexes :

Proposition 21. *Soit $f : \mathbb{R} \to \mathbb{R}$. f est convexe si et seulement si son épigraphe $A = \{(x,y) \in \mathbb{R}^2 \,|\, y \leq f(x)\}$ est une partie convexe de $\mathbb{R}^2$.*

Preuve :
($\Longrightarrow$) Supposons que f est convexe. Soit $a = (x_1, y_1), b = (x_2, y_2) \in A, t \in [0,1]$. Montrons que $ta + (1-t)b \in A$ c'est-à-dire $ty_1 + (1-t)y_2 \geq f(tx_1 + (1-t)x_2)$.
Or, par convexité,

$$f(tx_1 + (1-t)x_2) \leq t\underbrace{f(x_1)}_{\leq y_1} + (1-t)\underbrace{f(x_2)}_{\leq y_2} \leq ty_1 + (1-t)y_2$$

($\Longleftarrow$) Supposons A est convexe. Montrons que f est convexe.
Soit $x, y \in \mathbb{R}, t \in [0,1]$. Montrons que $\forall x, y \in R, t \in [0,1], f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Comme $(x, f(x))$ et $(y, f(y)) \in A$. Par convexité, $\forall t \in [0,1], t(x, f(x)) + (1-t)(y, f(y)) \in A$ d'où $tf(x) + (1-t)f(y) \geq f(tx + (1-t)y)$ donc f est convexe. $\square$

Définition 22. On dit qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est **concave** sur $I \subset \mathbb{R}$ si $-f$ est convexe sur I. Ainsi, f est concave si et seulement si

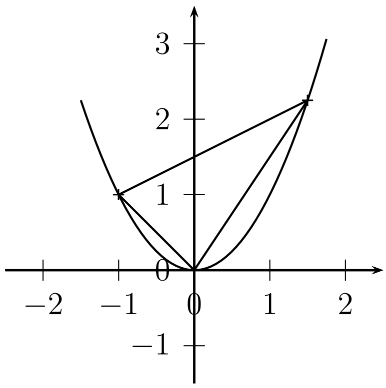
$$\forall x, y \in I, t \in [0,1], f(tx + (1-t)y) \geq tf(x) + (1-t)f(y)$$

2.2 Propriétés des fonctions convexes

Le lemme suivant est important pour démontrer toute sorte de résultat sur les fonctions convexes :

Théorème 23 (Inégalité des trois pentes)**.** $f : \mathbb{R} \to \mathbb{R}$ *est convexe si et seulement si pour tout $x < y < z$ trois points distincts de $\mathbb{R}$,*

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}$$



Preuve :
$(\implies)$ Supposons que f est convexe et $x < y < z \in \mathbb{R}$. On a $y \in [x, z]$ donc y est combinaison convexe de x et z. Cette combinaison convexe est

$$y = \frac{z - y}{z - x}x + \frac{y - x}{z - x}z$$

Par convexité, on a donc

$$f(y) \leq \frac{z - y}{z - x}f(x) + \frac{y - x}{z - x}f(z)$$

D'où

$$\begin{aligned}\frac{f(y) - f(x)}{y - x} &\leq \frac{1}{y - x}\left(\frac{z - y}{z - x}f(x) + \frac{y - x}{z - x}f(z) - f(x)\right)\\ &\leq \frac{1}{y - x}\left(\frac{x - y}{z - x}f(x) + \frac{y - x}{z - x}f(z)\right)\\ &\leq \frac{f(z) - f(x)}{z - x}\end{aligned}$$

et comme

$$\frac{z - y}{z - x}f(x) \geq f(y) + \frac{x - y}{z - x}f(z)$$

$$\begin{aligned}\frac{f(z)-f(x)}{z-x} &\leq \frac{1}{z-x}\left(f(z)+\frac{x-z}{z-y}f(y)+\frac{y-x}{z-y}f(z)\right)\\ &\leq \frac{1}{z-x}\left(\frac{z-x}{z-y}f(z)+\frac{x-z}{z-y}f(y)\right)\\ &\leq \frac{f(z)-f(y)}{z-y}\end{aligned}$$

($\Longleftarrow$) Supposons l'inégalité des trois pentes vérifiées. Montrons que f est convexe. Soit $x, z \in I, t \in [0,1]$. Montrons que $f(tz+(1-t)x) \leq tf(z)+(1-t)f(x)$.
L'inégalité est vérifiée pour $t=0$ et $t=1$. Pour $t \notin \{0,1\}$, on a en posant $y = tz+(1-t)x$, $x < y < z$. Donc par inégalité des pentes :

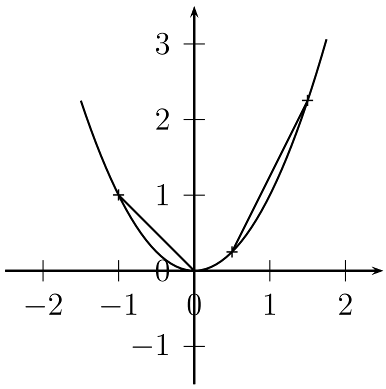
$$f(tz+(1-t)x) = f(y) \leq \frac{z-y}{z-x}f(x)+\frac{y-x}{z-x}f(z) \leq tf(z)+(1-t)f(x)$$

□

Un corollaire de ce théorème permet d'établir une inégalité sur les taux d'accroissements :

Corollaire 24. *Soit f une fonction convexe, pour $x_1 < x_2 < y_1 < y_2$ quatres réels distincts. On a l'inégalité :*

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(y_2)-f(y_1)}{y_2-y_1}$$



Preuve : Appliquons l'inégalité des trois pentes avec $x_1 < x_2 < y_1$:

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(y_1)-f(x_1)}{y_1-x_1} \leq \frac{f(y_1)-f(x_2)}{y_1-x_2}$$

et avec $x_2 < y_1 < y_2$:

$$\frac{f(y_1)-f(x_2)}{y_1-x_2} \leq \frac{f(y_2)-f(x_2)}{y_2-x_2} \leq \frac{f(y_2)-f(y_1)}{y_2-y_1}$$

En rassemblant les inégalités, on trouve le résultat. □

Ces inégalités permettent de donner des propriétés de régularité aux fonctions convexes : Par exemple, pour pour tout segment $]\alpha, \beta[$ strictement inclus dans $I = [a,b]$, on a pour $x \in]\alpha, \beta[$ et $h > 0$ tel que $x + h < \alpha$, on a d'après l'inégalité précédente :

$$f(x+h)-f(x) \leq \underbrace{h\left(\frac{f(b)-f(\beta)}{b-\beta}\right)}_{\xrightarrow[h\to 0]{}0}$$

Pour $h < 0$ tel que $x + h < \alpha$, on a l'inégalité précédente :

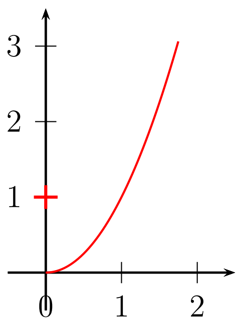
$$f(x+h) - f(x) \geq \underbrace{h\left(\frac{f(\alpha) - f(a)}{\alpha - a}\right)}_{\xrightarrow[h\to 0]{} 0}$$

Donc f est continue sur $]\alpha, \beta[$.

Remarque 25. Une fonction convexe sur I est donc continue sur tout segment ouvert inclus dans I mais n'est pas forcément continue sur tout I comme le montre le contre-exemple suivant :

$$\begin{array}{rccc} f & : & \mathbb{R}^+ & \to & \mathbb{R} \\ & & x & \mapsto & \left\{\begin{array}{ll} 1 & \text{si } x = 0 \\ x^2 & \text{sinon} \end{array}\right. \end{array}$$

f est convexe sur $\mathbb{R}^+$ mais n'est pas continue en 0.



Plus précisément, une fonction est même lipschitzienne sur tout segment ouvert inclus dans son domaine de convexité.

Proposition 26. *Soit $f : I \to \mathbb{R}$ convexe sur $I = [a, b]$. Pour tout $\alpha < \beta$ tel que $]\alpha, \beta[\subset I$. f est lipschitzienne sur $]\alpha, \beta[$*

Preuve : Comme précédemment, on a $f(y) - f(x) \leq (y - x)\left(\frac{f(b)-f(\beta)}{b-\beta}\right)$ si $y > x$ et $f(y) - f(x) \geq (y - x)\left(\frac{f(\alpha)-f(a)}{\alpha-a}\right)$ si $y < x$ d'où

$$|f(y) - f(x)| \leq k|x - y|$$

avec $k = \max(\left(\frac{f(b)-f(\beta)}{b-\beta}\right), \left(\frac{f(\alpha)-f(a)}{\alpha-a}\right))$. $\square$

Remarque 27. Ainsi, une fonction convexe est uniformément continue sur tout segment ouvert inclus dans son domaine de convexité et une suite de fonction convexe qui converge simplement converge uniformément sur tout segment ouvert inclus dans son domaine de convexité.

Un dernier corollaire de l'inégalité des pentes permet de faire le lien entre convexité et la monotonie des taux d'accroissements :

Proposition 28. *Soit $f : \mathbb{R} \to \mathbb{R}$, f est convexe sur $I \subset \mathbb{R}$ si et seulement si pour tout $x_0 \in I$, la fonction taux d'accroissement* $\begin{array}{rccc} \sigma_{x_0} & : & I \setminus \{x_0\} & \to & \mathbb{R} \\ & & x & \mapsto & \frac{f(x)-f(x_0)}{x-x_0} \end{array}$ *est croissante.*

Preuve :
($\Longrightarrow$) Supposons f convexe. Soit $x < y \in I \setminus \{x_0\}$, montrons que $\sigma(x) < \sigma(y)$. Si $x < x_0 < y$, l'inégalité des trois pentes donne :

$$\frac{f(x)-f(x_0)}{x-x_0} < \frac{f(y)-f(x)}{y-x} < \frac{f(y)-f(x_0)}{y-x_0}$$

donc $\sigma_{x_0}(x) < \sigma_{x_0}(y)$. Si $x < y < x_0$, de même l'inégalité des trois pentes donne :

$$\frac{f(y)-f(x)}{y-x} < \frac{f(x)-f(x_0)}{x-x_0} < \frac{f(y)-f(x_0)}{y-x_0}$$

d'où $\sigma_{x_0}(x) < \sigma_{x_0}(y)$ et de même si $x_0 < x < y$.

($\Longleftarrow$) Supposons que pour tout x_0, σ_{x_0} est croissante. Montrons que f est convexe. Il suffit de montrer que f vérifie l'inégalité des trois pentes.
Soit $x < y < z \in I$, La croissance de σ_x donne :

$$\frac{f(y)-f(x)}{y-x} < \frac{f(z)-f(x)}{z-x}$$

et la croissance de σ_z donne :

$$\frac{f(z)-f(x)}{z-x} < \frac{f(z)-f(y)}{z-y}$$

ce qui donne l'inégalité des trois pentes. □

2.3 Propriétés des fonctions convexes dérivables

Soit f une fonction convexe sur I et $x \in I$. La fonction taux d'accroissement de f : σ_x est croissante et majorée au voisinage de x (car continue au voisinage de x), donc admet une limite finie à droite et à gauche en x. Ainsi, toute fonction convexe f est dérivable à droite et à gauche en tout x. Cependant, ces limites peuvent ne pas coïncider comme c'est le cas de la fonction valeur absolue en 0.

Néanmoins, pour les fonctions dérivables, il existe des équivalences de la convexité particulièrement simple :

Proposition 29. *Soit f une fonction dérivable sur $I \subset \mathbb{R}$, f est convexe si et seulement si f' est croissante.*

Preuve :
($\Longrightarrow$) Supposons f convexe. Montrons que f' est croissante. Soit $x < y \in I$. Montrons que $f'(x) < f'(y)$.
Soit $h > 0$ tel que $[x-h, x+h] \cap [y-h, y+h] = \emptyset$. Par inégalité des pentes, on a

$$\frac{f(x+h)-f(x)}{h} \leq \frac{f(y+h)-f(y)}{h}$$

Comme f est dérivable en x et y, les deux termes ont une limite pour $h \to 0$. En passant à la limite, on a $f'(x) \leq f'(y)$.
($\Longleftarrow$) Supposons que f' est croissante. Montrons que f est convexe. Montrons que les taux d'accroissements sont croissants. Soit $x_0 \in I$, soit $x < y \neq x_0$. Montrons que $\sigma_{x_0}(x) < \sigma_{x_0}(y)$.

Supposons $x < x_0 < y$,
f est continue sur I et dérivable sur l'intérieur de I donc d'après le théorème des accroissements finis, il existe $c_x \in]x, x_0[$ tel que $\sigma_{x_0}(x) = f'(c_x)$. De même il existe $c_y \in]x_0, y[$ tel que $\sigma_{x_0}(y) = f'(c_y)$.
Comme $c_x < c_y$ et f' croissant, on a $f'(c_x) \leq f'(c_y)$ d'où $\sigma_{x_0}(x) \leq \sigma_{x_0}(y)$.
Les cas $x_0 < x < y$ et $x < y < x_0$ se traitent de la même façon. □

Ainsi, si f est deux fois dérivable, comme f' est croissante sur I si et seulement si f'' est positive sur I, on a :

Proposition 30. *Soit $f : \mathbb{R} \to \mathbb{R}$ deux fois dérivable sur I un segment de $\mathbb{R}$, f est convexe si et seulement si $f'' \geq 0$ sur I*

Exemple 31.
- $\exp'' > 0$ donc la fonction exponentielle est strictement convexe sur $\mathbb{R}$.
- $\forall x \in \mathbb{R}^{+*}, \ln''(x) = -\dfrac{1}{x^2} < 0$ donc ln est strictement concave sur $\mathbb{R}^{+*}$.

La convexité de ces fonctions permet d'établir des inégalité classiques, dites "inégalité de convexité" :

Proposition 32. *Soit $f : \mathbb{R} \to \mathbb{R}$ dérivable, f est convexe sur $I \subset \mathbb{R}$ si et seulement si f se situe au dessus de ses tangentes en tout point de I. C'est-à-dire,*

$$\forall x_0, x \in I, f(x) \geq f'(x_0)(x - x_0) + f(x_0)$$

Preuve :
($\Longrightarrow$) Supposons f convexe et $x_0 \in I$,
Soit $\varphi : x \mapsto f(x) - (f'(x_0)(x - x_0) + f(x_0))$, φ est dérivable par hypothèse, et pour tout $x \in I$, $\varphi'(x) = f'(x) - f'(x_0)$.
Comme f est convexe, f' est croissante, donc φ' est négatif pour $x < x_0$ et est positif pour $x > x_0$ et s'annule en $x = x_0$ donc atteint son maximum en x_0. Ainsi, $\forall x \in I, \varphi(x) \geq \varphi(x_0) = 0$ d'où le résultat.
($\Longleftarrow$) Supposons que f se situe au dessus de ses tangentes. Soit $x < y$, alors on a $f(y) \geq f'(x)(y-x)+f(y)$ d'où $\frac{f(y)-f(x)}{y-x} \geq f'(x)$. De plus $f(x) \geq f'(y)(x-y)+f(y)$ d'où $\frac{f(y)-f(x)}{y-x} \leq f'(y)$ d'où $f'(x) \leq f'(y)$. f' est donc croissante d'où f est convexe. □

Inégalité de convexité de fonctions usuelles
- exp est convexe sur $\mathbb{R}$ donc se situe au dessus de sa tangente en 0 : Ainsi, on a
$$\forall x \in \mathbb{R}, e^x \geq 1 + x$$
- $x \mapsto \ln(1 + x)$ est concave sur $\mathbb{R}^+$ donc se situe en dessous de sa tangente en 0 : Ainsi, on a
$$\forall x \in \mathbb{R}^+, \ln(1 + x) \leq x$$
- Sur $[0, \pi]$, on a $\sin'' = -\sin \leq 0$, donc sin se situe en dessous de sa tangente en 0 : Ainsi, on a
$$\forall x \in [0, \pi], \sin x \leq x$$

2.4 Optimisation des fonctions convexes

La convexité donne une régularité forte sur la fonction qui permet de rendre suffisante des conditions qui sont normalement suffisantes :

Proposition 33. *Soit $f : \mathbb{R} \to \mathbb{R}$ de classe C^2 et strictement convexe sur I et $x_0 \in I$: $f'(x_0) = 0$ (x_0 est un point critique de f) si et seulement si f admet un minimum local en x_0*

Preuve : Il est toujours vrai que si une fonction (pas forcément convexe) admet un minimum local en x_0 alors $f'(x_0) = 0$. La réciproque est vraie pour les fonctions convexes :
Au voisinage de x_0 en effectuant un développement limité de f on a

$$f(x) - f(x_0) = f''(x_0)\frac{(x-x_0)^2}{2} + \mathrm{o}((x-x_0)^2)$$

Comme $f''(x_0)\frac{(x-x_0)^2}{2} > 0$ alors $f(x) - f(x_0) \sim f''(x_0)\frac{(x-x_0)^2}{2} > 0$ donc de même signe au voisinage de x_0 donc $f(x) > f(x_0)$ au voisinage de x_0. $\square$

On peut également être plus précis, en effet, dans le cas des fonctions convexes, les notions de minimum local et global coïncident :

Proposition 34. *Soit f une fonction convexe sur I. Si f admet un minimum local en $x_0 \in I$ alors f admet un minimum global en $x_0 \in I$.*

Preuve : Supposons que x_0 est un minimum local de f, donc il existe $r > 0$ tel que $\forall x \in]x_0 - r, x_0 + r[\cap I, f(x) \geq f(x_0)$.
Supposons par l'absurde que x_0 n'est pas globale, alors il existe $y \in I$ tel que $f(y) < f(x_0)$. Par convexité, pour tout $t \in [0,1]$, $f(tx_0 + (1-t)y) \leq tf(x_0) + (1-t)f(y) \leq f(x_0)$ car $f(y) \leq f(x_0)$. Or, comme $tx_0 + (1-t)y \xrightarrow[t \to 1]{} x_0$, il existe t_0 tel que $tx_0 + (1-t)y \in]x_0 - r, x_0 + r[\cap I$. Absurde car x_0 est un minimum local de f. $\square$

3 Convexité des fonctions de plusieurs variables

3.1 Définition et propriétés

Définition 35. Soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction définit sur $\mathbb{R}^d$. On dit que f est convexe si pour tout $x, y \in \mathbb{R}^d$, $t \in [0,1]$, $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$

Remarque 36. Ainsi, une fonction est convexe si les restrictions de f sur chaque segment de la forme $[x, y]$ (qui est donc une fonction d'une seule variable) est convexe.

Pour déduire des résultats sur les fonctions de plusieurs variables à partir des fonctions d'une seule variable on utilise la caractérisation suivante :

Proposition 37. *Soit $f : \mathbb{R}^d \to \mathbb{R}$, f est convexe si et seulement si $\forall x, y \in \mathbb{R}^d$, la fonction*

$$\begin{array}{rccc} \varphi_{xy} & : & \mathbb{R} & \to & \mathbb{R} \\ & & t & \mapsto & f(tx + (1-t)y) \end{array}$$

est convexe.

Remarque 38. Il est important de réaliser que l'on relie ici la convexité d'une fonction de plusieurs variables à la convexité d'une infinité de fonctions d'une seule variable.

Preuve : ($\Longrightarrow$) Supposons f convexe. Soit $x, y \in \mathbb{R}^d$, montrons que φ_{xy} est convexe. Soit $t_1, t_2 \in \mathbb{R}, \lambda \in [0,1]$,

$$\begin{aligned} \varphi_{xy}(\lambda t_1 + (1-\lambda)t_2) &= f((\lambda t_1 + (1-\lambda)t_2)x + (1 - (\lambda t_1 + (1-\lambda)t_2))y) \\ &= f(\lambda(t_1 x + (1-t_1)y) + (1-\lambda)(t_2 x + (1-t_2)y) \\ &\leq \lambda f(t_1 x + (1-t_1)y) + (1-\lambda)f(t_2 x + (1-t_2)y) \\ &\leq \lambda \varphi_{xy}(t_1) + (1-\lambda)\varphi_{xy}(t_2) \end{aligned}$$

($\Longleftarrow$) Supposons que pour tout $x, y \in \mathbb{R}^d$, $t \in [0,1]$, φ_{xy} est convexe. En utilisant que la fonction φ_{xy} est en dessous de sa corde de $[0,1]$: pour tout $t \in [0,1]$:

$$\begin{aligned} \varphi(t) &\leq t\varphi(1) + (1-t)\varphi(0) \\ f(tx + (1-t)y) &\leq tf(x) + (1-t)f(y) \end{aligned}$$

$\square$

On peut remarquer que l'on peut lier la différentiabilité de f en tout point à la dérivabilité des fonctions φ_{xy} avec le théorème des fonctions composées :
En effet, soit $x, y \in \mathbb{R}^d$, comme la fonction $g : t \mapsto tx + (1-t)y$ est dérivable sur $\mathbb{R}$, par théorème des fonctions composées on a que f est différentiable sur $\mathbb{R}^d$ si et seulement si φ_{xy} est dérivable sur $[0,1]$ avec la relation suivante : pour $a, h \in \mathbb{R}^d$

$$\begin{aligned} \varphi'_{xy}(t) &= \mathrm{d}f(g(t)).(g'(t)) \\ &= \langle \nabla f(g(t)), g'(t) \rangle \\ &= \langle \nabla f(tx + (1-t)y), x - y \rangle \end{aligned}$$

Ainsi,

Proposition 39. *Soit $f : \mathbb{R}^d \to \mathbb{R}$ différentiable sur $\mathbb{R}^d$. f est convexe si et seulement si $\forall x, y \in \mathbb{R}^d$,*

$$f(y) \geq f(x) + \mathrm{d}f(x).(y-x) = f(x) + \langle \nabla f(x), y-x \rangle$$

Preuve : Il suffit de passer par la fonction auxiliaire : f est convexe si et seulement si $\forall x, y \in \mathbb{R}^d, \varphi_{xy}$ est convexe et φ_{xy} est convexe si et seulement si la courbe de φ_{xy} se situe au dessus de ses tangentes.
Si $\forall x, y \in \mathbb{R}^d, \varphi_{xy}$ se trouve au dessus de ses tangentes, en particulier, elle se situe au dessus de sa tangente en 0 d'où $\forall x, y \in \mathbb{R}^d, \varphi_{xy}(0) \geq \varphi'_{xy}(1)(0-1) + \varphi_{xy}(1)$ d'où $f(x) \geq -\mathrm{d}f(x).(x-y) + f(y)$ d'où $f(y) \geq f(x) + \mathrm{d}f(x).(y-x)$.
Réciproquement, $\nabla^2 f(x)$ $\square$