Chapitre 7

Extrema

Trouver les *extrema* d'une fonction de la variable réelle et à valeurs réelles se fait en général en utilisant son tableau de variations. Pour cela, on dérive la fonction (après s'être assuré que celle-ci est dérivable), on détermine ses *points critiques* (c'est-à-dire les valeurs pour lesquelles la dérivée s'annule) et on décide du sens de variation de la fonction sur chaque intervalle borné par deux points critiques. Cette méthode permet d'obtenir les *extrema locaux* et les *extrema globaux* de la fonction.

Dans ce chapitre, on veut obtenir les extrema de fonctions de plusieurs variables. Pour que cela ait un sens, elle sera à valeurs réelles. On va dans un premier temps suivre la méthode de la dimension 1 en déterminant les points critiques de la fonction, mais la suite n'est en revanche pas adaptable. On distinguera deux cas, suivant si le domaine de définition de la fonction est ouvert ou compact.

Dans toute la suite, E désigne un $\mathbb{R}$-espace vectoriel de dimension finie et U un ouvert de E.

7.1 Définitions

On commence par donner les définitions des extrema. On distingue les extrema globaux (ceux dont la valeur majore ou minore la fonction sur tout le domaine de définition) des extrema locaux (ceux dont la valeur majore ou minore la fonction au voisinage de l'extremum).

Définition 1. *Soit D une partie quelconque de E et $f : D \longrightarrow \mathbb{R}$ une fonction à valeurs réelles.*

*(i) La fonction f admet un **minimum global** en $a \in D$ si $f(x) \geq f(a)$ pour tout $x \in D$.*

*(ii) La fonction f admet un **maximum global** en $a \in D$ si $f(x) \leq f(a)$ pour tout $x \in D$.*

*(iii) La fonction f admet un **minimum local** en $a \in D$ si $f(x) \geq f(a)$ dans un voisinage de a (c'est-à-dire s'il existe $r > 0$ tel que $f(x) \geq f(a)$ pour tout $x \in B(a, r) \cap D$).*

*(iv) La fonction f admet un **maximum local** en $a \in D$ si $f(x) \leq f(a)$ dans un voisinage de a (c'est-à-dire s'il existe $r > 0$ tel que $f(x) \leq f(a)$ pour tout $x \in B(a, r) \cap D$).*

*Un **extremum** (local ou global) de f est un minimum ou un maximum (local ou global). Il est dit strict si l'inégalité est stricte pour $x \neq a$ (c'est-à-dire $f(x) > f(a)$ pour un minimum strict et $f(x) < f(a)$ pour un maximum strict).*

Remarque :

Les extrema globaux (qu'on appelle parfois extrema absolus) sont en particuliers des extrema locaux (qu'on appelle parfois extrema relatifs) mais la réciproque est fausse. Par exemple, le point d'abscisse 1 est un minimum local non global ; le point d'abscisse 3 est un maximum local non global et le point d'abscisse -2 est un maximum local et global.

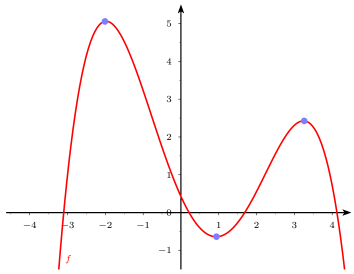

FIGURE 7.1 – Des extrema locaux non globaux.

7.2 Extrema sur un ouvert

7.2.1 Points critiques

Le but de cette section est de déterminer des conditions pour que f admette un extremum local en un point d'un ouvert de E. Commençons par rappeler ce qu'il se passe pour les fonctions d'une seule variable réelle.

Proposition 1 (Extremum d'une fonction dérivable)**.** *Soit I un intervalle* ***ouvert*** *et $f : I \longrightarrow \mathbb{R}$ une fonction dérivable. Si la fonction f admet un extremum local en $c \in I$, alors $f'(c) = 0$.*

Démonstration. L'idée est de regarder le signe du taux d'accroissement de f en l'extremum, de calculer la limite par la gauche et par la droite pour montrer que celle-ci est nécessairement nulle. $\square$

Définition 2. *On dit qu'une application $f : U \subset E \longrightarrow \mathbb{R}$ différentiable admet un* ***point critique*** *en $a \in U$ si $\mathrm{d}f(a)$ est l'application nulle de E dans $\mathbb{R}$, notée $0_{\mathcal{L}(E,\mathbb{R})}$.*

Proposition 2. *Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de E, $f : U \subset E \longrightarrow \mathbb{R}$ différentiable et $a \in U$. On a équivalence entre :*

(i). a est un point critique de f,

(ii). $\forall i \in [\![1;n]\!]$, $\partial_i f(a) = 0$,

(iii). la Jacobienne de f en a est la matrice nulle

en calculant les dérivées partielles et la Jacobienne de f relativement à la base $\mathcal{B}$.

Démonstration. Par construction de la matrice Jacobienne de f en a, il est clair que les deux dernières propriétés sont équivalentes.

- $(i) \Rightarrow (ii)$: On sait que pour tout $i \in [\![1;n]\!]$,

$$\partial_i f(a) = \mathrm{d}f(a)(e_i) = 0.$$

- $(ii) \Rightarrow (i)$: Provient du fait que pour tout $h = \sum_{i=1}^{n} h_i e_i \in E$, on a

$$\mathrm{d}f(a)(h) = \sum_{i=1}^{n} h_i \partial_i f(a) = 0.$$

On aurait aussi pu procéder par implications circulaires en utilisant le fait que la matrice Jacobienne de f en a correspond à la matrice de $\mathrm{d}f(a)$ relative aux bases $\mathcal{B}$ et (1).

□

La proposition suivante est une généralisation de la proposition 1 dans le cas des fonctions de plusieurs variables.

Proposition 3 (Condition nécessaire d'extremum sur un ouvert)**.** *Soit U un* ***ouvert*** *de E et $f : U \longrightarrow \mathbb{R}$ une fonction différentiable. Si f admet un extremum local en $a \in U$, alors a est un point critique de f.*

Démonstration. Soit $h \in E$. On considère la fonction $\varphi : t \longmapsto f(a + th)$. Puisque U est ouvert, il existe $r > 0$ tel que $B(a, r) \subset U$, ce qui montre que φ est définie sur $] - r/\|h\|; r/\|h\|[$ si $h \neq 0_E$ et sur $\mathbb{R}$ sinon. De plus, puisque f admet un extremum local en a, la fonction φ admet un extremum local en 0. On applique donc le résultat 1, ce qui démontre que $\varphi'(0) = 0$. Or la formule de composition montre que

$$\varphi'(0) = \mathrm{d}f(a)(h)$$

ce qui montre que $\mathrm{d}f(a)$ est la fonction nulle. On peut aussi utiliser le fait que la dérivée en 0 de φ correspond à la dérivée directionnelle de f en a selon le vecteur h, qui n'est rien d'autre que $\mathrm{d}f(a)(h)$ dans le cas où f est différentiable en a. □

Remarque.

1. Un point a est un point critique si et seulement si $\mathrm{d}f_a = 0_{\mathcal{L}(\mathbb{R}^n,\mathbb{R})}$ si et seulement si la matrice jacobienne $J_f(a)$ de f en a est nulle si et seulement si le gradient $\mathrm{grad} f(a)$ de f en a est nul si et seulement si les dérivées partielles de f en a sont nulles.
2. La proposition 3 est fausse si l'on ne se place pas sur un ouvert. Pour le cas d'une fonction réelle, il suffit de prendre l'identité sur $[0; 1]$ pour s'en rendre compte. Pour des fonctions de deux variables, considérons la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2$ qui est de classe C^1 sur $\mathbb{R}^2$. Alors la restriction de f au cercle unité $\mathcal{C}$ est constante égale à 1 donc elle admet des extrema en tout point de $\mathcal{C}$. Pourtant $\dfrac{\partial f}{\partial x}(x, y) = 2x$ et $\dfrac{\partial f}{\partial y}(x, y) = 2y$ pour tout $(x, y) \in \mathbb{R}^2$, donc on ne peut pas avoir $\dfrac{\partial f}{\partial x}(x, y) = 0$ et $\dfrac{\partial f}{\partial x}(x, y) = 0$ pour un point $(x, y) \in \mathcal{C}$.
3. La proposition 3 montre que les seuls points "candidats" à être des extrema locaux de f sont les points critiques. Mais la réciproque de la proposition 3 est fausse, même pour des fonctions d'une variable comme le montre la fonction cube en 0. Pour des fonctions de deux variables, on peut également avoir ce que l'on appellera des *points selles*. Par exemple, si $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est définie par $f(x, y) = x^2 - y^2$, alors $\dfrac{\partial f}{\partial x}(0, 0) = \dfrac{\partial f}{\partial y}(0, 0) = 0$ et pourtant $(0, 0)$ n'est ni un point de maximum ni de minimum puisque $f(x, 0) = x^2 \geq 0$ et $f(0, x) = -x^2 \leq 0$ pour tout x au voisinage de 0.

Exemple : Déterminer les extrema (locaux et globaux) de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + xy - 1$.

Exercice 1. Déterminer les extrema (locaux et globaux) de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + 4xy - 1$.

7.2.2 Matrice hessienne

Dans toute la suite, on supposera que $E = \mathbb{R}^n$ pour simplifier. On pourrait énoncer des résultats similaires dans le cas d'un espace vectoriel normé de dimension finie en identifiant la fonction avec la fonction de $\mathbb{R}^n$ donnée par

$$\widetilde{f} : (x_1, \dots, x_n) \longmapsto f\left(\sum_{i=1}^{n} x_i e_i\right).$$

Dans l'exercice 1, nous avons cherché à la main les extrema d'une fonction relativement simple. La méthode utilisée sera parfois impraticable sur des exemples plus complexes. Nous allons donc voir un critère donnant des conditions suffisantes pour qu'un point critique soit un extremum local.

Définition 3. *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \longrightarrow \mathbb{R}$ une application. La* ***matrice hessienne*** *de f au point $a \in U$ est, si elle est définie, la matrice des dérivées partielles secondes de f en a :*

$$\begin{aligned} H_f(a) &= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \le i,j \le n} \\ &= \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \dots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} \end{aligned}$$

Remarque. Si f est de classe C^2 sur U, alors elle admet des dérivées partielles d'ordre 2 donc la matrice hessienne de f est définie en tout point de U. De plus, cette matrice est symétrique (et à coefficients réels) par le théorème de Schwarz.

Exercice 2. Déterminer la matrice hessienne de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + 4xy - 1$ en tout point de $\mathbb{R}^2$.

On admet la *formule de Taylor-Young* à l'ordre 2 pour les fonctions de plusieurs variables. Elle est similaire à la formule de Taylor-Young pour les fonctions d'une variable :

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + o\left(h^2\right) \text{ quand } h \to 0 \\ &= f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + h^2\varepsilon(h) \text{ avec } \lim_{h \to 0} \varepsilon(h) = 0 \end{aligned}$$

Cette formule va nous être utile pour décider si un point critique est un extremum.

Théorème 1 (Taylor-Young à l'ordre 2)**.** *Soit U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une fonction de classe C^2 et $a \in U$. Alors il existe une fonction ε, définie sur un voisinage de $0_{\mathbb{R}^n}$ et à valeurs dans $\mathbb{R}$, vérifiant :*

$$f(a+h) = f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \|h\|^2 \varepsilon(h) \text{ avec } \lim_{h \to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Remarque. Remarquons que le terme $\displaystyle\sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}$ correspond simplement à $\mathrm{d}f(a)(h)$. On pourra aussi noter $\|h\|^2\varepsilon(h)$ sous la forme $o(\|h\|^2)$ lorsque $h \to 0_{\mathbb{R}^n}$.

Remarque. Notons $J_f(a)$ la matrice jacobienne de f en a. Soit H le vecteur des coefficients de h dans la base canonique de $\mathbb{R}^n$. La formule de Taylor-Young à l'ordre 2 s'écrit sous forme matricielle :

$$f(a+h) = f(a) + J_f(a)H + \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0$$

en identifiant $\mathcal{M}_1(\mathbb{R})$ avec $\mathbb{R}$.

Exercice 3. Écrire la formule de Taylor-Young à l'ordre 2 de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = x^2 + y^2 + 4xy - 1$ en tout point de $\mathbb{R}^2$ puis en $(0,0)$. Écrire la formule matricielle.

On prépare le résultat suivant qui va nous permettre, dans la plupart des cas, de savoir si un point critique est un point d'extremum et s'il s'agit alors d'un maximum ou d'un minimum. On donne le principe de la preuve de ce résultat dans le cas d'une fonction de deux variables. Soit U un ouvert de $\mathbb{R}^2$ et $f : U \longrightarrow \mathbb{R}$ de classe C^2. La formule de Taylor-Young matricielle à l'ordre 2 donne en $a = (a_1, a_2) \in U$:

$$f(a+h) = f(a) + J_f(a)H + \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^2}} \varepsilon(h) = 0.$$

Si a est un point critique de f, alors la proposition 3 montre que la matrice jacobienne $J_f(a)$ de f en a est nulle. La formule de Taylor-Young devient :

$$f(a+h) - f(a) = \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^2}} \varepsilon(h) = 0.$$

Le signe de $f(a+h) - f(a)$ est donné par le signe de $\frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h)$ au voisinage de $0_{\mathbb{R}^2}$.

Si de plus la matrice hessienne $H_f(a)$ de f en a est diagonale, de la forme $H_f(a) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ avec par exemple $\lambda_1 \leq \lambda_2$, alors :

$$\begin{aligned} q(h) &= \frac{1}{2}{}^tHH_f(a)H \\ &= \frac{1}{2}\left(\lambda_1 h_1^2 + \lambda_2 h_2^2\right) \end{aligned}$$

- Si $\lambda_1 > 0$ et $\lambda_2 > 0$, alors :

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1 h_1^2 + \lambda_2 h_2^2\right) \\ &\geq \frac{1}{2}\left(\lambda_1 h_1^2 + \lambda_1 h_2^2\right) \\ &= \frac{1}{2}\lambda_1 \|h\|_2^2 \end{aligned}$$

Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \\ &\geq \frac{1}{2}\lambda_1\|h\|_2^2 + \|h\|^2\varepsilon(h) \\ &= \|h\|_2^2\left(\frac{1}{2}\lambda_1 + \varepsilon(h)\right) \\ &\geq 0 \end{aligned}$$

pour h assez proche de $0_{\mathbb{R}^2}$ (si on prend $h \neq 0$, on a même $f(a+h) - f(a) > 0$ pour h proche de 0). Ainsi la fonction f atteint un minimum local (strict) en a.

- Si $\lambda_1 < 0$ et $\lambda_2 < 0$, alors :

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1 h_1^2 + \lambda_2 h_2^2\right) \\ &\leq \frac{1}{2}\left(\lambda_2 h_1^2 + \lambda_2 h_2^2\right) \\ &= \frac{1}{2}\lambda_2 \|h\|_2^2 \end{aligned}$$

Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^t H H_f(a) H + \|h\|^2 \varepsilon(h) \\ &\leq \frac{1}{2}\lambda_2 \|h\|_2^2 + \|h\|^2 \varepsilon(h) \\ &= \|h\|_2^2 \left(\frac{1}{2}\lambda_2 + \varepsilon(h)\right) \\ &\leq 0 \end{aligned}$$

pour h assez proche de $0_{\mathbb{R}^n}$ (idem). Ainsi la fonction f atteint un maximum local (strict) en a.

- Si $\lambda_1 < 0$ et $\lambda_2 > 0$, alors d'une part :

$$\begin{aligned} f(a + (0, h_2)) - f(a) &= \frac{1}{2}{}^t H H_f(a) H + \|(0, h_2)\|^2 \varepsilon(h) \\ &= \frac{1}{2}\lambda_2 h_2^2 + h_2^2 \varepsilon(0, h_2) \\ &= h_2^2 \left(\frac{\lambda_2}{2} + \varepsilon(0, h_2)\right) \\ &> 0 \end{aligned}$$

pour h_2 non nul proche de 0. D'autre part :

$$\begin{aligned} f(a + (h_1, 0)) - f(a) &= \frac{1}{2}{}^t H H_f(a) H + \|(h_1, 0)\|^2 \varepsilon(h) \\ &= \frac{1}{2}\lambda_1 h_1^2 + h_1^2 \varepsilon(h_1, 0) \\ &= h_1^2 \left(\frac{\lambda_1}{2} + \varepsilon(h_1, 0)\right) \\ &< 0 \end{aligned}$$

pour h_1 non nul proche de 0. Ainsi a n'est pas un extremum de f.

Cependant la matrice hessienne est rarement diagonale. Mais on peut toujours la diagonaliser puisqu'elle est symétrique et à coefficients réels.

Proposition 4 (Condition suffisante d'extremum sur un ouvert)**.** *Soit U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et $a \in U$ un point critique de f.*

(i) Si toutes les valeurs propres de $H_f(a)$ sont strictement positives, alors f admet un minimum local strict en a.

(ii) Si toutes les valeurs propres de $H_f(a)$ sont strictement négatives, alors f admet un maximum local strict en a.

(iii) *Si $H_f(a)$ admet au moins une valeur propre strictement positive et une strictement négative, alors f n'admet pas d'extremum local en a. On dit que a est un* ***point col****, ou* ***point selle****.*

Démonstration. La preuve peut être passée, elle n'est pas difficile mais un peu longue, et fait le lien avec le cours d'algèbre sur la réduction des matrices symétriques réelles. La formule de Taylor-Young matricielle à l'ordre 2 donne en $a \in U$:

$$f(a+h) = f(a) + J_f(a)H + \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Comme a est un point critique de f, la proposition 3 montre que la matrice jacobienne $J_f(a)$ de f en a est nulle. D'où :

$$f(a+h) - f(a) = \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0_{\mathbb{R}^n}} \varepsilon(h) = 0.$$

Le signe de $f(a+h) - f(a)$ est donné par le signe de $\frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h)$ au voisinage de $0_{\mathbb{R}^n}$.

Comme la matrice $H_f(a)$ est symétrique et à coefficients réels, il existe une matrice orthogonale $P \in O_n(\mathbb{R})$ telle que $P^{-1}H_f(a)P = D$ est diagonale, de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$. Soit $h \in \mathbb{R}^n$. On note H sa matrice coordonnées dans la base canonique de $\mathbb{R}^n$ et H' sa matrice coordonnées dans la base orthonormée formés par les vecteurs propres de $H_f(a)$ (qui sont aussi les colonnes de la matrice P). La formule de changement de base pour les vecteurs donne $H = PH'$. Comme P est une matrice orthogonale, les vecteurs H et H' ont la même norme euclidienne. Alors :

$$\begin{aligned} q(h) &= \frac{1}{2}{}^tHH_f(a)H \\ &= \frac{1}{2}{}^t(PH')H_f(a)(PH') \\ &= \frac{1}{2}{}^tH'{}^tPH_f(a)PH' \\ &= \frac{1}{2}{}^tH'DH' \\ &= \frac{1}{2}\left(\lambda_1^2h_1'^2 + \cdots + \lambda_n^2h_n'^2\right) \end{aligned}$$

- Si toutes les valeurs propres de $H_f(a)$ sont strictement positives, c'est-à-dire $\lambda_i > 0$ pour tout $i \in \{1, \ldots, n\}$, alors en supposant par exemple $\lambda_1 = \min(\lambda_1, \ldots, \lambda_n)$:

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1h_1'^2 + \cdots + \lambda_nh_n'^2\right) \\ &\geq \frac{1}{2}\left(\lambda_1h_1'^2 + \cdots + \lambda_1'h_n'^2\right) \\ &= \frac{1}{2}\lambda_1\|H'\|_2^2 \\ &= \frac{1}{2}\lambda_1\|H\|_2^2 \\ &= \frac{1}{2}\lambda_1\|h\|_2^2 \end{aligned}$$

Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^tHH_f(a)H + \|h\|^2\varepsilon(h) \\ &\geq \frac{1}{2}\lambda_1\|h\|_2^2 + \|h\|^2\varepsilon(h) \\ &= \|h\|_2^2\left(\frac{1}{2}\lambda_1 + \varepsilon(h)\right) \\ &\geq 0 \end{aligned}$$

pour h assez proche de $0_{\mathbb{R}^n}$. Ainsi la fonction f atteint un minimum local en a.

- Si toutes les valeurs propres de $H_f(a)$ sont strictement négatives, c'est-à-dire $\lambda_i < 0$ pour tout $i \in \{1,\dots,n\}$, alors en supposant par exemple $\lambda_n = \max(\lambda_1,\dots,\lambda_n)$:

$$\begin{aligned} q(h) &= \frac{1}{2}\left(\lambda_1 h_1'^2 + \cdots + \lambda_n h_n'^2\right) \\ &\leq \frac{1}{2}\left(\lambda_n h_1'^2 + \cdots + \lambda_n h_n'^2\right) \\ &= \frac{1}{2}\lambda_n \|H'\|_2^2 \\ &= \frac{1}{2}\lambda_n \|H\|_2^2 \\ &= \frac{1}{2}\lambda_n \|h\|_2^2 \end{aligned}$$

Alors :

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}{}^t\!H H_f(a) H + \|h\|^2 \varepsilon(h) \\ &\leq \frac{1}{2}\lambda_n \|h\|_2^2 + \|h\|^2 \varepsilon(h) \\ &= \|h\|_2^2 \left(\frac{1}{2}\lambda_n + \varepsilon(h)\right) \\ &\leq 0 \end{aligned}$$

pour h assez proche de $0_{\mathbb{R}^n}$. Ainsi la fonction f atteint un maximum local en a.

- Si par exemple $\lambda_1 < 0$ et $\lambda_2 > 0$, alors d'une part en notant h_1 un vecteur propre normé de $H_f(a)$ associé à la valeur propre λ_1 (dont on note H_1 la matrice coordonnées dans la base canonique de $\mathbb{R}^n$) et $t \in \mathbb{R}$ assez petit :

$$\begin{aligned} f(a+th_1) - f(a) &= \frac{1}{2}{}^t\!(tH_1) H_f(a) (tH_1) + \|th_1\|_2^2 \varepsilon(th_1) \\ &= \frac{1}{2}t^2\, {}^t\!H_1 \lambda_1 H_1 + |t|^2 \|h_1\|_2^2 \varepsilon(th_1) \\ &= \frac{1}{2}t^2 \lambda_1 \|h_1\|_2^2 + t^2 \varepsilon(th_1) \\ &= t^2 \left(\frac{\lambda_1}{2} + \varepsilon(th_1)\right) \\ &\leq 0 \end{aligned}$$

D'autre part en notant h_2 un vecteur propre normé de $H_f(a)$ associé à la valeur propre λ_2 (dont on note H_2 la matrice coordonnées dans la base canonique de $\mathbb{R}^n$) et $t \in \mathbb{R}$ assez petit :

$$\begin{aligned} f(a+th_2) - f(a) &= \frac{1}{2}{}^t\!(tH_2) H_f(a) (t) H_2 + \|th_2\|_2^2 \varepsilon(th_2) \\ &= \frac{1}{2}t^2\, {}^t\!H_2 \lambda_2 H_2 + |t|^2 \|h_2\|_2^2 \varepsilon(th_2) \\ &= \frac{1}{2}t^2 \lambda_2 \|h_2\|_2^2 + t^2 \varepsilon(th_2) \\ &= t^2 \left(\frac{\lambda_2}{2} + \varepsilon(th_2)\right) \\ &\geq 0 \end{aligned}$$

Ainsi a n'est pas un extremum de f.

□

Remarque : dans tous les autres cas, la proposition 4 ne permet pas de conclure.

On peut utiliser la proposition 4 dans $\mathbb{R}^2$ sans même calculer les valeurs propres de la matrice hessienne.

Corollaire 1. *Soit U un ouvert de $\mathbb{R}^2$, $f : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction de classe C^2 et a un point critique de f.*

(*i*) *Si* $\det(H_f(a)) > 0$ *et* $\mathrm{Tr}(H_f(a)) > 0$*, alors f admet un minimum local strict en a.*

(*ii*) *Si* $\det(H_f(a)) > 0$ *et* $\mathrm{Tr}(H_f(a)) < 0$*, alors f admet un maximum local strict en a.*

(*iii*) *Si* $\det(H_f(a)) < 0$*, alors f n'admet pas d'extremum en a.*

Démonstration. On commence par relier les valeurs propres de la proposition 4 au déterminant et à la trace de la matrice hessienne. La matrice $H_f(a)$ est symétrique et à coefficients réels, donc le théorème spectral donne $P \in O_2(\mathbb{R})$ telle que $P^{-1}H_f(a)P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, avec λ_1 et λ_2 les valeurs propres de $H_f(a)$. Comme le déterminant et la trace sont invariants par changement de base, on a :

$$\det(H_f(a)) = \lambda_1\lambda_2 \text{ et } \mathrm{Tr}(H_f(a)) = \lambda_1 + \lambda_2.$$

(*i*) Si $\det(H_f(a)) > 0$ et $\mathrm{Tr}(H_f(a)) > 0$, alors la première condition montre que λ_1 et λ_2 ont le même signe et la deuxième condition montre que λ_1 et λ_2 sont strictement positifs. La proposition 4 montre que f admet un minimum local en a.

(*ii*) Si $\det(H_f(a)) > 0$ et $\mathrm{Tr}(H_f(a)) < 0$, alors la première condition montre que λ_1 et λ_2 ont le même signe et la deuxième condition montre que λ_1 et λ_2 sont strictement négatifs. La proposition 4 montre que f admet un maximum local en a.

(*iii*) Si $\det(H_f(a)) < 0$, alors λ_1 et λ_2 ont des signes opposés. La proposition 4 montre que f n'admet pas d'extremum en a.

□

Remarque. Si $\det(H_f(a)) = 0$, on ne peut pas conclure avec le corollaire. Il faut alors étudier "à la main" le signe de $f(x) - f(a)$ au voisinage de a afin de déterminer si f admet un extremum (local) en a.

Exercice 4. Déterminer les extrema locaux de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^3 + y^3 - 3xy$.

Exercice 5. Déterminer les extrema locaux de la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x, y) = x^2\,(1 + x + y)$.