SÉANCE 6-7 : PROPRÉTÉS DES FONCTIONS DIFFÉRENTIABLES

6.1. Plus de résultats sur la différentiabilité

Soient E un espace vectoriel de rang fini sur $\mathbb{R}$. On fix une base $(e_i)_{i=1}^d$ de E. Pour tout $i \in \{1,\ldots,d\}$, on désigne par $e_i^\vee$ l'application linéaire de E vers $\mathbb{R}$ qui envoie $\lambda_1 e_1 + \cdots + \lambda_d e_d$ en λ_i.

Théorème 6.1. — *Soit f une application définie sur un ouvert non-vide U de E et à valeurs dans un espace vectoriel normé. On suppose que, pour tout $i \in \{1,\ldots,d\}$, la dérivée partielle $\partial_{e_i} f$ existe et définit une application continue de U vers F. Alors l'application f est différentiable sur U et on a*

$$Df(x) = \sum_{i=1}^d \partial_{e_i} f(x) e_i^\vee$$

quel que soit $x \in U$.

Démonstration. — Sans perte de généralité, on suppose que E est muni de la norme $\|.\|$ telle que $\|\lambda_1 e_1 + \cdots + \lambda_d e_d\| := \max\{|\lambda_1|,\ldots,|\lambda_d|\}$. Soit h un vecteur dans E de la forme $\lambda_1 e_1 + \cdots + \lambda_d e_d$. Soit a un élément de U. On a

$$f(a+h) - f(a) = \sum_{i=1}^d f(a + \lambda_1 e_1 + \cdots + \lambda_i e_i) - f(a + \lambda_1 e_1 + \cdots + \lambda_{i-1} e_{i-1}).$$

Par conséquent, on a

$$\begin{aligned} &\Big\| f(a+h) - f(a) - \sum_{i=1}^d \lambda_i \partial_{e_i} f(a) \Big\| \\ &\leqslant \sum_{i=1}^d \Big\| f\Big(a + \sum_{j \leqslant i} \lambda_j e_j\Big) - f\Big(a + \sum_{j<i} \lambda_j e_j\Big) - \lambda_i \partial_{e_i} f(a) \Big\|. \end{aligned}$$

Comme $\partial_{e_i} f$ est une application continue, on obtient

$$\begin{aligned}
&\Big\| f\Big(a + \sum_{j \leqslant i} \lambda_j e_j\Big) - f\Big(a + \sum_{j<i} \lambda_j e_j\Big) - \lambda_i \partial_{e_i} f(a)\Big\| \\
&\leqslant \Big\| f\Big(a + \sum_{j \leqslant i} \lambda_j e_j\Big) - f\Big(a + \sum_{j<i} \lambda_j e_j\Big) - \lambda_i \partial_{e_i} f\Big(a + \sum_{j<i} \lambda_i e_j\Big)\Big\| \\
&\qquad + |\lambda_i| \cdot \Big\| \partial_{e_i} f\Big(a + \sum_{j<i} \lambda_i e_j\Big) - \partial_{e_i} f(a)\Big\| \\
&= o(\|h\|)
\end{aligned}$$

puisque $\partial_{e_i} f$ est continue, d'où le résultat. □

Exemple 6.2. — Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie comme $f(x, y) = x^2 + y^2$. On a $\frac{\partial f}{\partial x} = 2x$ et $\frac{\partial f}{\partial y} = 2y$. Ce sont des fonctions continues sur $\mathbb{R}^2$. On en déduit donc que la fonction f est de classe C^1 sur $\mathbb{R}^2$ et que la différentielle de f est de la forme

$$Df(x, y) = 2x\mathrm{d}x + 2y\mathrm{d}y,$$

où $\mathrm{d}x$ et $\mathrm{d}y$ désignent les applications linéaires de $\mathbb{R}^2$ vers $\mathbb{R}$ définies par les projection en les deux coordonnées respectivement.

Proposition 6.3. — *Soient F un espace vectoriel normé et $(f_i)_{i=1}^d$ une famille de fonctions continues (resp. différentiables, dérivable dans une direction $h \in F$) définies sur un ouvert non-vide V de F, alors la fonction $f : V \to E$ définie comme*

$$(x \in V) \longmapsto f_1(x)e_1 + \cdots + f_d(x)e_d$$

est continue (resp. différentiable, dérivable dans la direction h).

Démonstration. — Sans perte de généralité, on peut supposer que la norme sur E satisfait à la relation $\|\lambda_1 e_1 + \cdots + \lambda_d e_d\| = \max(|\lambda_1|, \ldots |\lambda_d|)$. On a alors

$$\|f(x + h) - f(x)\| = \max_{1 \leqslant i \leqslant d} |f_i(x + h) - f_i(x)|.$$

Donc f est continue dès que les f_i sont continues. De même, si les f_i sont différentiables, alors on a

$$\Big\| f(x + h) - f(x) - \sum_{i=1}^{d} Df_i(x)(h)e_i \Big\| = \max_{1 \leqslant i \leqslant d} |f_i(x + h) - f_i(x) - Df_i(x)(h)|.$$

Donc f est différentiable en x et on a $Df(x) = \sum_{i=1}^d Df_i(x)e_i$. La démonstration du troisième énoncé est similaire, on la laisse comme un exercice. □

Corollaire 6.4. — *Soient F un espace vectoriel normé et $(f_i)_{i=1}^d$ une famille de fonctions définies sur un ouvert non-vide V de F. Soit $f : V \to E$ la fonction définie comme $(x \in V) \longmapsto f_1(x)e_1 + \cdots + f_d(x)e_d$. Soit g une application définie sur un voisinage de $f(x)$ est à valeur dans un autre espace vectoriel normé G. Soit a un* *élément de V. Si l'application f est continue en a et dérivable en a dans la direction de $h \in F$ est si la fonction g est différentiable en $f(a)$, alors la fonction composée $g \circ f$ est dérivable en a dans la direction de h. En outre, on a*

$$\partial_h(g \circ f)(a) = \sum_{i=1}^{d} \partial_h f_i(a) \partial_{e_i} g(f(a)).$$

Démonstration. — Comme la fonction g est différentiable en $f(a)$, elle est continue en $f(a)$. Donc la fonction composée $g \circ f$ est bien définie dans un voisinage de a. En outre, pour $t \in \mathbb{R}$ dont la valeur absolue est suffisamment petite, on a

$$\begin{aligned}(g \circ f)(a + th) - g(f(a)) &= g(f(a) + (f(a + th) - f(a))) - g(f(a)) \\ &= Dg(f(a))(f(a + th) - f(a)) + o(\|f(a + th) - f(a)\|)\end{aligned}$$

Comme f est dérivable en a dans la direction de h, on a

$$f(a + th) - f(a) = t\partial_h f(a) + o(|t|) = \sum_{i=1}^{d} t\partial_h f_i(a) e_i + o(|t|).$$

Par conséquent, $(g \circ f)(a + th) - g(f(a)) = t \sum_{i=1}^{d} \partial_h f_i(a) \partial_{e_i} g(f(a)) + o(|t|)$. $\square$

Remarque 6.5. — Soit E un espace vectoriel de rang fini sur $\mathbb{R}$ et $(e_i)_{i=1}^d$ une base de E. Pour tout $i \in \{1, \ldots, d\}$, soit $e_i^\vee : E \to \mathbb{R}$ la projection sur la i-ème coordonnée: elle envoie $\lambda_1 e_1 + \cdots + \lambda_d e_d$ vers λ_i. C'est une application linéaire entre des espaces vectoriels de rang fini sur $\mathbb{R}$, donc est continue. On en déduit qu'elle est différentiable. Si f est une application définie sur un ouvert d'un espace vectoriel normé F vers E qui est différentiable, alors il en est de même de chaque $f_i = \mathrm{pr}_i \circ f$. Sachant que la fonction f s'écrit sous la forme $f = f_1 e_1 + \cdots + f_d e_d$, on obtient que la fonction f est différentiable si et seulement si chaque projection f_i ($i \in \{1, \ldots, d\}$) l'est, compte tenu de la proposition 6.3.

6.2. Matrice jacobienne

Soit $d \geqslant 1$ un entier. On désigne par $(e_i)_{i=1}^d$ la base canonique de $\mathbb{R}^d$. Considérons une application f définie sur un ouvert non-vide U de $\mathbb{R}^d$ et à valeurs dans $\mathbb{R}^d$. L'application f s'écrit sous la forme $f_1 e_1 + \cdots + f_d e_d$, où chaque f_i est une fonction sur U. Si f est une application différentiable en $a \in U$ (ou de façon équivalente, chacune des fonctions f_i est différentiable en a), on appelle matrice jacobienne la matrice carrée de la forme

$$J_f(a) := \begin{pmatrix} \partial_{e_1} f_1(a) & \partial_{e_2} f_1(a) & \cdots & \partial_{e_d} f_1(a) \\ \partial_{e_1} f_2(a) & \partial_{e_2} f_2(a) & \cdots & \partial_{e_d} f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{e_1} f_d(a) & \partial_{e_2} f_d(a) & \cdots & \partial_{e_d} f_d(a) \end{pmatrix}$$

On observe aussitôt que $Df(a)$ est inversible si et seulement si le déterminant de la matrice jacobienne est non-nul.

Exemple 6.6. — Considérons l'application $f :]0,+\infty[\times\mathbb{R} \to \mathbb{R}^2$ qui envoie (r,θ) en $(r\cos(\theta), r\sin(\theta))$. On a

$$J_f(r,\theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

Le déterminant de $J_f(r,\theta)$ est r, qui est strictement positif.

Exemple 6.7. — Considérons l'application $g :]0,+\infty[\times\mathbb{R}^2 \to \mathbb{R}^3$ qui envoie (r,θ,ψ) en $(r\cos(\theta)\cos(\psi), r\cos(\theta)\sin(\psi), r\sin(\theta))$. On a

$$J_g(r,\theta,\psi) = \begin{pmatrix} \cos(\theta)\cos(\psi) & -r\sin(\theta)\cos(\psi) & -r\cos(\theta)\sin(\psi) \\ \cos(\theta)\sin(\psi) & -r\sin(\theta)\sin(\psi) & r\cos(\theta)\cos(\psi) \\ \sin(\theta) & r\cos(\theta) & 0 \end{pmatrix}$$

Donc

$$\begin{aligned} \text{dét}(J_g(r,\theta,\psi)) &= \sin(\theta)(-r^2\sin(\theta)\cos(\theta)) - r\cos(\theta)(r\cos(\theta)^2) \\ &= -r^2\cos(\theta). \end{aligned}$$

6.3. Norme d'opérateur, dérivées supérieures

Soient E et F deux espaces vectoriels normés. On désine par $\mathscr{L}(E,F)$ l'espace vectoriel des applications linéaire continue de E vers F. Pour tout $\varphi \in \mathscr{L}(E,F)$, on définit

$$\|\varphi\| := \sup_{x\in E\setminus\{0\}} \frac{\|\varphi(x)\|}{\|x\|} = \sup_{y\in E,\, \|y\|=1} \|\varphi(y)\|.$$

C'est un élément dans $\mathbb{R}_{\geqslant 0}$, compte tenu du corollaire 2.7.

Proposition 6.8. — *L'application $\|.\| : \mathscr{L}(E,F) \to \mathbb{R}_{\geqslant 0}$ qui envoie φ en $\|\varphi\|$ est une norme sur $\mathscr{L}(E,F)$.*

Démonstration. — Si $\|\varphi\| = 0$, alors pour tout $x \in E\setminus\{0\}$ on a $\|\varphi(x)\| \leqslant 0$ (et donc $\|\varphi(x)\| = 0$). On en déduit que $\varphi = 0$. En outre, pour tout $\lambda\in\mathbb{R}$ et tout $y \in E$ on a $(\lambda\varphi)(y) = \lambda\varphi(y)$. Donc

$$\|\lambda\varphi\| = \sup_{y\in E,\, \|y\|=1} |\lambda|\cdot\|\varphi(y)\| = |\lambda| \sup_{y\in E,\, \|y\|=1} \|\varphi(y)\| = |\lambda|\cdot\|\varphi\|.$$

Enfin, si φ et ψ sont deux applications linéaires continues, pour tout $y \in E$ tel que $\|y\| = 1$ on a

$$\|(\varphi+\psi)(y)\| = \|\varphi(y)+\psi(y)\| \leqslant \|\varphi(y)\| + \|\psi(y)\|.$$

Donc

$$\|\varphi+\psi\| = \sup_{y\in E,\, \|y\|=1} \|\varphi(y)+\psi(y)\| \leqslant \|\varphi\| + \|\psi\|.$$

□

Définition 6.9. — Soient E et F deux espaces linéaires normés sur $\mathbb{R}$, et f une application définie sur un ouvert U de E et à valeurs dans F. On dit que l'application f est de classe C^0 si elle est continue. De façon récurrsive, pour tout entier $n \geqslant 1$, on dit que f est de classe C^n si elle est différentiable en tout point de U, et l'application $Df : U \to \mathscr{L}(E, F)$ est une application de classe C^{n-1}. On dit que l'application f est de classe C^∞ si elle est de classe C^n pour tout entier $n \geqslant 0$.

Remarque 6.10. — On note $\mathscr{L}^{(1)}(E, F) = \mathscr{L}(E, F)$, et pour tout $n \geqslant 2$, on note (de façon récurssive) $\mathscr{L}^{(n)}(E, F) = \mathscr{L}(E, \mathscr{L}^{(n-1)}(E, F))$. Si $f : U \to E$ est une application de classe C^n, alors l'application

$$D^n f := \underbrace{D(D(D \cdots D}_{n \text{ fois}}(f) \cdots)) : U \longrightarrow \mathscr{L}^{(n)}(E, F)$$

est bien définie et est continue sur U. En outre, pour tout $a \in U$ et tout $(h_1, \ldots, h_n) \in E^n$, on désigne par $D^n f(a)(h_1, \ldots, h_n)$ le vecteur $D^n f(a)(h_1)(h_2) \cdots (h_n) \in F$. L'application $D^n f(a) : E^n \to F$ est n-linéaire (c'est-à-dire qu'elle est linéaire par rapport à chaque coordonnée) et continue. Si h est un élément de E, on utilise l'expression $D^n f(a)(h^{(n)})$ pour désigner $D^n f(a)(\underbrace{h, \ldots, h}_{n \text{ copies}})$.

Définition 6.11. — Soit $\varphi : E^n \to F$ une application n-linéaire. On dit que φ est *symétrique* si, pour toute bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ on a

$$\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varphi(x_1, \ldots, x_n).$$

Proposition 6.12. — *Soit $\varphi : E^n \to F$ une application n-linéaire continue symétrique. Soit $f : E \to F$ l'application qui envoie $x \in E$ en $\varphi(x, \ldots, x)$. Alors l'application f est de classe C^∞, et on a*

$$D^k f(x)(h_1, \ldots, h_k) = \frac{n!}{k!} \varphi(\underbrace{x, \ldots, x}_{n-k \text{ copies}}, h_1, \ldots, h_k), \quad k \in \{0, \ldots, n\};$$

et $D^k f = 0$ lorsque $k > n$.

Démonstration. — On raisonne par récurrence sur n. Le cas où $n = 1$ est trivial. Supposons maintenant que le résultat a été démontré pour $n - 1$ (où $n \geqslant 2$). Pour tous éléments x et h dans E, on a

$$f(x + h) - f(x) = \sum_{k=1}^{n} \binom{n}{k} \varphi(x^{(n-k)}, h^{(k)}) = \varphi(x^{(n-1)}, h) + o(\|h\|).$$

On obtient donc que f est de classe C^1 et que $Df(x)(h) = \varphi(x^{(n-1)}, h)$. Soit

$$\varphi_1 : E^{n-1} \longrightarrow \mathscr{L}(E, F)$$

l'application $(n-1)$-linéaire qui envoie $(x_1,\ldots,x_{n-1})$ en l'application $x_n \mapsto \varphi(x_1,\ldots,x_{n-1},x_n)$, alors on a $Df(x) = \varphi_1(x,\ldots,x)$. On applique l'hypthèse de récurrence à Df et φ_1 et obtient le résultat. $\square$

Théorème 6.13. — *Soient E et F deux espaces vectoriels normés sur $\mathbb{R}$ et f une application définie sur un ouvert U de E et à valeurs dans F, qui est de classe C^n. Alors pour tout $a \in U$ l'application n-linéaire $D^n f(a) : E^n \to F$ est symétrique. En outre, pour tout $(h_1,\ldots,h_n) \in E^n$, l'application $\partial_{h_1}\cdots\partial_{h_n} f : U \to F$ est bien définie et on a*

$$\partial_{h_1}\cdots\partial_{h_n} f(x) = D^n f(x)(h_1,\ldots,h_n)$$

quel que soit $x \in U$.

Démonstration. — Quitte à passer au complété de l'espace vectoriel normé F on peut supposer que F est complet. Comme toute permutation de $\{1,\ldots,n\}$ s'écrit comme composé de transpositions, il suffit de montrer le résultat dans le cas où $n=2$ (appliqué à $D^{n-2}f$ lorsque l'on considère le cas général). Pour tous les nombres réels λ et μ suffisamment petits, soit

$$A(\lambda,\mu) := f(a+\lambda h_1+\mu h_2) - f(a+\lambda h_1) - f(a+\mu h_2) + f(a).$$

La fonction A est différentialble dans un voisinage ouvert de $(0,0)$, et on a

$$\partial_\mu A(\lambda,\mu) = \partial_{h_2} f(a+\lambda h_1+\mu h_2) - \partial_{h_2} f(a+\mu h_2),$$

d'où

$$\begin{aligned} A(\lambda,\mu) &= \mu \int_0^1 \partial_{h_2} f(a+\lambda h_1+s\mu h_2) - \partial_{h_2} f(a+s\mu h_2)\,\mathrm{d}s \\ &= \lambda\mu \int_0^1\int_0^1 \partial_{h_1}\partial_{h_2} f(a+t\lambda h_1+s\mu h_2)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

Comme f est de classe C^2, on obtient $A(\lambda,\mu) = \lambda\mu Df(a)(h_1,h_2) + o(\lambda\mu)$. Par le même argument (en intervertisant les rôles de λ et μ), on obtient

$$A(\lambda,\mu) = \lambda\mu Df(a)(h_2,h_1) + o(\lambda\mu).$$

Le résultat est donc obtenu. $\square$

6.3.1. Formule de Taylor. — La formule de Taylor décrit des propriétés locales fines d'une application différentiable à l'ordre supérieur.

Lemme 6.14. — *Soient E et F deux espaces linéaires normés sur $\mathbb{R}$, et g une application définie sur un ouvert U de E et à valeurs dans F, qui est de classe C^n. Si a est un point de U avec $D^k f(a)$ soit l'application nulle pour tout $k \in \{0,\ldots,n\}$, alors on a $f(x) = o(\|x-a\|^n)$ lorsque $x \mapsto a$.*

Démonstration. — On raisonne par récurrence sur n. Le cas où $n = 0$ est trivial. Supposons que le résultat est démontré pour $n - 1$ ($n \geqslant 1$). Comme f est de classe C^n, elle est continuement différentiable, d'où

$$f(a+h) - f(a) = \int_0^1 \partial_h f(a+th)\,\mathrm{d}t.$$

L'hypothèse de récurrence montre que $\|Df(a+th)\| = o(\|h\|^{n-1})$. Donc

$$\|\partial_h f(a+th)\| = \|Df(a+th)(h)\| = o(\|h\|^n).$$

Le résultat est donc démontré. □

Théorème 6.15. — *Soient E et F deux espaces linéaires normés sur $\mathbb{R}$, et f une application définie sur un ouvert U de E et à valeurs dans F, qui est de classe C^n. Alors pour tout $a \in U$ on a*

$$f(a+h) = \sum_{k=0}^{n} \frac{1}{k!} D^k f(a)(h^{(n)}) + o(\|h\|^n).$$

Démonstration. — L'application

$$g(x) := f(x) - \sum_{k=0}^{n} \frac{1}{k!} D^k f(a)((x-a)^{(k)})$$

de U vers F est de classe C^n, et on a $D^k g(a) = 0$ pour tout $k \in \{0, \ldots, n\}$, compte tenu de la proposition 6.12. Par le lemme 6.14, on obtient le résultat. □

6.4. Comportement local d'une fonction de classe C^2

Dans ce paragraphe, on fixe un espace vectoriel de rang fini E. Soit f une fonction définie sur un sous-ensemble non-vide A de E. On dit qu'un point $a \in A$ est un maximum (resp. minimum) local de la fonction f s'il existe un voisiange U de a dans A tel que $f(a) \geqslant f(x)$ (resp. $f(a) \leqslant f(x)$) quel que soit $x \in U$. On dit que a est un maximum (resp. minimum) global de la fonction f si pour tout $x \in A$ on a $f(a) \geqslant f(x)$ (resp. $f(a) \leqslant f(x)$). On dit que a est un extremum local (resp. global) s'il est un maximum local (resp. global) ou un minimum local (resp. global).

Proposition 6.16. — *Soit f une fonction définie sur un ouvert U de E qui est dérivable en $a \in U$ dans la direction de $h \in E$. Si a est un extremum local de f, alors* $\partial_h f(a) = 0$.

Démonstration. — Sans perte de la généralité, on peut supposer que a est un maximum local de f. Pour tout nombre λ suffisamment petit, on a $f(a+\lambda h) \leqslant f(a)$. Donc

$$\partial_h f(a) = \lim_{\lambda \to 0+} \frac{f(a+\lambda h) - f(a)}{\lambda} \geqslant 0 \quad \text{et} \quad \partial_h f(a) = \lim_{\lambda \to 0-} \frac{f(a+\lambda h) - f(a)}{\lambda} \leqslant 0,$$

d'où $\partial_h f(a) = 0$. □

Remarque 6.17. — Soient f un fonctoin définie sur un ouvert non-vide U et a un point de U. On dit que a est un *point critique*de f si la fonction f est différentiable en a et si $Df(a) = 0$. La proposition précédente montre que, si a est un extremum local de la fonction f et si la fonction f est différentiable en a, alors a est un point critique de f.

Soit $\varphi : E \times E \to \mathbb{R}$ une forme bilinéaire symétrique sur E. On dit que φ est définie positive (resp. négative) si pour tout élément non-nul $x \in E$ on a $\varphi(x, x) > 0$ (resp. $\varphi(x, x) < 0$)

Théorème 6.18. — *Soit f une fonction de classe C^2 définie sur un ouvert non-vide U de E. Si a est un point critique de f et que $D^2 f(a)$ soit définie positive (resp. négative), alors a est un minimum (resp. maximum) local de la fonction f.*

Démonstration. — Pour tout vecteur $h \in E$ tel que $\|h\|$ soit assez petit, la formule de Taylor d'ordre 2 montre que

$$f(a + h) = f(a) + Df(a)(h) + D^2 f(a)(h, h) + o(\|h\|^2).$$

Donc $f(a + h) - f(a) = D^2 f(a)(h, h) + o(\|h\|^2)$ puisque a est un point critique de f. Supposons que $D^2 f(a)$ est défini positif. Comme le cercle unité de E est compact, il existe $\varepsilon > 0$ tel que $\|D^2 f(a)(h, h)\| \geqslant \varepsilon \|h\|^2$ pour tout $h \in E$. Par conséquent, si $\|h\|$ est assez petit, on a $f(a + h) - f(a) \geqslant 0$, et donc a est un minimum local de f. La démonstration de l'autre assertion est très similaire et est laissée comme un exercice. □

Définition 6.19. — Soit f une fonction de classe C^2 définie sur un ouvert non-vide U de $\mathbb{R}^d$. Si a est un point de U, on désigne par $H_f(a)$ la matrice symétrique

$$\begin{pmatrix} \partial_{e_1}^2 f(a) & \partial_{e_1}\partial_{e_2} f(a) & \cdots & \partial_{e_1}\partial_{e_d} f(a) \\ \partial_{e_2}^2 f(a) & \partial_{e_2}^2 & \cdots & \partial_{e_2}\partial_{e_d} f(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{e_d}^2 f(a) & \partial_{e_d}\partial_{e_d} f(a) & \cdots & \partial_{e_d}^2 f(a) \end{pmatrix},$$

appelée la matrice hessienne de f en a, où $(e_1, \ldots, e_d)$ désigne la base canonique de $\mathbb{R}^d$. Par définition, si $h = a_1 e_1 + \ldots + a_d e_d$ est un élément de $\mathbb{R}^d$, alors on a

$$D^2 f(a)(h, h) = (a_1, \ldots, a_d) H_f(a) \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix}$$

En particulier, $D^2 f(a)$ est défini positive si et seulement si toutes les valeurs propres de $H_f(a)$ sont strictement positives.

Remarque 6.20. — Considérons une matrice symétrique carrée

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Son polynôme caractéristique est $P(\lambda) = \lambda^2 - (a+c)\lambda + (ac - b^2)$. Donc les valeurs propres de cette matrice sont strictement positives si et seulement si $a + c > 0$ et $ac - b^2 > 0$.