MT242, Cours n^o 9, Lundi 6 Mars 2000.

Exercice traité. Dans un espace vectoriel réel de dimension finie, si $v_n \to v$ et $w_n \to w$, alors $v_n + w_n \to v + w$; si de plus $\lambda_n \to \lambda$ dans $\mathbb{R}$, alors $\lambda_n v_n \to \lambda v$.

Rappel : voisinages et non-voisinages. Soit E un espace vectoriel de dimension finie ; pour que $\mathrm{W} \subset \mathrm{E}$ soit un “non-voisinage” du point $w \in \mathrm{E}$, il faut et il suffit qu'il existe une suite $(v_n) \subset \mathrm{E}$ qui tende vers w, et dont tous les termes sont en dehors de W,

$$w = \lim_n v_n \quad \text{et} \quad \forall n \geq 0, \quad v_n \notin \mathrm{W}.$$

2.3.1. Ensembles fermés

Définition 2.3.4. On dit qu'un sous-ensemble Φ de E est *fermé dans* E si pour toute suite (v_n) de points de Φ qui converge dans E vers un point v, la limite v reste dans Φ.

Exemples évidents.

L'ensemble vide est fermé dans E, E est fermé dans E. Tout sous-ensemble de E réduit à un seul point est fermé dans E.

Proposition 2.3.2. *Soit $f : \Phi \to \mathbb{R}$ une fonction continue définie sur une partie fermée Φ de* E *; l'ensemble*

$$\Phi_1 = \{v \in \Phi : f(v) \geq 0\}$$

est fermé dans E.

Démonstration. Si (v_n) est une suite dans Φ_1 qui converge vers $w \in \mathrm{E}$, on a d'abord $w \in \Phi$ parce que Φ est fermé dans E, puis $f(w) \geq 0$ par passage à la limite des inégalités larges $f(v_n) \geq 0$ (puisque f est continue au point $w \in \Phi$), donc $w \in \Phi_1$.

Plus généralement si $\Phi \subset \mathrm{E}$ est fermé dans E et si F est un espace vectoriel de dimension finie, l'image inverse d'un fermé Ψ de F par une fonction continue $f : \Phi \to \mathrm{F}$, à savoir

$$\Phi_1 = \{v \in \Phi : f(v) \in \Psi\}$$

est fermée dans E (consulter Liret-Martinais, ou bien traiter en exercice).

Exemples.

a. Toute boule fermée est un ensemble fermé dans E. Une sphère $\mathrm{S}(x, r) = \{y : d(x, y) = r\}$ est fermée dans E. La parabole d'équation $y = x^2$ est un fermé de $\mathbb{R}^2$.

b. Un ensemble peut très bien n'être ni ouvert ni fermé : $\mathrm{X} =]0, 1]$ par exemple n'est ni ouvert ni fermé dans $\mathbb{R}$. Il n'est pas ouvert parce que X n'est pas un voisinage de 1, et pas fermé parce que $0 \notin \mathrm{X}$ est limite de points de X.

Proposition 2.3.3. *Un ensemble Φ est fermé dans* E *si et seulement si son complémentaire* $\mathrm{E} \setminus \Phi$ *est ouvert dans* E.

Démonstration. Supposons $\Omega = \mathrm{E} \setminus \Phi$ non ouvert dans E. Il existe alors un point $w \in \Omega$ tel que $\mathrm{W} = \Omega$ ne soit pas voisinage de w. D'après le critère de non-voisinage, il existe une suite (v_n) formée de points extérieurs à Ω, c'est à dire $(v_n) \subset \Phi$ et qui converge vers w, avec $w \notin \Phi$. Donc Φ n'est pas fermé.

Inversement si Φ n'est pas fermé dans E, il existe une suite $(v_n) \subset \Phi$ qui converge vers une limite $w \notin \Phi$. Alors $w \in \Omega$ et Ω n'est pas un voisinage de w d'après la partie inverse du critère de non-voisinage. Donc Ω n'est pas voisinage de tous ses points, donc Ω n'est pas ouvert.

Exercice proposé, non traité. Toute intersection d'ensembles fermés est un ensemble fermé. Toute réunion **finie** de fermés est fermée.

Proposition 2.3.4. *Si Φ est un fermé non vide et majoré de $\mathbb{R}$, il possède un plus grand élément.*

Démonstration. Puisque Φ est majoré et non vide, il possède une borne supérieure $s = \sup \Phi$. On peut trouver une suite (x_n) d'éléments de Φ qui converge vers la borne supérieure s. Puisque Φ est fermé, on en déduit que $s \in \Phi$. Par définition de la borne supérieure, s majore tous les éléments de Φ. Puisque $s \in \Phi$, c'est le plus grand élément de Φ.

2.4. Compacité dans les espaces vectoriels de dimension finie

Définition 2.4.1. Un espace métrique $(\mathrm{X}, d_{\mathrm{X}})$ est *compact* si toute suite (x_n) de points de X admet une sous-suite (x_{n_k}) qui converge vers un point $x \in \mathrm{X}$, c'est à dire que l'on a $\lim_k d_{\mathrm{X}}(x_{n_k}, x) = 0$.

Exemples. L'ensemble $\emptyset$ est compact, de même que tout ensemble réduit à un seul point (trivial), ou tout ensemble fini (facile). L'intervalle fermé $[0, 1]$ est compact (avec la distance usuelle ; c'est Bolzano-Weierstrass). Mais $\mathrm{X} =]0, 1[$ n'est pas compact (la suite $(2^{-n}) \subset \mathrm{X}$ n'a pas de sous-suite convergente vers une limite **dans** X).

On considère dans la suite de cette section un cas très particulier d'espace métrique : on prendra un espace vectoriel réel E de dimension finie muni d'une norme. On a remarqué que les notions de convergence, d'ensemble borné, ne dépendent pas de la norme choisie.

Définition 2.4.2. Un sous-ensemble K de E est *compact* si toute suite (x_n) de points de K admet une sous-suite qui converge dans E vers un point $x \in \mathrm{K}$.

Théorème 2.4.1. *Soit* E *un espace vectoriel réel de dimension finie ; un sous-ensemble* K *de* E *est compact si et seulement s'il est fermé et borné.*

Démonstration. Supposons K compact, et vérifions d'abord que K est fermé. Soit donc (v_n) une suite de points de K, qui converge dans E vers un point v ; nous devons montrer que v appartient à K. Mais puisque K est compact, nous savons qu'il existe une sous-suite (v_{n_k}) de (v_n) qui converge vers un vecteur $w \in \mathrm{K}$; puisque v_n tend vers v, la sous-suite (v_{n_k}) converge aussi vers v, ce qui implique que $w = v$, donc $v \in \mathrm{K}$, et on a montré que K est fermé. Montrons maintenant que K est borné. Sinon, on pourrait trouver dans K une suite (v_n) telle que $\|v_n\|$ tende vers $+\infty$. On aurait aussi $\lim_k \|v_{n_k}\| = +\infty$ pour toute sous-suite, ce qui impliquerait qu'aucune sous-suite ne pourrait converger, d'où une contradiction de la propriété de compacité, ce qui montre que K doit être borné.

La partie réciproque utilise le théorème de Bolzano-Weierstrass dans un espace de dimension finie.

De toute suite bornée (v_n) dans un espace vectoriel réel de dimension finie E *on peut extraire une sous-suite convergente.*

Terminons la démonstration du fait que tout fermé borné $\mathrm{K} \subset \mathrm{E}$ est compact. Soit (v_n) une suite de points de K ; puisque K est borné, la suite (v_n) est bornée, donc d'après Bolzano-Weierstrass, on peut trouver une sous-suite convergente $(v_{n_k})_k$, de limite $v \in \mathrm{E}$. Mais puisque la sous-suite est formée de points de K et que K est fermé, on a $v \in \mathrm{K}$, donc on a bien trouvé une sous-suite qui converge vers un point de K.

Conséquence. Tout compact *non vide* de $\mathbb{R}$ possède un plus grand et un plus petit élément.

2.4.1. Propriétés des fonctions continues sur un compact

Théorème 2.4.2. *Soient* E, F *deux espaces vectoriels réels de dimension finie,* K *un compact de* E *et* $f : \mathrm{K} \to \mathrm{F}$ *une application continue ; l'image* $f(\mathrm{K})$ *est une partie compacte de* F*.*

Démonstration. Soit (y_n) une suite de points de $f(\mathrm{K})$; par définition de l'image de K par l'application f, on peut trouver $x_n \in \mathrm{K}$ tel que $y_n = f(x_n)$, et par compacité il existe une sous-suite $(x_{n_k})_k$ qui converge vers un point $x \in \mathrm{K}$. Puisque f est continue sur K, donc continue au point x, on sait que $y_{n_k} = f(x_{n_k})$ converge vers $y = f(x) \in f(\mathrm{K})$, et on a montré que $f(\mathrm{K})$ est compact.

Corollaire 2.4.1. *Si* f *est une fonction réelle définie et continue en tout point d'un compact non vide* $\mathrm{K} \subset \mathrm{E}$ *(où* E *est de dimension finie), elle est bornée et elle atteint son maximum et son minimum sur* K*.*

Démonstration. L'image $f(\mathrm{K}) \subset \mathbb{R}$ est compacte, donc bornée et fermée, et elle est non vide puisque K est non vide. Mais on sait qu'un compact non vide de $\mathbb{R}$ possède un plus grand élément (et un plus petit élément), donc il existe $\mathrm{M} \in f(\mathrm{K})$ qui est le plus grand élément de $f(\mathrm{K})$. Par définition de $f(\mathrm{K})$, il existe $x \in \mathrm{K}$ tel que $f(x) = \mathrm{M}$, et f atteint donc son maximum sur K au point x. Le raisonnement est le même pour le minimum.

Corollaire 2.4.2. *Soit* Φ *un fermé non borné de* E *(vectoriel de dimension finie) et soit* f *une fonction réelle continue sur* Φ *telle que* $f(v)$ *tende vers* $+\infty$ *lorsque* $\|v\| \to +\infty$ *(par valeurs* $v \in \Phi$*). La fonction* f *atteint son minimum sur l'ensemble* Φ*.*

Démonstration. Puisque Φ est non vide, on peut trouver un point $v_0 \in \Phi$. Posons

$$\mathrm{K} = \{v \in \Phi : f(v) \leq f(v_0)\}.$$

Cet ensemble est fermé (défini par une inégalité large et une fonction continue), non vide ($v_0 \in \mathrm{K}$) et borné parce que f tend vers $+\infty$ lorsque $\|v\| \to +\infty$, ce qui entraîne qu'il existe R tel que $(\|v\| > \mathrm{R}) \Rightarrow \big(f(v) > \mathrm{A} = f(v_0)\big)$, donc $\|v\| > \mathrm{R}$ implique $v \notin \mathrm{K}$, ce qui revient à dire que tout les points de K vérifient $\|v\| \leq \mathrm{R}$. Il résulte de tout ça que K est compact et non vide, donc f atteint son minimum sur K en un point $v_1 \in \mathrm{K}$. Il est clair que $f(v_1)$ est aussi la plus petite valeur de f sur Φ tout entier, puisque pour tout $v \in \Phi \setminus \mathrm{K}$ on a $f(v_1) \leq f(v_0) < f(v)$.

Exercice. Théorème de D'Alembert. On identifie $\mathbb{C}$ avec $\mathbb{R}^2$. Soit $\mathrm{P}(z)$ un polynôme de degré ≥ 1 ; il est facile de voir que la fonction $f(x, y) = |\mathrm{P}(x + iy)|$ est continue sur $\mathbb{R}^2$, et qu'elle tend vers $+\infty$ quand $\sqrt{x^2 + y^2} = |z| \to +\infty$, donc cette fonction atteint son minimum sur $\mathbb{R}^2$ en un point (x_0, y_0) qui correspond à un $z_0 = x_0 + iy_0$. Nous devons montrer que $f(x_0, y_0) = |\mathrm{P}(z_0)| = 0$, c'est à dire que z_0 est une racine de P. Dans le cas contraire, posons $\mathrm{Q}(h) = \mathrm{P}(z_0 + h)/\mathrm{P}(z_0)$. C'est un polynôme en h de degré ≥ 1, tel que $\mathrm{Q}(0) = 1$ et dont le module atteint son minimum en $h = 0$. Ecrivons

$$\mathrm{Q}(h) = 1 + a_k h^k + \cdots + a_n h^n,$$

avec $k > 0$ et $a_k \neq 0$. Posons $a_k = r\,\mathrm{e}^{i\theta}$, $r > 0$, et utilisons des points h tendant vers 0 de la forme

$$h_t = t\,\mathrm{e}^{-i(\theta+\pi)/k},$$

pour t réel > 0 petit. On voit que $a_k h_t^k = rt^k \,\mathrm{e}^{-i\pi} = -rt^k$ pour un tel choix de h_t. En prenant t assez petit pour que $rt^k < 1$, on aura

$$|\mathrm{Q}(h_t)| \le (1 - rt^k) + t^{k+1}|\mathrm{B}(t)|,$$

où $\mathrm{B}(t)$ est une expression polynomiale qui reste bornée quand t tend vers 0. Il en résulte que $|\mathrm{Q}(h_k)| < 1$ pour $t > 0$ petit, une contradiction du fait que $1 = \mathrm{Q}(0) = |\mathrm{Q}(0)|$ était le minimum de $|\mathrm{Q}|$ sur $\mathbb{C}$.

Cours n° 10, Mercredi 8 Mars 2000.

Chapitre 3. Introduction au calcul différentiel

3.1. Dérivées partielles

Commençons par un exemple très simple. On définit une fonction f en tout point $(s,t) \in \mathbb{R}^2$ par la formule

$$f(s,t) = (s + 1 - t)^2 + (s - t^2)^2.$$

Que veut-on dire par l'expression "calculer la dérivée par rapport à la variable s", en un point fixé (s_0, t_0) ? C'est très simple. On fixe la valeur de l'autre variable (la variable t) à sa valeur au point $a_0 = (s_0, t_0)$, et on fait varier seulement s. On obtient ainsi une fonction g_1 d'une seule variable, définie par

$$g_1(s) = f(s, t_0) = (s + 1 - t_0)^2 + (s - t_0^2)^2$$

dont on calcule la dérivée ordinaire, pour la valeur $s = s_0$, ce qui donne

$$g_1'(s_0) = 2(s_0 + 1 - t_0) + 2(s_0 - t_0^2).$$

On dit cette valeur est la *dérivée partielle de f, au point a_0, par rapport à la première variable* et on note

$$\frac{\partial f}{\partial s}(a_0) = \frac{\partial f}{\partial s}(s_0, t_0) = 2(s_0 + 1 - t_0) + 2(s_0 - t_0^2).$$

On a de même

$$\frac{\partial f}{\partial t}(a_0) = -2(s_0 + 1 - t_0) - 4t_0(s_0 - t_0^2).$$

Passons au cas général. Soient $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, V un voisinage de a et $f : \mathrm{V} \to \mathbb{R}$ une fonction réelle ; considérons

$$\lim_{t \to 0} \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, \ldots, a_n)}{t} \in \mathbb{R}.$$

Il s'agit de la dérivée au point $a_1 \in \mathbb{R}$, au sens habituel, de la fonction d'une seule variable $s \to f(s, a_2, \ldots, a_n)$, fonction qui est définie pour s voisin de a_1, et qui est obtenue en fixant les variables $a_2, \ldots, a_n$ et en ne laissant varier que la première coordonnée ; on appelle cette fonction la première *fonction partielle* associée à la fonction donnée f. On appelle la quantité

$$\lim_{t \to 0} \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, \ldots, a_n)}{t}$$

(quand elle existe) la *première dérivée partielle* de f au point a, et on la note $\dfrac{\partial f}{\partial x_1}(a)$. On définit de même les autres dérivées partielles $\dfrac{\partial f}{\partial x_j}(a)$, pour $j = 2, \ldots, n$. Il est parfois commode d'employer une autre notation, $\mathrm{D}_j f(a)$,

$$\mathrm{D}_j f(a) = \frac{\partial f}{\partial x_j}(a).$$

Exemples traités.

1. Calcul des trois dérivées partielles de la fonction f définie sur $\mathbb{R}^3$ par $f(x, y, z) = z\,\mathrm{e}^{x+y^2}$.

2. Dérivées partielles en $(0, 0)$ pour la fonction $f(x, y) = xy/(x^2 + y^2$ si $(x, y) \neq 0$, et $f(0, 0) = 0$. Dans un tel exemple on ne se lance pas dans le calcul algébrique des dérivées partielles, calcul possible en tous les points $\neq (0, 0)$. Pour le point $(0, 0)$, on revient à la définition. On voit que $f(x, 0) = 0$ pour tout x et $f(0, y) = 0$ pour tout y, donc les deux dérivées partielles existent au point $(0, 0)$ et elles sont nulles. Pourtant on avait vu que cet exemple est une fonction qui n'est pas continue au point $(0, 0)$. Cet exemple montre que l'existence de toutes les dérivées partielles de f en un point n'entraîne même pas la continuité de f en ce point.

3.1.1. Condition nécessaire pour un extremum local

Supposons que g soit une fonction réelle d'une variable réelle, définie dans un voisinage V_1 de x_0. Si on sait que $g(x) \leq g(x_0)$ pour tout point $x \in \mathrm{V}_1$ et si la dérivée $g'(x_0)$ existe, il en résulte que $g'(x_0) = 0$.

Rappelons très rapidement le principe de la démonstration. Puisque V_1 est un voisinage de x_0, il existe $r > 0$ tel que $]x_0 - r, x_0 + r[\subset \mathrm{V}_1$. On peut donc tendre vers x_0 par les deux côtés. Si $x_0 < x < x_0 + r$, on aura $(g(x) - g(x_0))/(x - x_0) \leq 0$ parce que $g(x) \leq g(x_0)$ et $x - x_0 > 0$, donc $g'(x_0) \leq 0$ en passant à la limite par la droite ; mais si $x_0 - r < x < x_0$, on aura $(g(x) - g(x_0))/(x - x_0) \geq 0$, ce qui donne $g'(x_0) \geq 0$ par limite à gauche, donc $g'(x_0) = 0$.

Proposition 3.1.1. *Si une fonction réelle f, définie sur un voisinage $\mathrm{V} \subset \mathbb{R}^n$ d'un point $a \in \mathbb{R}^n$, vérifie*

$$\forall v \in \mathrm{V}, \quad f(v) \leq f(a)$$

et si la dérivée partielle $\mathrm{D}_i f(a)$ au point a existe, alors

$$\mathrm{D}_i f(a) = 0.$$

Démonstration. Prenons $n = 2$ et $i = 1$; posons $a = (x_0, y_0)$. Puisque V est un voisinage de a, il existe une boule ouverte euclidienne $\mathrm{B}(a, r)$, $r > 0$ qui est contenue dans V. Il en résulte que

$$\mathrm{V}_1 = \{x \in \mathbb{R} : (x, y_0) \in \mathrm{V}\}$$

contient $]x_0 - r, x_0 + r[$, donc V_1 est un voisinage de x_0. Posons aussi $g(x) = f(x, y_0)$. Pour tout $x \in \mathrm{V}_1$, on a $g(x) = f(x, y_0) \leq f(x_0, y_0) = g(x_0)$ (parce que $v = (x, y_0) \in \mathrm{V}$) et $g'(x_0)$ existe, et vaut $\mathrm{D}_1 f(x_0, y_0)$. D'après le principe en dimension un, on sait que $g'(x_0) = 0$, donc $\mathrm{D}_1 f(a) = 0$.

Bien sûr le principe est valable pour un minimum aussi (considérer $-f$) et dans la plupart des exemples toutes les dérivées partielles existeront. On cherchera donc les candidats extrema de f parmi les points qui annulent toutes les dérivées partielles de f.

Un exemple

On considère dans le plan $\mathbb{R}^2$ la parabole (P) d'équation $y = x^2$ et la droite (D) d'équation $y = x - 1$. On voudrait trouver le couple de points (d, p) d'un point $d \in \mathrm{D}$ et $p \in \mathrm{P}$ qui réalise le minimum de la distance possible.

On paramétrise tous les points de P par $p_t = (t, t^2)$, $t \in \mathbb{R}$, et tous les points de D par $d_s = (s + 1, s)$, $s \in \mathbb{R}$. Alors

$$f(s, t) = \|p_t - d_s\|_2^2 = (s + 1 - t)^2 + (s - t^2)^2$$

est la fonction dont on a calculé les dérivées partielles précédemment. Le couple optimal doit correspondre à une paire (s, t) qui annule les deux dérivées partielles de f. On trouve $t = 1/2$ et $s = -1/8$, d'où $d = (7/8, -1/8)$ et $p = (1/2, 1/4)$. On vérifie ensuite que le vecteur $\overrightarrow{dp} = (-3/8, 3/8)$ est orthogonal à la droite (D), et aussi orthogonal à la tangente à la parabole au point p.

3.2. La différentielle

Que peut-on espérer quand tout va bien ? On pourra passer du point $a = (x_0, y_0)$ à un point voisin $(x_0 + s, y_0 + t)$ en deux coups, en deux petits déplacements le long de chaque axe de coordonnées. On aura

$$f(x_0 + s, y_0) = f(x_0, y_0) + \mathrm{D}_1 f(a)\, s + \mathrm{E}_1(s),$$
$$f(x_0 + s, y_0 + t) = f(x_0 + s, y_0) + \mathrm{D}_2 f(x_0 + s, y_0)\, t + \mathrm{E}_2(s, t)$$

où E_1 et E_2 sont des "petites erreurs". Si tout va bien on pourra dire que $\mathrm{D}_2 f(x_0 + s, y_0)$ est presque égal à $\mathrm{D}_2 f(a)$, et on aura en posant $h = (s, t)$ et $d_j = \mathrm{D}_j f(a)$

$$f(a + h) = f(a) + d_1\, s + d_2\, t + \mathrm{E}(h),$$

où l'opération n'a d'intérêt que si le terme d'erreur $\mathrm{E}(h)$ est négligeable par rapport aux termes déjà écrits. Ce qui est important est de remarquer que le terme principal qui décrit l'accroissement de f est une *fonction linéaire* du déplacement h. C'est cette idée qui conduit à la définition suivante.

Définition 3.2.1. Soient E et F deux espaces vectoriels de dimension finie, a un point de E, V un voisinage de a et f une application de V dans F ; on dit que f est *différentiable au point a* s'il existe une application linéaire $\ell \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ telle que

$$f(a + h) = f(a) + \ell(h) + \|h\|\, \overrightarrow{\varepsilon(h)}$$

où $\|\overrightarrow{\varepsilon(h)}\| \to 0$ quand h tend vers 0_E dans l'espace vectoriel E (noter que $\overrightarrow{\varepsilon(h)}$ est un vecteur de F).

On peut encore dire que f est différentiable au point a si l'erreur

$$\mathrm{E}(h) = f(a + h) - f(a) - \ell(h)$$

commise quand on remplace $f(a + h)$ par son approximation affine $f(a) + \ell(h)$, a une norme majorée par une formule du type

$$\|\mathrm{E}(h)\| \le \|h\|\, \varepsilon(h),$$

où ε est maintenant une fonction à valeurs réelles ≥ 0, qui doit tendre vers 0 lorsque h tend vers 0_E.

Quand f est différentiable au point a, l'application linéaire ℓ est uniquement déterminée ; en effet, pour tout vecteur $v \in \mathrm{E}$, on a $\ell(v) = \lim_{t\to 0} t^{-1}\big(f(a+tv) - f(a)\big)$, ce qui montre que ℓ est complètement connue. Quand la limite

$$\lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}$$

existe pour un vecteur $v \in \mathrm{E}$, on l'appelle la *dérivée de f au point a dans la direction du vecteur v* et on peut la noter $\mathrm{D}_v f(a)$.

Notation. On notera $(df)_a$ l'application linéaire ℓ ci-dessus (on note aussi souvent df_a –sans parenthèses– et parfois $d_a f$). On dit que $(df)_a$ est la *différentielle de f au point a.*

Remarque. On vient de voir que pour tout vecteur $v \in \mathrm{E}$,

$$(df)_a(v) = \mathrm{D}_v f(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}.$$

3.2.1. Différentielle et dérivées partielles

Soient $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, V un voisinage de a et $f : \mathrm{V} \to \mathbb{R}$ une fonction réelle différentiable au point a ; l'application $(df)_a$ est une forme linéaire sur $\mathbb{R}^n$, dont on peut calculer la matrice (de taille $1 \times n$) par rapport à la base canonique $\mathbf{e}$ de $\mathbb{R}^n$ (définie par $e_1 = (1, 0, \ldots), \ldots$). Cette matrice qu'on notera $(\mathrm{J}f)_a$ (matrice jacobienne de f au point a) contient dans chaque colonne l'image par l'application linéaire $(df)_a$ du vecteur de base correspondant. Ainsi, dans la première colonne de $\mathrm{J} = (\mathrm{J}f)_a$ on a

$$\mathrm{J}_{1,1} = (df)_a(e_1) = \lim_{t\to 0} \frac{f(a_1 + t, a_2, \ldots, a_n) - f(a_1, \ldots, a_n)}{t} \in \mathbb{R}.$$

On reconnaît la première dérivée partielle de f au point a, que nous avons notée $\dfrac{\partial f}{\partial x_1}(a)$. On trouve de même les autres coefficients de J, qui sont les autres dérivées partielles $\dfrac{\partial f}{\partial x_j}(a)$, pour $j = 2, \ldots, n$. La matrice (ligne) de l'application linéaire $(df)_a$ est donc égale (dans la base canonique de $\mathbb{R}^n$) à

$$(\mathrm{J}f)_a = \begin{pmatrix} \mathrm{D}_1 f(a) & \ldots & \mathrm{D}_n f(a) \end{pmatrix}$$

et l'image d'un vecteur $h = (h_1, \ldots, h_n)$ s'obtient par produit matriciel usuel ligne-colonne

$$(df)_a(h) = \mathrm{D}_1 f(a)\, h_1 + \cdots + \mathrm{D}_n f(a)\, h_n,$$

de sorte que si on désigne par A, H les matrices colonne de a, h et si on pose aussi $\mathrm{F}(\mathrm{V}) = f(v)$ pour toute matrice colonne V représentant un vecteur v, on aura

$$\mathrm{F}(\mathrm{A} + \mathrm{H}) = \mathrm{F}(\mathrm{A}) + (\mathrm{J}f)_a\, \mathrm{H} + \mathrm{E}(\mathrm{H}).$$

En résumé

Si $f : \mathrm{V} \to \mathbb{R}$ est différentiable au point a, toutes les dérivées partielles existent au point a. La matrice de la forme linéaire $(df)_a$ dans la base canonique est la matrice ligne contenant les dérivées partielles de f au point a.

Le théorème suivant est très important. C'est lui qui permet de passer dans l'autre sens, des dérivées partielles à la différentiabilité. On a dit que la seule existence des dérivées partielles de f au point a ne peut pas garantir la différentiabilité. Il faut donc supposer plus : l'existence dans tout un voisinage, et la continuité au point a pour toutes les dérivées partielles.

Théorème 3.2.1. *Soit* V *un voisinage d'un point* $a \in \mathbb{R}^n$ *et soit* $f : \mathrm{V} \to \mathbb{R}$ *; si* f *admet en tout point* v *du voisinage* V *des dérivées partielles* $\mathrm{D}_j f(v)$*, pour tout* $j = 1, \ldots, n$ *et si toutes ces fonctions dérivées partielles* $v \to \mathrm{D}_j f(v)$ *sont continues au point* a*, la fonction* f *est différentiable au point* a*.*

Démonstration. Ecrivons la démonstration dans le cas $n = 2$ pour simplifier. Posons $a = (x_0, y_0) \in \mathbb{R}^2$ et $h = (s, t)$. On va utiliser le théorème des accroissements finis pour les fonctions d'une variable réelle g, continues sur $[x_0, x_1]$ et dérivables sur $]x_0, x_1[$, qui dit que

$$g(x_1) - g(x_0) = (x_1 - x_0)\, g'(c)$$

où c est un point entre x_0 et x_1. On va passer de $a = (x_0, y_0)$ à $a + h = (x_0 + s, y_0 + t)$ en deux coups, comme on l'a dit au début de la section. En utilisant $g_1(x) = f(x, y_0)$ entre x_0 et $x_1 = x_0 + s$, on trouve

$$f(x_0 + s, y_0) - f(x_0, y_0) = g_1(x_0 + s) - g_1(x_0) = s\, g_1'(c(s)) = s\, \mathrm{D}_1 f(\mathrm{A}_s)$$

où $c(s)$ est un réel entre x_0 et $x_0 + s$ et $\mathrm{A}_s = (c(s), y_0)$ un point de $\mathbb{R}^2$, situé sur le segment horizontal entre (x_0, y_0) et $(x_0 + s, y_0)$. De même en utilisant $g_2(y) = f(x_0 + s, y)$ entre y_0 et $y_1 = y_0 + t$ on trouve que

$$f(x_0 + s, y_0 + t) - f(x_0 + s, y_0) = t\, g_2'(d(s, t)) = t\, \mathrm{D}_2 f(\mathrm{B}_{s,t})$$

où $\mathrm{B}_{s,t} = (x_0 + s, d(s, t))$ est un point du segment vertical entre $(x_0 + s, y_0)$ et $(x_0 + s, y_0 + t)$. Alors

$$f(a + h) - f(a) = s\, \mathrm{D}_1 f(\mathrm{A}_s) + t\, \mathrm{D}_2 f(\mathrm{B}_{s,t})$$

et

$$\mathrm{E}(h) = f(a + h) - f(a) - \mathrm{D}_1 f(a)\, s - \mathrm{D}_2 f(a)\, t$$

est égal à

$$\big(\mathrm{D}_1 f(\mathrm{A}_s) - \mathrm{D}_1 f(a)\big)\, s + \big(\mathrm{D}_2 f(\mathrm{B}_{s,t}) - \mathrm{D}_2 f(a)\big)\, t$$

qu'on majore par

$$|\mathrm{E}(h)| \le \Big(\big|\mathrm{D}_1 f(\mathrm{A}_s) - \mathrm{D}_1 f(a)\big| + \big|\mathrm{D}_2 f(\mathrm{B}_{s,t}) - \mathrm{D}_2 f(a)\big|\Big)\, \max(|s|, |t|);$$

ceci est bien une majoration de l'erreur qui est de la forme $\|h\|_\infty\, \varepsilon(h)$, avec

$$\varepsilon(h) = \big|\mathrm{D}_1 f(\mathrm{A}_s) - \mathrm{D}_1 f(a)\big| + \big|\mathrm{D}_2(\mathrm{B}_{s,t}) - \mathrm{D}_2 f(a)\big|$$

qui tend vers 0 quand $h \to 0_\mathrm{E}$, parce que les dérivées partielles sont continues au point a, et que $\|\mathrm{A}_s - a\|_\infty$ et $\|\mathrm{B}_{s,t} - a\|_\infty$ sont majorés par $\|h\|_\infty$, donc tendent vers 0 avec h.