OCVX1 : un peu de calcul différentiel

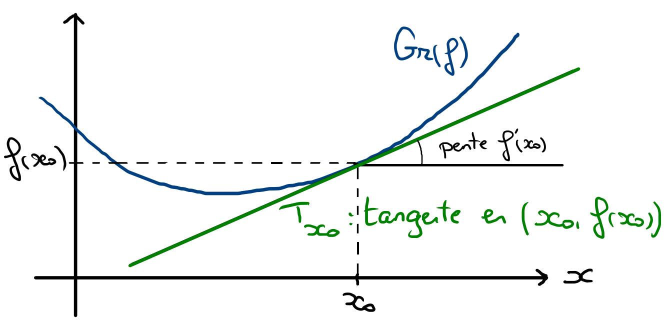
Vu lors du cours d'introduction : pour trouver le point optimal de $f(x) = ax^2+bx+c$ $(a>0)$, on cherche le point d'annulation de la dérivée : $f'(x) = 2ax+b \to f'(x^*) = 0 \iff x^* = -\frac{b}{2a}$

Pour $f(x) = x^4 - 2x^2$ en revanche, les zéros de la dérivée ne sont pas forcément des points optimaux ($f'(0) = 0$ mais $x=0$ n'est pas un point optimal). Par contre, les points optimaux ($x^* = \pm 1$) sont des zéros de la dérivée

L'équivalence x^* optimal $\iff f'(x^*) = 0$ dans le premier cas n'est vraie que parce que la fonction est convexe. Mais quoi qu'il en soit, il y a un lien à expliciter entre point critique (zéro de la dérivée), point optimal, dérivée en un point et convexité de la fonction.

1) Rappels sur la dérivée : le cas d'une fonction $f : \mathbb{R} \to \mathbb{R}$

On appelle nombre dérivé de f en x_0 la limite, si elle existe, du rapport $\boxed{f'(x_0) = \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}}$



$\to f'(x_0) \equiv$ pente de la tangente au graphe de f au point $(x_0, f(x_0))$

<u>Equation de la tangente</u> : $T_{x_0} : x \mapsto f(x_0) + (x-x_0)f'(x_0)$

$T_{x_0} : h \mapsto f(x_0) + hf'(x_0)$

<u>Propriété</u> : Si le nombre dérivé $f'(x_0)$ existe, alors f est continue en x_0

On appelle <u>dérivée</u> de f la fonction $f' : x \mapsto f'(x)$ f' associe à tout point $x \in D_f$ la pente de la tangente au point $(x, f(x))$ du graphe de f $Gr(f)$

Mais la notion de nombre dérivé (et de fonction dérivée) ne s'étend pas aux fonctions $f : \mathbb{R}^n \to \mathbb{R}$ puisque l'écriture $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x) \to \in \mathbb{R} \checkmark\checkmark}{h \to \in \mathbb{R}^n \times\times}$ ne se généralise pas si $f : \mathbb{R}^n \to \mathbb{R}$

<u>Rappel</u> : On dit que f est négligeable devant g au point a, et on note $f = o_a(g)$ ("petit o") s'il existe une fonction $\varepsilon : \mathbb{R} \to \mathbb{R}$ avec $\varepsilon(x) \xrightarrow[x\to a]{} 0$ telle que $f(x) = \varepsilon(x)g(x)$ au voisinage de a

$\to$ c'est équivalent à $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$ si g ne s'annule pas au voisinage de a (sauf éventuellement en a)

$\to$ les notations $f = o_a(g)$ et $f(x) = \varepsilon(x)g(x)$ sont équivalentes

<u>Lien entre nombre dérivé et développement limité</u>

Si f est dérivable en x_0, de dérivée $f'(x_0)$: $\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) \iff \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)-hf'(x_0)}{h} = 0$

$\iff f(x_0+h) - f(x_0) - hf'(x_0) = o_0(h)$

Donc $f(x_0+h) = \underbrace{f(x_0)+hf'(x_0)}_{\text{équation de la tangente } T_{x_0}} + o_0(h)$ c'est l'écriture du développement limité à l'ordre 1 de f en x_0

$\to$ l'approximation de f par sa tangente est d'autant meilleure que h est petit

⚠ Dans l'écriture du DL1, la variable est h

2) La différentielle

Dans l'écriture du DL1 $f(x_0+h) = f(x_0) + hf'(x_0) + o_0(h)$, on retrouve la tangente $h \mapsto hf'(x_0)$ (application linéaire)
Cette idée s'étend en dimension n et permet de définir la notion de <u>différentiabilité</u> d'une fonction $f: \mathbb{R}^n \to \mathbb{R}$

<u>Définition</u> : on dit que $f: \mathbb{R}^n \to \mathbb{R}$ est <u>différentiable</u> en x_0 s'il existe une application linéaire $df_{x_0}: \mathbb{R}^n \to \mathbb{R}$ (se note aussi $d_{x_0}f$) telle que : $\boxed{f(x_0+h) = f(x_0) + df_{x_0}(h) + \|h\|\varepsilon(h)}$ avec $\begin{array}{l}\varepsilon: \mathbb{R}^n \to \mathbb{R} \\ \quad h \mapsto \varepsilon(h)\end{array}$ et $\varepsilon(h) \xrightarrow[\|h\| \to 0]{} 0$

<u>Pour $f: \mathbb{R} \to \mathbb{R}$</u>. Si f est différentiable en x_0 : $f(x_0+h) = f(x_0) + df_{x_0}(h) + o_0(h)$ avec df_{x_0} une application linéaire
$\to df_{x_0}: h \mapsto \alpha h$ avec $\alpha \in \mathbb{R}$
Donc $f(x_0+h) = f(x_0) + \alpha h + o_0(h) \iff \dfrac{f(x_0+h) - f(x_0)}{h} = \alpha + o_0(1)$ $\to \varepsilon(h) \xrightarrow[h \to 0]{} 0$

$$\iff \frac{f(x_0+h) - f(x_0)}{h} \xrightarrow[h \to 0]{} \boxed{\alpha = f'(x_0)}$$

Donc si f est différentiable en x_0, f est dérivable en x_0 et $df_{x_0}: h \mapsto hf'(x_0)$
f continue en x_0 (puisque dérivable $\Rightarrow$ continue)

⚠ Différentielle et dérivée sont deux choses différentes (même si les deux notions sont liées)

<u>Exemple</u> $\begin{array}{l} f: \mathbb{R} \to \mathbb{R} \\ \quad x \mapsto x^2 \end{array}$ En un point x_0 : $f'(x_0) = 2x_0$ $\to$ c'est la pente de la tangente au graphe de f au point $(x_0, x_0^2 = f(x_0))$

Que vaut la différentielle en x_0 de $f: x \mapsto x^2$?
$f(x_0+h) = (x_0+h)^2 = x_0^2 + 2x_0 h + h^2 = f(x_0) + 2x_0 h + h^2$ $\to h \times h = h\,\varepsilon(h)$ avec $\varepsilon(h) \xrightarrow[h \to 0]{} 0$
↳ c'est une application linéaire en h $\to$ c'est donc la différentielle
Donc la différentielle de f en x_0 est <u>$df_{x_0}: h \mapsto 2x_0 h = hf'(x_0)$</u>

Pour une fonction $f: \mathbb{R} \to \mathbb{R}$, la différentielle est donc l'application linéaire "tangente" en un point du graphe donné, alors que la dérivée en ce point est la pente de cette tangente

<u>Pour $f: \mathbb{R}^n \to \mathbb{R}$</u> : pour trouver la différentielle en un point donné, il faut (pour le moment) en revenir à la définition : on écrit / linéarise $f(x_0+h)$ et on identifie le terme linéaire par rapport à h

<u>Exemple</u> : $\begin{array}{l} f: \mathbb{R}^n \to \mathbb{R} \\ \quad x \mapsto x^T x \end{array}$ $f(x_0+h) = (x_0+h)^T(x_0+h) = x_0^T x_0 + x_0^T h + h^T x_0 + h^T h = \underbrace{x_0^T x_0}_{f(x_0)} + \underbrace{2x_0^T h}_{\text{linéaire en } h} + \underbrace{\|h\|^2}_{o_0(h)}$

La différentielle en x_0 est donc l'application linéaire <u>$df_{x_0}: h \mapsto 2x_0^T h$</u>

3) Lien entre différentielle et gradient

<u>Rappel</u> : Dérivée partielle en un point d'une fonction $f: \mathbb{R}^n \to \mathbb{R}$
Soit $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$. On dit que la k^e dérivée partielle de f en a existe, et on note $\frac{\partial f}{\partial x_k}(a)$ $(\equiv \partial_k f(a))$ si l'application $\varphi: t \mapsto f(a_1, \dots, a_k+t, \dots, a_n)$ est dérivable en 0, auquel cas $\frac{\partial f}{\partial x_k}(a) = \varphi'(0)$