Chapitre 20

Fonctions convexes

L'interprétation géométrique sera toutefois la même.

L'objectif de ce chapitre est de définir plus rigoureusement les notions de fonctions convexes et concaves puis de démontrer les résultats sur ces fonctions, déjà évoqués dans le chapitre 4, et de les compléter.

Soit I un intervalle non vide de $\mathbb{R}$ non réduit à un point. Soit f une fonction définie sur I et à valeurs réelles.

I Préliminaires géométriques

On munit le plan d'un repère orthonormé.

1) Paramétrage de segments

Ce résultat assure que $[a\,;b]$ peut être paramétré par les réels de $[0\,;1]$. Le réel t en question est unique et vaut $t = \dfrac{x-a}{b-a}$.

Proposition. *Soit $(a,b) \in I^2$ avec $a < b$. On a*

$$M \in [a\,;b] \quad\Longleftrightarrow\quad \exists t \in [0\,;1]\,,\quad x = (1-t)a + tb.$$

DÉMONSTRATION.

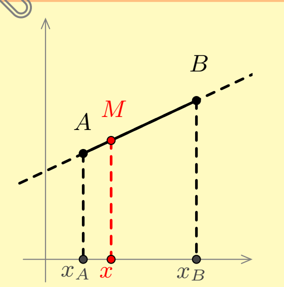
$\square$

Ce résultat assure que le segment $[AB]$ peut être paramétré par les réels de $[0\,;1]$. Le réel t en question est unique et vaut $\dfrac{x - x_A}{x_B - x_A}$.

Proposition. *Soient $A(x_A, y_A)$, $B(x_B, y_B)$ et $M(x,y)$ trois points du plan tels que $x_A < x_B$. On a :*

$$M \in [AB] \quad\Longleftrightarrow\quad \exists t \in [0\,;1]\,,\quad \begin{cases} x = (1-t)x_A + t x_B \\ y = (1-t)y_A + t y_B \end{cases}$$

DÉMONSTRATION.



$\square$

2) Arcs et cordes d'une fonction

Soit $(a,b) \in I^2$ avec $a < b$. Notons $A(a, f(a))$ et $B(b, f(b))$.

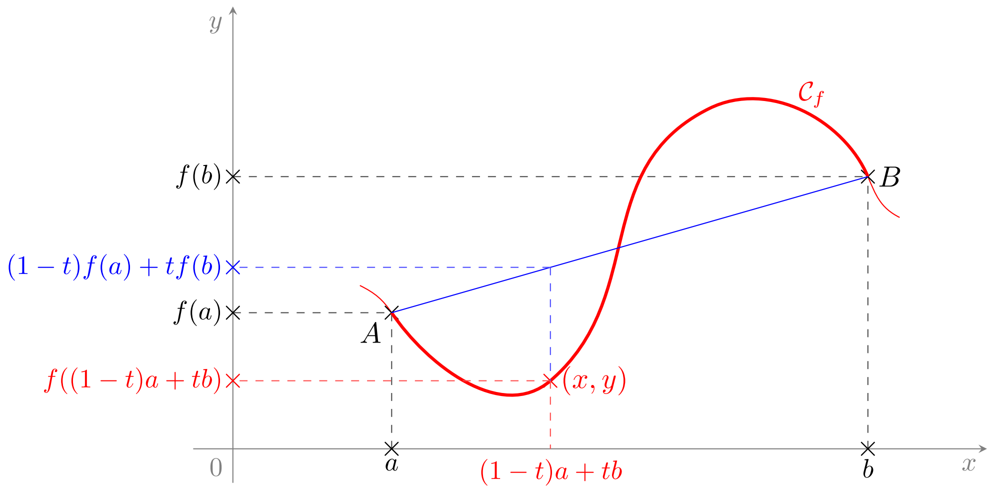
L'arc de $\mathcal{C}_f$ entre a et b est la portion de $\mathcal{C}_f$ comprise entre A et B.

Définition. *La partie $\{(x, f(x)) \mid x \in [a\,;b]\}$ du plan est appelé arc de $\mathcal{C}_f$ entre a et b. Le segment $[AB]$ est appelé corde ou sécante de cet arc.*

On déduit des propositions précédentes que, pour tout $x \in [a\,;b]$:

- Il existe $t \in [0\,;1]$ tel que $x = (1-t)a + tb$.
- Le point d'abscisse x sur la courbe de f a donc pour ordonnée $f(x) = f((1-t)a+tb)$.
- Le point d'abscisse x du segment $[AB]$ a pour ordonnée $(1-t)f(a) + tf(b)$.

> Sur cette figure, la partie de la courbe en rouge gras correspond à l'arc de f entre a et b. Le segment en bleu est la corde de cet arc.



II Résultats généraux sur les fonctions convexes

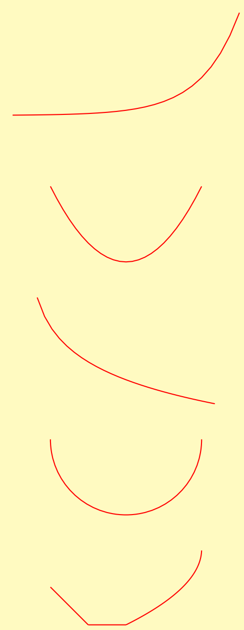
1) Notion de fonctions convexes et concaves

> L'inégalité dans la définition d'une fonction convexe est appelée inégalité de convexité.

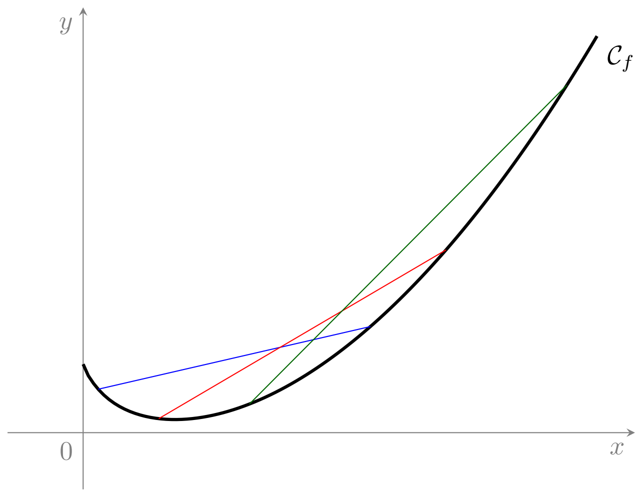
Définition (fonction convexe). *On dit que f est convexe sur I si*

$$\forall (a,b) \in I^2, \quad \forall t \in [0\,;1]\,, \qquad f\big((1-t)a+tb\big) \leqslant (1-t)f(a) + tf(b).$$

Il découle du paragraphe préliminaire une caractérisation géométrique d'une fonction convexe :

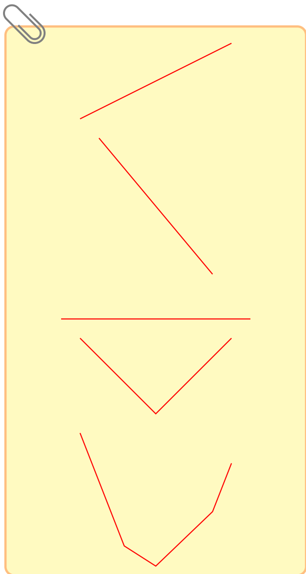
> Une fonction convexe est une fonction « qui sourit ». Quelques allures de graphes de fonctions convexes :
>
> 

Théorème. *La fonction f est convexe si et seulement si tout arc de $\mathcal{C}_f$ est en-dessous de sa corde.*



Exemples :

- *Il est immédiat que les fonctions affines (i.e. les fonctions du type $x \longmapsto \alpha x + \beta$, avec $(\alpha,\beta) \in \mathbb{R}^2$) sont convexes (l'inégalité de convexité est alors une égalité).*



- *La fonction $x \longmapsto |x|$ est convexe sur $\mathbb{R}$.*

  En effet, pour tous $(a,b) \in I^2$ et $t \in [0\,;1]$, l'inégalité triangulaire donne

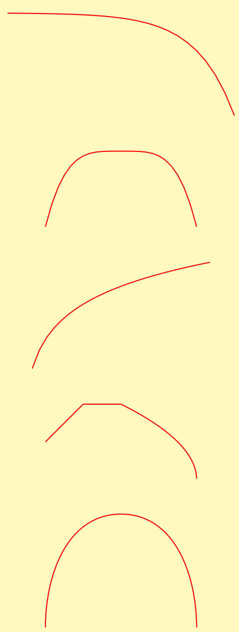
  $$|(1-t)a + tb| \leqslant \big|(1-t)a\big| + |tb| = (1-t)|a| + t|b|,$$

  puisque $t \geqslant 0$ et $1 - t \geqslant 0$.
- *La fonction $x \longmapsto x^2$ est convexe sur $\mathbb{R}$.*

Remarques :

- L'inégalité dans la définition est large : un arc de la courbe peut être confondu avec sa corde (comme on le voit dans les six derniers exemples ci-contre). Il existe une notion de fonction strictement convexe mais nous ne l'utiliserons pas.
- ⚠ Comme on peut le voir avec la fonction valeur absolue, une fonction convexe n'est pas forcément monotone ni dérivable.

Une fonction concave est une fonction « qui fait la tête ». Quelques allures de graphes de fonctions concaves :

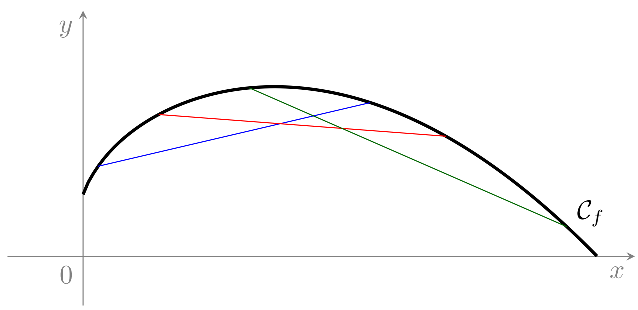


Définition (fonction concave). *On dit que f est concave sur I si*

$$\forall (a,b) \in I^2, \quad \forall t \in [0\,;1]\,, \qquad f\big((1-t)a + tb\big) \geqslant (1-t)f(a) + tf(b).$$

Il découle du paragraphe préliminaire une caractérisation géométrique d'une fonction concave :

Théorème. *La fonction f est concave si et seulement si tout arc de $\mathcal{C}_f$ est au-dessus de sa corde.*



La proposition suivant est immédiate :

Proposition. *La fonction f est concave si et seulement si la fonction $-f$ est convexe.*

Exemple : *Les fonctions $x \longmapsto -|x|$ et $x \longmapsto -x^2$ sont concaves sur $\mathbb{R}$.*

En vertu de cette proposition, nous allons nous intéresser surtout aux fonctions convexes dans ce chapitre puisqu'on se ramène facilement à une fonction concave en prenant l'opposé.

Remarques :

- Les fonctions affines sur I sont les seules fonctions qui sont à la fois convexes et concaves sur I.

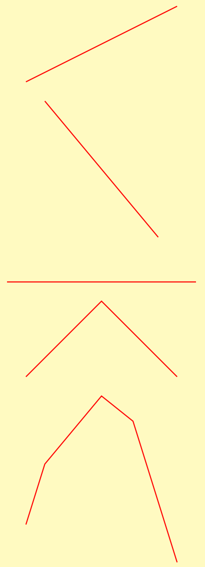
  $\rightsquigarrow$ Démonstration laissée en exercice.

- ⚠ Une fonction qui n'est pas convexe n'est pas forcément concave. Une fonction qui n'est pas concave n'est pas forcément convexe. En effet il existe des fonctions qui ne sont ni convexes ni concaves (autrement dit la négation de « f est convexe » n'est pas « f est concave »).

  Par exemple $f : x \longmapsto x^3$ n'est ni convexe, ni concave sur $\mathbb{R}$. En effet, si on prend $a = -2$ et $b = 2$,

  — *si* $t = \dfrac{1}{4}$, $f((1-t)a + tb) = f(-1) = -1 > -4 = (1-t)f(a) + tf(b)$,

  — *si* $t = \dfrac{3}{4}$, $f((1-t)a + tb) = f(1) = 1 < 4 = (1-t)f(a) + tf(b)$.



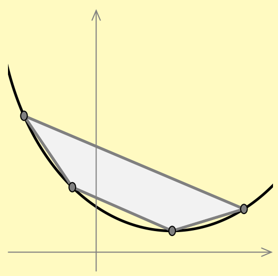
2) Inégalité de Jensen

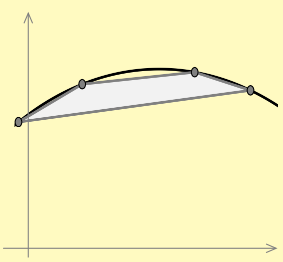
Théorème. *La fonction f est convexe sur I si et seulement si pour tout $n \geqslant 2$, pour tous $(x_1, \ldots, x_n) \in I^n$ et $(\lambda_1, \ldots, \lambda_n) \in [0\,;1]^n$ vérifiant $\lambda_1 + \cdots + \lambda_n = 1$, on a*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i).$$

Remarque : Ce théorème est toujours valable pour une fonction concave, mais en changeant le sens des inégalités (ce qui s'obtient en considérant $-f$ au lieu de f puisque f est concave si et seulement si $-f$ est convexe).

Interprétation géométrique. Si on choisit des points sur le graphe de f, le polygone qu'ils forment est au-dessus du graphe pour une fonction convexe, et en dessous pour une fonction concave.





Démonstration. Le sens indirect est immédiat (il suffit de prendre $n = 2$, $\lambda_1 = 1 - t$, $\lambda_2 = t$, $x_1 = a$ et $x_2 = b$ et on retrouve la définition d'une fonction convexe). Montrons le sens direct.

□

Remarque : Si $n=2$, l'inégalité devient

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

avec $\lambda_1 + \lambda_2 = 1$. Si maintenant $x_1 = a$, $x_2 = b$ et $\lambda_2 = t$, alors $\lambda_1 = 1-t$ et donc

$$f((1-t)a + tb) \leqslant (1-t)f(a) + tf(b).$$

Exemple :

Soit $n \in \mathbb{N}\backslash\{0,1\}$. Si on prend $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \dfrac{1}{n}$, on a $\lambda_1 + \cdots + \lambda_n = 1$ et on obtient un cas particulier très utile du (sens direct du) théorème :

L'image par une fonction convexe de la moyenne des points est inférieure à la moyenne des images.

Corollaire. *Si f est une fonction convexe sur I alors, pour tout $n \in \mathbb{N}^*$,*

$$\forall (x_1, \ldots, x_n) \in I^n, \qquad f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leqslant \frac{f(x_1) + \cdots + f(x_n)}{n}.$$

Exemples :

- *Puisque la fonction carré est convexe sur $\mathbb{R}$,*

- *Nous avons vu dans le chapitre 4 (avec des résultats admis certes mais que nous allons démontrer dans le paragraphe III de ce chapitre) que la fonction $\ln$ est concave sur $\mathbb{R}_+^*$.*

La quantité $A(x_1, \ldots, x_n)$ est appelée la moyenne arithmétique des réels $x_1, \ldots, x_n$ et la quantité $G(x_1, \ldots, x_n)$ la moyenne géométrique des réels $x_1, \ldots, x_n$. Cette inégalité est appelée « inégalité arithmético-géométrique ».

3) Opérations sur les fonctions convexes

En général les propriétés de convexité ne se propagent pas quand on fait des opérations sur les fonctions. Voyons juste un cas qui est classique en pratique :

Un produit, un quotient, une composée de fonctions convexes n'est pas forcément convexe. Nous en verrons des contre-exemples en exercice.

Proposition. *Soit $n \in \mathbb{N}^*$. Soient $f_1, \ldots, f_n$ des fonctions convexes sur I et $\alpha_1, \ldots, \alpha_n$ des réels* ***positifs****. Alors $\alpha_1 f_1 + \cdots + \alpha_n f_n$ est convexe sur I.*

Démonstration. Soient $(x,y) \in I^2$ et $\lambda \in [0\,;1]$. Soit $i \in [\![1\,;n]\!]$. La fonction f_i étant convexe,

$$f_i(\lambda x + (1-\lambda)y) \leqslant \lambda f_i(x) + (1-\lambda) f_i(y)$$

De plus, $\alpha_i \geqslant 0$ donc $\alpha_i f_i(\lambda x + (1-\lambda)y) \leqslant \lambda \alpha_i f_i(x) + (1-\lambda)\alpha_i f_i(y)$. Par somme

$$\left(\sum_{i=1}^{n} \alpha_i f_i\right)(\lambda x + (1-\lambda y)) \leqslant \lambda \left(\sum_{i=1}^{n} \alpha_i f_i\right)(x) + (1-\lambda)\left(\sum_{i=1}^{n} \alpha_i f_i\right)(y).$$

Ainsi $\alpha_1 f_1 + \cdots + \alpha_n f_n$ est convexe sur I. □

En d'autres termes, une combinaison linéaire à coefficients positifs de fonctions convexes est une fonction convexe.

⚠ Le résultat de la proposition est faux si les coefficients ne sont pas positifs ! En effet, si f est convexe et si $\alpha < 0$ alors αf est concave donc n'est pas convexe (sauf si f est affine).

Corollaire. *Une somme de fonctions convexes est convexe.*

La proposition ci-dessus et ce corollaire ci-contre sont encore valables avec des fonctions concaves (et toujours des coefficients **positifs**).

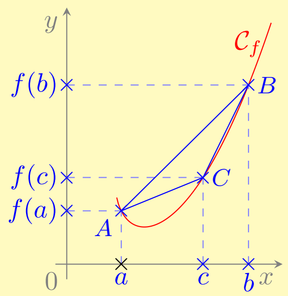
4) Croissance des pentes

Lemme (inégalité des pentes). *Soit f une fonction convexe sur I. Soit $(a,b,c) \in I^3$ est tel que $a < c < b$, alors*

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

Interprétation graphique : Si on note A le point de coordonnées $(a, f(a))$, B le point de coordonnées $(b, f(b))$ et C le point de coordonnées $(c, f(c))$, alors

pente de (AC) $\leqslant$ pente de (AB) $\leqslant$ pente de (BC)



Remarque : Si f est concave sur I, les deux inégalités du lemme sont dans l'autre sens (ce qui s'obtient en considérant $-f$ au lieu de f).

Démonstration.

□

Proposition (croissance des pentes). *La fonction f est convexe sur I si et seulement si, pour tout $x_0 \in I$, la fonction*

$$\tau_{x_0} : \begin{cases} I\backslash\{a\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{f(x)-f(x_0)}{x-x_0} \end{cases}$$

est croissante sur $I\backslash\{x_0\}$.

Remarque : La fonction f est concave sur I si et seulement si, pour tout $x_0 \in I$, τ_{x_0} est décroissante sur $I\backslash\{x_0\}$ (ce qui s'obtient en considérant $-f$ au lieu de f).

Démonstration.

$\square$

5) Régularité des fonctions convexes (HP)

Avant de commencer, notons $\overset{o}{I}$, l'intérieur de I, c'est-à-dire l'intervalle I privé de ses éventuelles bornes réelles.

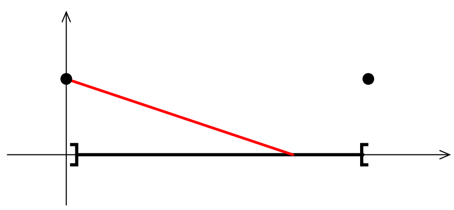
Plus précisément, pour tous réels a et b tels que $a < b$,
- Si $I = [a\,;b]$, $[a\,;b[$, $]a\,;b]$ ou $]a\,;b[$, alors $\overset{o}{I} =]a\,;b[$.
- Si $I = [a\,;+\infty[$ ou $]a\,;+\infty[$, alors $\overset{o}{I} =]a\,;+\infty[$.
- Si $I =]-\infty\,;a]$ ou $]-\infty\,;a[$, alors $\overset{o}{I} =]-\infty\,;a[$.
- Si $I = \mathbb{R}$, alors $\overset{o}{I} = \mathbb{R}$.

Les résultats de ce paragraphe sont hors-programme mais classiques et utiles dans de nombreux exercices. Ils sont de plus très accessibles. Il est donc vivement conseillé de savoir les redémontrer.

On a vu plus haut qu'une fonction convexe n'est pas forcément dérivable (c'est le cas de la fonction valeur absolue). Elle n'est pas non plus continue a propri.

Ci-dessous le graphe d'une fonction convexe non continue sur $[0\,;4]$. Plus précisément, elle est discontinue en 0 et en 4, mais est tout de même convexe sur $[0\,;4]$ (je vous laisse démontrer que son graphe est bien en-dessous de ses cordes).

Il s'agit de la fonction nulle sur $]0\,;4[$ qui vaut 1 en 0 et en 4. La courbe est en noir et, en rouge, nous avons tracé une corde.



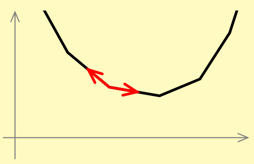
Cependant une fonction convexe est tout de même assez régulière :

Proposition (HP). *Une fonction convexe sur I est dérivable à gauche et à droite en tout point de $\overset{o}{I}$.*

C'est encore vrai pour une fonction concave.

Démonstration.

Une fonction dérivable à droite et à gauche n'est pas forcément dérivable (contrairement aux fonctions continues). Ci-dessous le graphe d'une fonction convexe non dérivable (mais on observe bien qu'elle admet des dérivées à gauche et à droite en doit point intérieur :



Rappelons qu'une fonction est dérivable en x_0 si et seulement si f est dérivable à droite et à gauche **et** si $f'_d(x_0) = f'_g(x_0)$.

□

Corollaire (HP). *Une fonction convexe sur I est continue sur $\overset{o}{I}$.*

Démonstration. Supposons que f est convexe sur I. Soit $x_0 \in \overset{o}{I}$. Puisque f est dérivable à gauche et à droite en x_0, elle est continue à gauche et à droite en x_0 et donc elle est continue en x_0. □

Terminons par un résultat qui fait le chemin inverse, en quelque sorte :

Proposition (HP). *Une fonction convexe sur $\overset{o}{I}$ qui est continue sur I est convexe sur I tout entier.*

Démonstration.

III Fonctions convexes ou concaves dérivables

1) Caractérisations des fonctions convexes ou concaves dérivables

Théorème. *Supposons que f est dérivable sur I. Alors :*

1. *f est convexe sur I si et seulement si f' est croissante sur I.*
2. *f est concave sur I si et seulement si f' est décroissante sur I.*

Démonstration. Montrons le point 1 (l'autre est analogue).

$\square$

Corollaire. *Supposons que f est deux fois dérivable sur I. Alors :*

1. *f est convexe sur I si et seulement si f'' est positive sur I.*
2. *f est concave sur I si et seulement si f'' est négative sur I.*

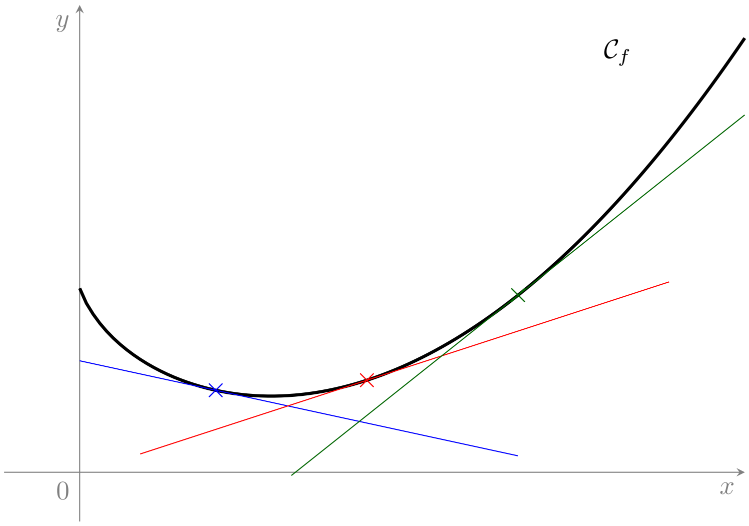
Ne pas confondre f'' et f', en d'autres termes, ne pas confondre une fonction convexe (qui vérifie $f'' \geqslant 0$) et une fonction croissante (qui vérifie $f' \geqslant 0$). Il n'y a aucun lien entre la convexité et la monotonie d'une fonction !

Exemples : *Nous renvoyons aux chapitres 4 et 5 pour des exemples.*

2) Convexité et tangentes

Théorème. *Supposons que f est dérivable sur I. Alors :*

1. *f est convexe si et seulement si son graphe est au-dessus de ses tangentes, c'est-à-dire,*
$$\forall (x_0, x) \in I^2, \qquad f(x) \geqslant f(x_0) + f'(x_0)(x - x_0).$$
2. *f est concave si et seulement si son graphe est en dessous de ses tangentes, c'est-à-dire,*
$$\forall (x_0, x) \in I^2, \qquad f(x) \leqslant f(x_0) + f'(x_0)(x - x_0).$$



DÉMONSTRATION. Montrons le point 1 (l'autre est analogue).

□

Exemples : *Nous renvoyons aux chapitres 4 et 5 pour des exemples.*

Pour une fonction concave f, on a, pour tout $x_0 \in I$, $f'(x_0) = 0$ si et seulement si f admet un maximum global en x_0.

Remarque : On a vu dans le chapitre 19 qu'une fonction dérivable sur I admettant un extremum local un point intérieur de I, a une dérivée nulle en ce point. C'est une condition nécessaire non suffisante d'existence d'un extremum local. Quand on travaille avec une fonction f convexe dérivable sur I, la situation est beaucoup plus simple, et on a même une CNS pour le minimum : pour tout $x_0 \in \mathring{I}$, $f'(x_0) = 0$ si et seulement si f admet un minimum global en x_0. Montrons-le : soit $x_0 \in I$.

- Supposons que $f'(x_0) = 0$. Comme $\mathcal{C}_f$ est au dessus de sa tangente en x_0, on a

$$\forall x \in I, \qquad f(x) \geqslant f(x_0) + f'(x_0)(x - x_0) = f(x_0).$$

  Ainsi f admet un minimum global en x_0.
- Réciproquement, supposons que f admet un minimum global en x_0. Comme x_0 est un point intérieur et f est dérivable en x_0, c'est un point critique : $f'(x_0) = 0$.

Cette implication est aussi vraie si x_0 n'est pas un point intérieur mais pas la réciproque, cf. remarque plus bas.

En d'autres termes : sauf aux bornes, une fonction convexe non localement constante (c'est-à-dire « qui ne fait pas de palier ») admet comme seul extremum local éventuel un minimum global (cela se voit bien sur le graphe de la fonction carré par exemple).

⚠ f peut n'admettre aucun point critique (par exemple l'exponentielle sur $\mathbb{R}$) et ce n'est plus valable si x_0 est une extrémité de I.

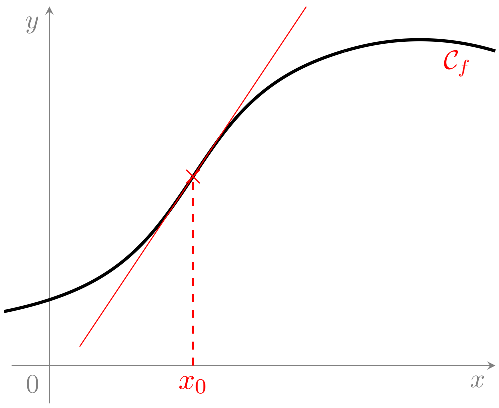
Par exemple, la fonction définie sur $[0\,;1]$ par $f : x \longmapsto e^x$ admet un maximum global en 1 et un minimum global en 0 alors que $f'(0) \neq 0$.

3) Point d'inflexion

Soit x_0 un point de I qui n'est pas une extrémité de I.

Définition (point d'inflexion). *On dit que $(x_0, f(x_0))$ est un point d'inflexion de la courbe représentative $\mathcal{C}_f$ de f (ou que $\mathcal{C}_f$ admet un point d'inflexion en x_0) si il existe $\alpha > 0$ tel que f est convexe sur l'intervalle $[x_0 - \alpha, x_0]$ et concave sur l'intervalle $[x_0, x_0 + \alpha]$, ou le contraire.*

Autrement dit $\mathcal{C}_f$ admet un point d'inflexion en x_0 si, au voisinage de x_0, f passe de convexe à concave ou de concave à convexe en x_0. Si f est dérivable sur I alors, graphiquement, cela se traduit par le fait que $\mathcal{C}_f$ traverse sa tangente en le point $(x_0, f(x_0))$.



⚠ Il existe des fonctions convexes qui ne sont pas deux fois dérivables. Dans ce cas, il faut revenir à la définition d'un point d'inflexion.

Il découle des résultats des paragraphes précédents que :

Proposition. *Supposons que f soit deux fois dérivable sur I. La fonction f'' s'annule et change de signe en x_0 si et seulement si la courbe représentative de f admet un point d'inflexion en x_0.*

Exemples : *Nous renvoyons aux chapitres 4 et 5 pour des exemples.*