MT242, Cours n° 13, Lundi 20 Mars 2000.

Exercice traité. Exprimer la matrice jacobienne de $\varphi : v \to \|f(v) - w_0\|_2^2$, où f est une application de classe C^1 d'un ouvert $\mathrm{U} \subset \mathbb{R}^n$ dans $\mathbb{R}^n$ et w_0 un point fixé de $\mathbb{R}^n$. On trouve

$$\forall v \in \mathrm{U}, \quad (\mathrm{J}\varphi)_v = 2\,{}^t(\mathrm{F}(v) - \mathrm{W}_0)\,(\mathrm{J}f)_v$$

où W_0 est la matrice colonne des coordonnées de w_0 et $\mathrm{F}(v)$ la matrice colonne des coordonnées du vecteur $f(v) \in \mathbb{R}^n$.

Rappel. Théorème des accroissements finis. Si on a $\|(df)_c\|_{\mathcal{L}(\mathrm{E},\mathrm{F})} \le \mathrm{M}$ en tout point c du segment $[a, b]$, alors $\|f(b) - f(a)\|_\mathrm{F} \le \mathrm{M}\,\|b - a\|_\mathrm{E}$.

La norme d'application linéaire est en général difficile à calculer exactement. On a quand même facilement $\|\mathrm{Id}\| = 1$. Si $\mathrm{E} = (\mathbb{R}^n, \|\,.\,\|_2)$ et si ℓ est un endomorphisme de E, de matrice L dans la base canonique, on a la majoration

$$(*) \qquad \|\ell\|_{\mathcal{L}(\mathrm{E})} \le \|\mathrm{L}\|_2.$$

3.3. Inversion locale et fonctions implicites

Lemme 3.3.1. *Si $\ell \in \mathcal{L}(\mathbb{R}^n)$ est inversible, il existe une constante $c > 0$ telle que $\|\ell(v)\|_2 \ge c\,\|v\|_2$ pour tout vecteur $v \in \mathbb{R}^n$.*

Démonstration. Puisque ℓ^{-1} existe, il existe une constante $\mathrm{C} > 0$ (la norme de ℓ^{-1} si on veut) telle que

$$\forall w \in \mathbb{R}^n, \quad \|\ell^{-1}(w)\|_2 \le \mathrm{C}\,\|w\|_2.$$

Puisque ℓ est bijective, tout vecteur $v \in \mathbb{R}^n$ peut s'écrire $v = \ell^{-1}(w)$ pour un $w \in \mathbb{R}^n$ unique, et la relation précédente s'écrit $\|\ell(v)\|_2 \ge \mathrm{C}^{-1}\,\|v\|_2$, ce qui donne le résultat voulu avec $c = \mathrm{C}^{-1}$.

Un exemple d'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

$$f(x, y) = (\sin(x) + y^2, \sin(y) + x^2) \in \mathbb{R}^2$$

Calcul de la matrice jacobienne au point $(0, 0)$: c'est la matrice identité, qui est inversible. Peut-on dire que f va suivre de si près son approximation

$$v \to f(0, 0) + (df)_{(0,0)}(v) = (df)_{(0,0)}(v)$$

qui est une application linéaire inversible, que f elle même sera injective ? Ca n'est pas vrai si on ne reste pas assez près du point $(0, 0)$: les points (π, π) et $(-\pi, -\pi)$ ont la même image par f. Le résultat théorique qui suit est intéressant car il serait en général très difficile de vérifier "à la main" que f est injective sur une petite boule autour du point $(0, 0)$.

Proposition 3.3.1. *Soient U un ouvert de $\mathbb{R}^n$, $a \in \mathrm{U}$ et $f : \mathrm{U} \to \mathbb{R}^n$ une fonction de classe C^1 sur U ; si la matrice jacobienne $(\mathrm{J}f)_a$ est inversible (ce qui équivaut à dire que l'application linéaire $(df)_a$ est inversible), il existe $r > 0$ et $\delta > 0$ tels que*

$$\|f(v) - f(v')\| \ge \delta\,\|v - v'\|$$

pour tous vecteurs $v, v' \in \mathrm{B}(a, r)$. En particulier f est injective sur $\mathrm{B}(a, r)$. De plus, $(\mathrm{J}f)_v$ est inversible pour tout $v \in \mathrm{B}(a, r)$.

Démonstration. Puisque $(df)_a$ est inversible, il existe $\gamma > 0$ tel que $(df)_a(h) \geq \gamma \|h\|$ pour tout $h \in \mathbb{R}^n$. On va choisir $r > 0$ assez petit pour que

$$|\mathrm{D}_i f(v) - \mathrm{D}_i f(a)| < \frac{\gamma}{4n}$$

pour tout $i = 1, \ldots, n$ et tout $v \in \mathrm{B}(a, r)$. Il en résulte que $\|(\mathrm{J}f)_v - (\mathrm{J}f)_a\|_2 < \gamma/4$ ce qui implique

$$(**) \qquad \|(df)_v - (df)_a\|_{\mathcal{L}(\mathrm{E})} < \gamma/4$$

d'après $(*)$, donc $\|(df)_v(h) - (df)_a(h)\|_2 < (\gamma/4)\,\|h\|_2$, donc $\|(df)_v(h)\|_2 \geq (3\gamma/4)\,\|h\|_2$ par l'inégalité triangulaire, pour tout $v \in \mathrm{B}(a, r)$. On fixe v_1, v_2 dans cette boule $\mathrm{B}(a, r)$; le segment $[v_1, v_2]$ est contenu dans la boule. On applique ensuite le théorème des accroissements finis sur le segment $[v_1, v_2]$ à l'application $g : \mathrm{U} \to \mathbb{R}^n$ définie par

$$g(v) = f(v) - f(v_1) - (df)_{v_1}(v - v_1).$$

La différentielle $(dg)_v$ est égale à $(df)_v - (df)_{v_1}$ qui est majorée en norme par $\gamma/2$ pour tout point v du segment d'après $(**)$. Alors les relations

$$\|f(v_2) - f(v_1) - (df)_{v_1}(v_2 - v_1)\| = \|g(v_2) - g(v_1)\| \leq \frac{\gamma}{2}\,\|v_2 - v_1\|$$

et $\|(df)_{v_1}(v_2 - v_1)\| \geq (3\gamma/4)\,\|v_2 - v_1\|$ donnent le résultat $\|f(v_2) - f(v_1)\| \geq \delta\,\|v_2 - v_1\|$ par l'inégalité triangulaire, avec $\delta = \gamma/4$.

Théorème 3.3.1. Théorème d'inversion locale. *Soient* U *un ouvert de* $\mathbb{R}^n$, $a \in \mathrm{U}$ *et* $g : \mathrm{U} \to \mathbb{R}^n$ *une fonction de classe* C^1 *sur* U *; si la matrice jacobienne* $(\mathrm{J}g)_a$ *est inversible, il existe un ouvert* U_0 *contenu dans* U *et contenant le point* a *et un ouvert* U'_0 *contenant* $g(a)$ *tels que la restriction de* g *à* U_0 *soit une bijection de* U_0 *sur* U'_0, *et tels que de plus, si* $\widehat{g}$ *désigne la restriction de* g *à* U_0, *l'application réciproque* $\widehat{g}^{-1} : \mathrm{U}'_0 \to \mathrm{U}_0$ *soit de classe* C^1 *sur* U'_0.

Esquisse de démonstration. On peut déjà trouver $r > 0$ et $\delta > 0$ tels que g soit injective sur $\mathrm{B}(a, r)$, et $(\mathrm{J}g)_v$ inversible pour tout $v \in \mathrm{B}(a, r)$. Posons $b = f(a)$. On va montrer que tout point $w_0 \in \mathrm{B}(b, \delta r/2)$ est dans l'image de $\mathrm{B}(a, r)$. On posera $r' = \delta r/2$ et

$$\mathrm{U}_0 = \{v \in \mathrm{B}(a, r) : g(v) \in \mathrm{B}(b, r')\}$$

et on aura bien que g est une bijection de U_0 sur son image $\mathrm{U}'_0 = \mathrm{B}(b, r')$. Une façon de montrer que tout $w_0 \in \mathrm{B}(b, r')$ est dans l'image de g est de minimiser la fonction $\varphi(v) = \|g(v) - w_0\|^2$ sur le compact $\overline{\mathrm{B}(a, r)}$ (la boule fermée de centre a et de rayon r). Le minimum existe par compacité. On va montrer que le minimum ne peut pas être atteint sur la sphère $\mathrm{S}(r)$ de centre a et de rayon r. En effet, pour tout point $v' \in \mathrm{S}(r)$ on a $\|g(v') - g(a)\| \geq \delta r = 2r'$, alors que $\|g(a) - w_0\| < r'$, donc $\|g(v') - w_0\| \geq r'$. Ceci montre que le point $v = a$ donne un meilleur résultat que n'importe quel $v' \in \mathrm{S}(r)$ pour la minimisation de φ. Le minimum de φ sur le compact $\overline{\mathrm{B}(a, r)}$ est donc atteint en un point $v_0 \in \mathrm{B}(a, r)$ (la boule ouverte). Il en résulte que le jacobien de φ s'annule au point v_0. On a calculé ce jacobien dans le premier exercice de cette séance. On a donc

$$(\mathrm{J}\varphi)_{v_0} = 2\,{}^t(\mathrm{G}(v_0) - \mathrm{W}_0)\,(\mathrm{J}g)_{v_0} = 0.$$

Mais puisque $(\mathrm{J}g)_{v_0}$ est inversible, on a $\mathrm{G}(v_0) - \mathrm{W}_0 = 0$, c'est à dire $g(v_0) = w_0$. On a ainsi montré que tout point $w_0 \in \mathrm{B}(b, r')$ est dans l'image de g.

Cours n° 14, Mercredi 22 Mars 2000.

Soit maintenant U un ouvert de $\mathbb{R}^3$, $a = (x_0, y_0, z_0) \in \mathrm{U}$ et $f : \mathrm{U} \to \mathbb{R}$ de classe C^1 telle que $f(a) = 0$. On s'intéresse à l'ensemble

(1) $$\mathrm{S} = \{v = (x, y, z) \in \mathrm{U} : f(v) = 0\}.$$

Dans les bons cas, cette équation représente une surface ou un morceau de surface. On dit que S est définie par une *équation implicite*.

Exemple. L'équation

(2) $$x^2 + y^2 + z^2 - 1 = 0$$

définit la sphère unité de $\mathbb{R}^3$.

On aimerait expliciter l'équation (1), en trouvant si possible une représentation de la forme $z = \varphi(x, y)$ pour les points de S. Pour l'équation (2) par exemple, si $a = (x_0, y_0, z_0)$ est sur la sphère unité et si $z_0 < 0$, on peut représenter les points de la sphère voisins de a par l'équation “explicite”

$$z = -\sqrt{1 - x^2 - y^2} = \varphi(x, y)$$

où φ est définie dans le voisinage V de $\widetilde{a} = (x_0, y_0)$ défini par $\mathrm{V} = \{\widetilde{v} = (x, y) : x^2 + y^2 < 1\}$. Mais cette représentation ne nous donnera que la moitié de la sphère, celle qui correspond à $z < 0$. De plus, si $z_0 = 0$, par exemple $a = (1, 0, 0)$ cette représentation peut donner le point a lui-même en prenant $\widetilde{v} = (1, 0)$, mais elle ne décrira pas les points de la sphère voisins de $(1, 0, 0)$ qui sont par exemple de la forme $(\cos\theta, 0, \sin\theta)$ avec $\theta > 0$ petit. Pour le point $(1, 0, 0)$, on cherchera plutôt à représenter un morceau de sphère contenant $(1, 0, 0)$ avec l'équation

$$x = \sqrt{1 - y^2 - z^2}.$$

Pour tout $v = (x, y, z) \in \mathbb{R}^3$, on notera $\widetilde{v} = (x, y) \in \mathbb{R}^2$.

Théorème 3.3.2. Théorème des fonctions implicites. *Soient* U *un ouvert de* $\mathbb{R}^3$, $a \in \mathrm{U}$ *et* $f : \mathrm{U} \to \mathbb{R}$ *une fonction de classe* C^1 *sur* U *telle que* $f(a) = 0$.

Si $\mathrm{D}_3 f(a) = \frac{\partial f}{\partial z}(a) \neq 0$, *il existe* $\varepsilon > 0$, $\delta > 0$ *et une fonction* $\varphi : \mathrm{B}(\widetilde{a}, \delta) \to]z_0 - \varepsilon, z_0 + \varepsilon[$ *tels que*

1. pour tout $\widetilde{v} \in \mathrm{B}(\widetilde{a}, \delta)$, *le point* $(\widetilde{v}, \varphi(\widetilde{v}))$ *est dans* S *(et il est en plus dans l'ouvert* $\mathrm{U}_1 = \mathrm{B}(\widetilde{a}, \delta)\times]z_0 - \varepsilon, z_0 + \varepsilon[$*)*

2. pour tout $v = (\widetilde{v}, z)$ *dans* U_1, *si* $v \in \mathrm{S}$ *alors* $z = \varphi(\widetilde{v})$.

Esquisse de preuve. On applique le théorème d'inversion locale à l'application $g : \mathrm{U} \to \mathbb{R}^3$ définie par $g(x, y, z) = (x, y, f(x, y, z))$. En écrivant la matrice jacobienne de g au point a, on voit facilement que son déterminant vaut $\mathrm{D}_3 f(a) \neq 0$, donc $(\mathrm{J}g)_a$ est inversible. On peut alors trouver U_0 sur lequel g est bijective, et on peut toujours supposer que

$$\mathrm{U}_0 = \mathrm{B}(\widetilde{a}, \varepsilon)\times]z_0 - \varepsilon, z_0 + \varepsilon[$$

pour un certain $\varepsilon > 0$. L'image U'_0 de U_0 par g est un voisinage de $g(a) = (\widetilde{a}, 0)$. Il existe donc $\delta > 0$ tel que $\mathrm{B}(f(a), \delta) \subset \mathrm{U}'_0$. En particulier, on a $(\widetilde{v}, 0) \in \mathrm{U}'_0$ pour tout $\widetilde{v} \in \mathrm{B}(\widetilde{a}, \delta)$. Pour un tel point, $w = (\widehat{g})^{-1}(\widetilde{v}, 0)$ est défini, et w est nécessairement de la forme $w = (\widetilde{v}, z)$ d'après la forme de g ; de plus $f(\widetilde{v}, z)$ est la troisième coordonnée de $g(w) = (\widetilde{v}, 0)$, donc $f(\widetilde{v}, z) = 0$. Puisque $\widehat{g}$ est bijective, z est uniquement défini en fonction de $\widetilde{v}$, et on peut poser $z = \varphi(\widetilde{v})$: pour tout $\widetilde{v} \in \mathrm{B}(\widetilde{a}, \delta)$, $\varphi(\widetilde{v})$ est la troisième coordonnée de $(\widehat{g})^{-1}(\widetilde{v}, 0)$.

3.4. Dérivées partielles secondes. Extrema locaux

Soit $f : \mathrm{U} \to \mathbb{R}$ une fonction de classe C^1, où U est un ouvert de $\mathbb{R}^n$; pour chaque $j = 1, \ldots, n$ il existe une *fonction* réelle définie sur U par $v \in \mathrm{U} \to (\mathrm{D}_j f)(v)$. On peut se demander si cette fonction $\mathrm{D}_j f$ admet elle-même des dérivées partielles.

Si la fonction $\mathrm{D}_j f$ admet une $i$ème dérivée partielle au point $a \in \mathrm{U}$, on appelle cette dérivée partielle la *dérivée partielle seconde* $\mathrm{D}_i(\mathrm{D}_j f)(a)$ de f au point a, qu'on peut noter encore

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)(a).$$

On dit que f est de classe C^2 dans U si f est de classe C^1 dans U et si toutes les fonctions dérivées partielles $\mathrm{D}_j f$, $j = 1, \ldots, n$ sont de classe C^1 dans U. Autrement dit, f est de classe C^2 si toutes les dérivées partielles secondes $\mathrm{D}_i(\mathrm{D}_j f)(v)$ existent pour tout $v \in \mathrm{U}$, et si toutes ces fonctions $v \to \mathrm{D}_i(\mathrm{D}_j f)(v)$ sont continues dans U.

Théorème 3.4.1. Lemme de Schwarz. *Soient* U *un ouvert de* $\mathbb{R}^n$, $f : \mathrm{U} \to \mathbb{R}$ *de classe* C^2 *dans* U. *Pour tous* $i, j = 1, \ldots, n$ *et tout point* $a \in \mathrm{U}$, *on a*

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)(a) = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big)(a).$$

Démonstration. On pose pour t assez petit

$$\Phi(t) = f(x_0 + t, y_0 + t) - f(x_0 + t, y_0) - f(x_0, y_0 + t) + f(x_0, y_0).$$

On va montrer que

$$\lim_{t \to 0} \frac{\Phi(t)}{t^2} = \frac{\partial}{\partial y}\Big(\frac{\partial f}{\partial x}\Big)(a).$$

Comme l'expression $\Phi(t)$ est symétrique en x et y, on pourra montrer de la même manière que $\Phi(t)/t^2$ tend vers $\frac{\partial}{\partial x}\big(\frac{\partial f}{\partial y}\big)(a)$, et le résultat sera établi.

Considérons pour t fixé

$$g_1(s) = f(x_0 + s, y_0 + t) - f(x_0 + s, y_0).$$

Alors

$$g_1'(s) = \frac{\partial f}{\partial x}(x_0 + s, y_0 + t) - \frac{\partial f}{\partial x}(x_0 + s, y_0).$$

et par le théorème des accroissements finis en dimension un

$$\Phi(t) = g_1(t) - g_1(0) = t\, g_1'(c(t)) = t\Big(\frac{\partial f}{\partial x}(x_0 + c(t), y_0 + t) - \frac{\partial f}{\partial x}(x_0 + c(t), y_0)\Big)$$

où $c(t)$ est un réel entre 0 et t. Posons maintenant

$$g_2(s) = \frac{\partial f}{\partial x}(x_0 + c(t), y_0 + s).$$

On a

$$\Phi(t) = t\,(g_2(t) - g_2(0)) = t^2\, g_2'(d(t))$$

où $d(t)$ est un réel entre 0 et t, ce qui donne finalement

$$\frac{\Phi(t)}{t^2} = \frac{\partial}{\partial y}\Big(\frac{\partial f}{\partial x}\Big)(x_0 + c(t), y_0 + d(t))$$

qui tend vers $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0, y_0)$ quand $t \to 0$, par la continuité des dérivées partielles secondes, et parce que $c(t) \to 0$ et $d(t) \to 0$.

Après le lemme de Schwarz, on introduit la notation définitive pour les dérivées partielles secondes. On posera

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \mathrm{D}_i(\mathrm{D}_j f)(a) \quad \text{si} \quad i \neq j, \qquad \frac{\partial^2 f}{\partial x_i^2}(a) = \mathrm{D}_i(\mathrm{D}_i f)(a).$$

Considérons une fonction réelle $f(x_1, x_2)$ de classe C^2 dans un ouvert U contenant un point $a = (a_1, a_2) \in \mathbb{R}^2$, et considérons un vecteur $h = (h_1, h_2)$. Etudions la fonction f en des points $a + th$ de la droite passant par a et de vecteur directeur h, avec t réel assez petit pour que $a + th \in \mathrm{V}$. Posons donc

$$\varphi(t) = f(a_1 + th_1, a_2 + th_2).$$

On sait que

$$\varphi'(t) = \frac{\partial f}{\partial x_1}(a_1 + th_1, a_2 + th_2)\, h_1 + \frac{\partial f}{\partial x_2}(a_1 + th_1, a_2 + th_2)\, h_2,$$

ce qui nous donnera en dérivant une fois de plus, en posant $a_t = a + th$

$$\varphi''(t_0) = \frac{\partial^2 f}{\partial x_1^2}(a_t)\, h_1^2 + \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_t)\, h_1 h_2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}(a_t)\, h_1 h_2 + \frac{\partial^2 f}{\partial x_2^2}(a_t)\, h_2^2 =$$

$$\frac{\partial^2 f}{\partial x_1^2}(a_t)\, h_1^2 + 2\frac{\partial^2 f}{\partial x_2 \partial x_1}(a_t)\, h_1 h_2 + \frac{\partial^2 f}{\partial x_2^2}(a_t)\, h_2^2.$$

On retiendra cette formule utile,

$$\frac{d^2}{dt^2} f(a + th)_{\big|t=t_0} = (\mathrm{hess}\, f)_{a+t_0 h}(h)$$

en désignant par $(\mathrm{hess}\, f)_v$ la forme quadratique dont la matrice dans la base canonique est la matrice des dérivées partielles secondes de f au point $v \in \mathrm{U}$. On notera $(\mathrm{Hess}\, f)_v$ cette matrice. On l'appelle la *matrice hessienne* de f au point v (d'après le nom de Ludwig Otto Hesse, mathématicien allemand, 1811–1874).

Théorème 3.4.2. Taylor-Lagrange. *Si f est de classe C^2 dans un ouvert U contenant le segment $[a, a + h]$, il existe un point c de $[a, a + h]$ tel que*

$$f(a + h) = f(a) + (df)_a(h) + \frac{1}{2}\, (\mathrm{hess}\, f)_c(h).$$

Si on exprime les choses matriciellement, en introduisant la matrice colonne H des coordonnées de h, et en identifiant les matrices 1×1 à des réels, on aura

$$f(a + h) = f(a) + (\mathrm{J}f)_a\, \mathrm{H} + \frac{1}{2}\, {}^t\mathrm{H}\, (\mathrm{Hess}\, f)_c\, \mathrm{H}.$$

Proposition 3.4.1. Taylor-Young. *Soient U un ouvert de $\mathbb{R}^n$, $a \in \mathrm{U}$ et $f : \mathrm{U} \to \mathbb{R}$ une fonction de classe C^2 dans U ; on peut écrire*

$$f(a + h) = f(a) + (df)_a(h) + \frac{1}{2}\, (\mathrm{hess}\, f)_a(h) + \|h\|^2 \varepsilon(h).$$

Démonstration. D'après Taylor-Lagrange, il existe un point c_h sur le segment $[a, a+h]$ tel que

$$f(a+h) = f(a) + (\mathrm{J}f)_a\,\mathrm{H} + \frac{1}{2}\,{}^t\mathrm{H}\,(\mathrm{Hess}\,f)_{c_h}\,\mathrm{H}$$

ce qui permet d'exprimer l'erreur

$$\mathrm{E}(h) = f(a+h) - f(a) - (\mathrm{J}f)_a\,\mathrm{H} - \frac{1}{2}\,{}^t\mathrm{H}\,(\mathrm{Hess}\,f)_a\mathrm{H}$$

dans l'approximation de Taylor-Young par

$$\mathrm{E}(h) = \frac{1}{2}\,{}^t\mathrm{H}\left((\mathrm{Hess}\,f)_{c_h} - (\mathrm{Hess}\,f)_a\right)\mathrm{H}.$$

Si on introduit la matrice $\mathrm{M}_h = (\mathrm{Hess}\,f)_{c_h} - (\mathrm{Hess}\,f)_a$, on voit que cette matrice M_h tend vers 0 quand $h \to 0_{\mathbb{R}^n}$, parce que les dérivées partielles secondes sont continues et que $c_h \to a$. D'autre part, on utilisant Cauchy-Schwarz et la définition de la norme de matrice,

$$2|\mathrm{E}(h)| = \left|{}^t\mathrm{H}\,\mathrm{M}_h\,\mathrm{H}\right| \leq \|\mathrm{H}\|_2\,\|\mathrm{M}_h\,\mathrm{H}\|_2 \leq \|\mathrm{M}_h\|\,\|\mathrm{H}\|_2^2$$

donc $|\mathrm{E}(h)| \leq \|h\|^2\,\varepsilon(h)$ avec $\varepsilon(h) = \frac{1}{2}\,\|\mathrm{M}_h\|$ qui tend vers 0 lorsque $h \to 0_{\mathbb{R}^n}$.

3.4.1. Conditions du second ordre pour un minimum local

On dit que $f : \mathrm{U} \to \mathbb{R}$ admet un maximum local au point $a \in \mathrm{U}$ s'il existe un voisinage V de a, $\mathrm{V} \subset \mathrm{U}$ tel que

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

Lemme 3.4.1. *Si* Q *est une forme quadratique définie positive sur* $\mathbb{R}^n$*, il existe une constante* $\delta > 0$ *telle que* $\mathrm{Q}(v) \geq \delta\,\|v\|^2$ *pour tout vecteur* $v \in \mathbb{R}^n$.

Démonstration omise.

Proposition 3.4.2. *Soient* U *un ouvert de* $\mathbb{R}^n$, $a \in \mathrm{U}$ *et* $f : \mathrm{U} \to \mathbb{R}$ *de classe* C^2 *dans* U *; On suppose que* a *est un point critique de* f*, c'est à dire que* $(df)_a = 0$.

1. Si la forme quadratique $(\mathrm{hess}\,f)_a$ *est définie positive (c'est à dire qu'elle est de signature* $(n, 0)$*) la fonction* f *admet au point* a *un minimum local.*

1 bis. Si la forme quadratique $(\mathrm{hess}\,f)_a$ *est définie négative (c'est à dire qu'elle est de signature* $(0, n)$*) la fonction* f *admet au point* a *un maximum local.*

2. Si $(\mathrm{hess}\,f)_a$ *est de signature* (s, t) *avec* $s\,t \neq 0$*, alors* $v \to f(v) - f(a)$ *change de signe au voisinage de* a*, donc* f *n'admet pas d'extremum local au point* a.

3. Si $(\mathrm{hess}\,f)_a$ *est de signature* $(s, 0)$ *avec* $s < n$ *ou bien* $(0, t)$ *avec* $t < n$*, on NE peut PAS conclure.*

Démonstration. Si $(\mathrm{hess}\,f)_a$ est définie positive, il existe un nombre $\delta > 0$ tel que $(\mathrm{hess}\,f)_a(h) \geq \delta\|h\|^2$ pour tout h. Ce qui donne avec le DL de Taylor-Young, compte tenu de $(df)_a = 0$

$$f(a+h) - f(a) \geq \|h\|^2(\delta + \varepsilon(h)),$$

quantité > 0 pour $h \neq 0$ assez petit. Si $(\mathrm{hess}\,f)_a$ est de signature (s, t) avec $s\,t \neq 0$, on peut trouver un vecteur v tel que $(\mathrm{hess}\,f)_a(v) = 1$ et un autre vecteur w tel que $(\mathrm{hess}\,f)_a(w) = -1$. Alors pour t petit non nul

$$f(a+tv) - f(a) = \frac{1}{2}\,(\mathrm{hess}\,f)_a(tv) + \|tv\|^2\varepsilon(tv) = \frac{1}{2}\,t^2 + t^2\|v\|^2\varepsilon(tv)$$

est > 0 alors que

$$f(a+tw) - f(a) = -\frac{1}{2}\,t^2 + t^2\|w\|^2\varepsilon(tw)$$

est < 0.

Un exemple du cas 3. Considérons

$$f_1(x,y) = x^4 + y^4 - 2(x-y)^2, \quad f_2(x,y) = -x^4 - y^4 - 2(x-y)^2.$$

Le point $(0,0)$ est critique pour les deux fonctions, elles ont la même matrice hessienne au point $(0,0)$,

$$(\operatorname{Hess} f)_{(0,0)} = \begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}.$$

La signature est $(0,1)$; la seconde fonction a visiblement un maximum local au point $(0,0)$ alors que la première n'a pas d'extremum en ce même point. Pourtant les informations sur les dérivées premières et secondes au point $(0,0)$ sont identiques pour les deux fonctions.