Convexité d'ordre {1,2} et points critiques

Le cours sur la différentielle a permis de généraliser les notions de nombre dérivé en un point et de dérivabilité d'une fonction à des fonctions à plusieurs variables. Mais le lien entre point optimal, dérivée d'une fonction et convexité reste encore à expliciter.

1) Points critiques d'une fonction

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable en $x_0 \in \mathbb{R}^n$

On dit que x_0 est un point critique de f si et seulement si $\boxed{\nabla f(x_0) = 0}$

On appelle lieu critique de f l'ensemble de ses points critiques

Remarque : pour $f : \mathbb{R} \to \mathbb{R}$, x_0 point critique $\Leftrightarrow f'(x_0) = 0$

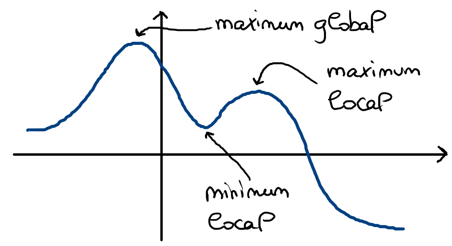
pour $f : \mathbb{R}^n \to \mathbb{R}^m$, x_0 point critique $\Leftrightarrow \mathrm{Jac} f(x_0) = 0 \Leftrightarrow \dfrac{\partial f_i}{\partial x_j}(x_0) = 0 \;\; \forall i = 1, \dots, m$ et $\forall j = 1, \dots, n$

Définition : extremum local

On dit que x_0 est un extremum local ssi $\exists R > 0$ tq $\forall y \in B(x_0, R)$:

$\to$ soit $f(x_0) \leq f(y)$ $\to$ x_0 est un minimum local

$\to$ soit $f(x_0) \geq f(y)$ $\to$ x_0 est un maximum local



Le lien entre extremum local et point critique se fait par la propriété suivante :

Propriété : Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable en x_0

Si f admet un extremum local en x_0, alors $\nabla f(x_0) = 0$ (s'appelle théorème de Fermat pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$)

$\to$ les extremums locaux d'une fonction sont des points critiques

⚠ La réciproque est fausse : un point peut être critique sans pour autant être un extremum local (par exemple $x \mapsto x^3$ en 0)

Dans le cas général, on va avoir besoin d'une caractérisation plus fine des points critiques que celle offerte uniquement par l'annulation du gradient

La convexité permet de simplifier nettement cette étude

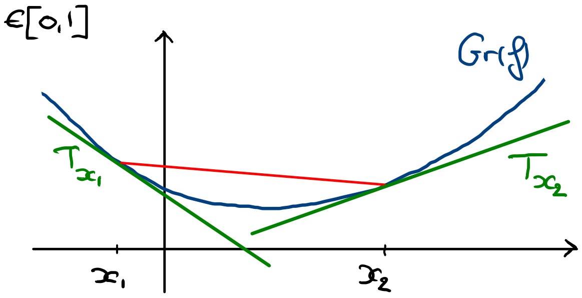
2) Caractérisation à l'ordre 1 de la convexité

Repartons de l'exemple d'une fonction $f : \mathbb{R} \to \mathbb{R}$ convexe

Caractérisation à l'ordre 0 : $f(tx_1 + (1-t)x_2) \leq t f(x_1) + (1-t) f(x_2) \;\; \forall x_1, x_2 \in D_f$ et $\forall t \in [0,1]$

$\to$ peu importe les points x_1 et x_2, le segment reliant $(x_1, f(x_1))$ et $(x_2, f(x_2))$ est au dessus du graphe de f

On peut voir aussi que les tangentes au graphe en $(x_1, f(x_1))$ et $(x_2, f(x_2))$ sont toujours au dessous du graphe



Rappel : l'équation de la tangente au graphe de f au point $(x_0, f(x_0))$ s'écrit

$T_{x_0} : x \mapsto f(x_0) + (x - x_0) f'(x_0)$ ou (avec $h = x - x_0$) $T_{x_0} : h \mapsto f(x_0) + h f'(x_0)$

→ Pour une fonction $f : \mathbb{R} \to \mathbb{R}$, quel que soit le point x où est calculée la tangente, le graphe de f est au dessus de la tangente

Ce qui s'écrit formellement : $\forall x, y,\ f(y) \geq T_x(y)$

$$\iff \forall x, y,\ \boxed{f(y) \geq f(x) + (y-x) f'(x)}$$

$$\forall x, y,\ \boxed{f(y) - f(x) \geq (y-x) f'(x)}$$

→ caractérisation à l'ordre 1 de la convexité pour $f : \mathbb{R} \to \mathbb{R}$

On peut en avoir une écriture équivalente :

$\forall x, y,\ f(y) \geq T_x(y)$ et $f(x) \geq T_y(x)$

$\to \quad f(y) \geq f(x) + (y-x) f'(x)$ et $f(x) \geq f(y) + (x-y) f'(y)$

$\iff \quad f(y) + f(x) \geq f(x) + f(y) + (y-x) f'(x) + (x-y) f'(y)$

$\iff \quad 0 \geq (y-x) f'(x) + (x-y) f'(y)$

$\iff \quad \boxed{(x-y)(f'(x) - f'(y)) \geq 0}$ → la fonction f' est monotone croissante : $x \leq y \Rightarrow f'(x) \leq f'(y)$

Ces propriétés se généralisent pour une fonction $f : \mathbb{R}^n \to \mathbb{R}$

Propriété : caractérisation à l'ordre 1 de la convexité

Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable. Les propositions suivantes sont équivalentes :

* f est convexe
* $\forall x, y,\ f(y) \geq f(x) + \nabla f(x)^T (y - x) \iff f(y) - f(x) \geq \nabla f(x)^T (y-x)$
* $\forall x, y,\ (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0$

La caractérisation à l'ordre 1 de la convexité permet d'obtenir un résultat immédiat sur la nature des points critiques d'une fonction convexe :

Soit x_0 un point critique de f convexe.

$\to \nabla f(x_0) = 0$ (par définition d'un point critique)

$\to \forall y,\ f(y) \geq f(x_0) + \underbrace{\nabla f(x_0)^T}_{=0} (y - x_0)$ (par caractérisation à l'ordre 1 de la convexité)

$\Rightarrow \forall y,\ \boxed{f(y) \geq f(x_0)}$ $\Rightarrow$ x_0 est un point optimal

Tout point critique de f est donc un point optimal, donc un minimum global

Pour les fonctions convexes, on a donc l'équivalence suivante :

$$\boxed{x_0 \text{ est optimal (minimum global)} \iff \nabla f(x_0) = 0 \text{ (point critique)}}$$