Cette définition correspond bien à l'intuition de la dérivée partielle : on dérive par rapport à la k^e variable en maintenant fixées les $(k-1)^e$ autres variables.

$(a_1, \dots, a_k + t, \dots, a_n)$ peut se réécrire $a + te_k$ avec $e_k = (0, \dots, 1, \dots, 0)^T \in \mathbb{R}^n$ le k^e vecteur de la base canonique

Donc $\left\| \begin{array}{l} \varphi(t) = f(a + te_k) \iff \text{variations de } f \text{ le long du } k^e \text{ axe} \\ \varphi'(0) = \dfrac{\partial f}{\partial x_k}(a) \iff \text{dérivée partielle au point } a \text{ selon l'axe } (0, e_k) \end{array} \right.$

⚠ En dimension 1, l'existence de $f'(x_0)$ implique la continuité de f en x_0

Cette propriété n'est plus vraie pour $f : \mathbb{R}^n \to \mathbb{R}$

<u>Exemple</u> : $f : \mathbb{R}^2 \to \mathbb{R}$

$$(x_1, x_2) \longmapsto \begin{cases} 0 & \text{si } (x_1, x_2) = (0,0) \\ \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{sinon} \end{cases}$$

On a vu (cf cours sur la convexité à l'ordre 0) que f n'est pas continue en $(0,0)$

Pourtant, $f(0,t) = f(t,0) = 0 \ \forall t \in \mathbb{R} \to \dfrac{\partial f}{\partial x_1}(0,0) = \dfrac{\partial f}{\partial x_2}(0,0) = 0$

Dérivée directionnelle pour $f : \mathbb{R}^n \to \mathbb{R}$

La dérivée partielle est une dérivée selon une direction (un axe de la base canonique)

Cette idée se généralise pour une direction quelconque : on dit que f admet une dérivée en x_0 suivant un vecteur $u \in \mathbb{R}^n \setminus \{0\}$ si l'application $\varphi : t \mapsto f(x_0 + tu)$ est dérivable en 0

La dérivée $\varphi'(0) = \lim\limits_{t \to 0} \dfrac{f(x_0 + tu) - f(x_0)}{t}$ se note $D_u f(x_0)$: c'est la <u>dérivée directionnelle</u> de f en x_0 suivant le vecteur u

Évidemment : $\dfrac{\partial f}{\partial x_k}(x_0) = D_{e_k} f(x_0)$

Si on remplace u par αu $(\alpha \neq 0)$: $\underline{D_{\alpha u} f(x_0) = \alpha D_u f(x_0)}$ → lorsque la dérivée directionnelle existe en un point suivant un vecteur, elle existe suivant tout vecteur de même direction (mais la valeur de la dérivée change)

→ on parle de dérivée <u>dans la direction</u> de u lorsque u est unitaire $(\|u\|_2 = 1)$

<u>Définition</u> : gradient de f en $x_0 \in \mathbb{R}^n$

les deux notations sont équivalentes

On appelle gradient de f en x_0 le vecteur $\nabla_{x_0} f = \nabla f(x_0) = \left(\dfrac{\partial f}{\partial x_1}(x_0), \dots, \dfrac{\partial f}{\partial x_n}(x_0) \right)^T$

<u>Propriétés</u> : Soit $f : \mathbb{R}^n \to \mathbb{R}$ et $x_0 \in \mathbb{R}^n$. Si f est différentiable en x_0, alors

→ f est continue en x_0 (déjà vu si $f : \mathbb{R} \to \mathbb{R}$)

→ f est dérivable en x_0 suivant tout vecteur $u \in \mathbb{R}^n \setminus \{0\}$, et la dérivée directionnelle vaut $\boxed{D_u f(x_0) = df_{x_0}(u)}$

→ en particulier pour tous les vecteurs de la base canonique e_k $k = 1, \dots, n$: $\dfrac{\partial f}{\partial x_k}(x_0) = D_{e_k} f(x_0) = df_{x_0}(e_k)$

→ toutes les dérivées partielles existent

Soit maintenant $h \in \mathbb{R}^n$ un vecteur de coordonnées $(h_1, \dots, h_n)^T$ dans la base canonique : $h = \sum\limits_{i=1}^{n} h_i e_i$

Alors $df_{x_0}(h) = df_{x_0}\left(\sum_{i=1}^{n} h_i e_i\right) = \sum_{i=1}^{n} h_i\, df_{x_0}(e_i) = \sum_{i=1}^{n} h_i \underbrace{D_{e_i} f(x_0)}_{\frac{\partial f}{\partial x_i}(x_0)}$

$\rightarrow \quad df_{x_0}(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0)$

$\rightarrow \quad \boxed{df_{x_0}(h) = \nabla f(x_0)^T h = \langle \nabla f(x_0), h\rangle}$ Si f est différentiable en x_0, alors $df_{x_0} : h \mapsto \langle \nabla f(x_0), h\rangle = \nabla f(x_0)^T h$

Le DL_1 d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ s'écrit donc $f(x_0+h) = f(x_0) + df_{x_0}(h) + o_0(h) = f(x_0) + \nabla f(x_0)^T h + o_0(h)$

<u>Interprétation géométrique du gradient</u>

On considère f différentiable en x_0. À quoi correspond le vecteur h tel que $df_{x_0}(h) = \nabla f(x_0)^T h = 0$?

$\rightarrow f(x_0+h) = f(x_0) + \nabla f(x_0)^T h + o_0(h)$ puisque f différentiable en x_0

$= f(x_0) + o_0(h)$ pour h tel que $\nabla f(x_0)^T h = 0$

Donc pour $\|h\| \to 0$ (h petit) tel que $\nabla f(x_0)^T h = 0$, $f(x_0+h) = f(x_0)$

$\rightarrow$ h est la direction de l'isocontour de valeur $f(x_0)$

$\rightarrow$ <u>$\nabla f(x_0)$ est orthogonal à la ligne de niveau $f(x_0)$ de f</u>

Dans quelle direction pointe le gradient ? Prenons $h = \alpha \nabla f(x_0)$ pour $\alpha > 0$ (donc h colinéaire à $\nabla f(x_0)$)

Alors $f(x_0+h) = f(x_0 + \alpha \nabla f(x_0)) = f(x_0) + \nabla f(x_0)^T h + o_0(h)$

$= f(x_0) + \nabla f(x_0)^T (\alpha \nabla f(x_0)) + \|h\| \varepsilon(h)$

$\rightarrow f(x_0+h) = f(x_0 + \alpha \nabla f(x_0)) = f(x_0) + \alpha \underbrace{\nabla f(x_0)^T \nabla f(x_0)}_{\|\nabla f(x_0)\|^2} + \alpha \|\nabla f(x_0)\| \varepsilon(\alpha \nabla f(x_0))$

$= f(x_0) + \alpha \|\nabla f(x_0)\| \left(\|\nabla f(x_0)\| + \varepsilon(\alpha \nabla f(x_0))\right)$

même si $\varepsilon(\alpha \nabla f(x_0)) < 0$

Or $\varepsilon(\alpha \nabla f(x_0)) \xrightarrow[\alpha \to 0]{} 0$, donc pour $\alpha > 0$ suffisamment petit, on est sûrs que $\|\nabla f(x_0)\| + \varepsilon(\alpha \nabla f(x_0)) > 0$

Donc $f(x_0 + \alpha \nabla f(x_0)) > f(x_0)$ pour $h = \alpha \nabla f(x_0)$ avec $\alpha > 0$ suffisamment petit

$\rightarrow$ <u>$\nabla f(x_0)$ pointe vers les valeurs croissantes de f</u>

$\rightarrow$ et de manière plus générale : <u>le gradient indique la direction de plus forte pente</u> (sera démontré plus tard)

4) Cas général : la matrice jacobienne

Dans le cas général où $f : \mathbb{R}^n \to \mathbb{R}^p$, $x \mapsto (f_1(x), \dots, f_p(x))$ avec $\forall i = 1, \dots p$ $f_i : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto f_i(x)$ "fonction-coordonnée"

f différentiable en $x_0 \in \mathbb{R}^n \iff$ ses p fonctions coordonnées f_i, $i = 1, \dots p$ sont différentiables en $x_0 \in \mathbb{R}^n$

$$f(x_0+h) = \begin{pmatrix} f_1(x_0+h) \\ \vdots \\ f_p(x_0+h) \end{pmatrix} = \begin{pmatrix} f_1(x_0) + df_{1x_0}(h) + o_0(h) \\ \vdots \qquad \vdots \qquad \vdots \\ f_p(x_0) + df_{p x_0}(h) + o_0(h) \end{pmatrix} = \begin{pmatrix} f_1(x_0) \\ \vdots \\ f_p(x_0) \end{pmatrix} + \begin{pmatrix} df_{1x_0}(h) \\ \vdots \\ df_{p x_0}(h) \end{pmatrix} + o_0(h) = f(x_0) + \begin{pmatrix} df_{1x_0}(h) \\ \vdots \\ df_{p x_0}(h) \end{pmatrix} + o_0(h)$$

$\|h\| \varepsilon(h)$ avec $\varepsilon : \mathbb{R}^n \to \mathbb{R}^p$ et $\|\varepsilon(h)\| \to 0$ quand $\|h\| \to 0$