Extremums

1. Formule de Taylor à l'ordre 2

- 1.1.
- 1.2.
- 1.3.
- 1.4.

Mini-exercices.

1

2. Dérivées partielles d'ordre 2, hessienne

Soit $f:(x,y,z)\mapsto f(x,y,z)$ une fonction de classe \mathscr{C}^1 sur l'ouvert U de \mathbb{R}^3 . Les tois dérivées partielles premières sont encore des fonctions de U sur \mathbb{R} . Si elles-mêmes sont de classe \mathscr{C}^1 sur U, on dit que f est de classe \mathscr{C}^2 sur U. Leurs dérivées partielles, au nombre de 9, sont les dérivées partielles secondes de f. Le théorème de Schwarz dit que le résultat ne dépend pas de l'ordre dans lequel on effectue les dérivations.

Théorème 1 (Théorème de Schwarz).

Soit $f: U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction de classe \mathscr{C}^2 . Pour tout i, j = 1, ..., n, on a :

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right).$$

La notation pour la dérivée partielle seconde par rapport à x_i et x_j est $\frac{\partial^2 f}{\partial x_i \partial x_j}$. Leur matrice est la matrice hessienne de f, qui est symétrique d'après le théorème de Schwarz.

Définition 1.

Soit $f:(x,y,z)\mapsto f(x,y,z)$ une fonction de classe \mathscr{C}^2 sur l'ouvert U de \mathbb{R}^3 . La matrice hessienne de f au point $A \in U$ est

$$\mathbf{H}_{f}(A) = \begin{pmatrix} \frac{\partial^{2} f}{\partial x^{2}}(A) & \frac{\partial^{2} f}{\partial y \partial x}(A) & \frac{\partial^{2} f}{\partial z \partial x}(A) \\ \frac{\partial^{2} f}{\partial x \partial y}(A) & \frac{\partial^{2} f}{\partial y^{2}}(A) & \frac{\partial^{2} f}{\partial z \partial y}(A) \\ \frac{\partial^{2} f}{\partial z \partial x}(A) & \frac{\partial^{2} f}{\partial z \partial y}(A) & \frac{\partial^{2} f}{\partial z^{2}}(A) \end{pmatrix}.$$

Exemple. Soit *f* la fonction définie par

$$f(x,y) = \frac{xy^3}{x^2 + y^2}$$
 si $(x,y) \neq (0,0)$ et $f(0,0) = 0$.

On vérifie que f est de classe \mathscr{C}^1 sur \mathbb{R}^2 et que

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y^5 - x^2 y^3}{(x^2 + y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

et

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{3x^3y^2 + xy^4}{(x^2 + y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Le taux d'accroissement

$$\frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y - 0} = 1 \longrightarrow_{y \to 0} 1$$

ce qui montre que $\frac{\partial^2 f}{\partial v \partial x}(0,0) = 1$. De même, le taux d'accroissement

$$\frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x - 0} = 0 \xrightarrow[x \to 0]{} 0$$

ce qui montre que $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 0$. On en déduit que l'une des dérivées partielles secondes $\frac{\partial^2 f}{\partial x \partial y}$ ou $\frac{\partial^2 f}{\partial v \partial x}$ n'est pas continue en (0,0).

Théorème 2 (Formule de Taylor).

Soit $f: U \to \mathbb{R}$ une fonction de classe \mathscr{C}^2 sur l'ouvert $U \subset \mathbb{R}^2$ et soit $(a,b) \in U$. Alors, pour tout

$$f(x,y) \in U,$$

$$f(x,y) = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

$$+ \frac{1}{2} \left[(x-a)^2 \frac{\partial^2 f}{\partial x^2}(a,b) + 2(x-a)(y-b) \frac{\partial^2 f}{\partial x \partial y}(a,b) + (y-b)^2 \frac{\partial^2 f}{\partial y^2}(a,b) \right]$$

$$+ o\left(\|(x-a,y-b)\|^2 \right).$$

On dit aussi que f admet un développement limité d'ordre 2 au point (a, b).

2.1. Pujo

Avant de donner les formules de Taylor, qui dépendent des différentielles d'ordre $n \ge 1$ dans différents cas de fonctions : fonctions de \mathbb{R} dans \mathbb{R} , fonctions de \mathbb{R} dans \mathbb{R}^q et fonctions de \mathbb{R}^p dans \mathbb{R}^q , nous allons donner un aperçu de ce qu'est une différentielle d'ordre 2 dans un premier temps, avec un théorème de symétrie important : le théorème de Schwarz. Nous définirons également la matrice Hessienne qui nous servira beaucoup dans le chapitre ?? sur les extrema.

2.2. Applications deux fois différentiables

Définition 2 (APPLICATIONS DEUX FOIS DIFFERENTIABLES).

Une fonction f définie sur un OUVERT (non vide) $U \subset \mathbb{R}^p$ et à valeurs dans \mathbb{R}^q est dite deux fois différentiable en $x \in U$ si

- 1. elle est différentiable dans un voisinage ouvert U_x de x et si,
- 2. sa différentielle $df: U_x \to \mathbb{E}(\mathbb{R}^p; \mathbb{R}^q)$ est différentiable en x.

On dit que f est deux fois différentiable dans U si elle est différentiable en tout point de U.

Remarque.

Par sa définition, la différentielle de df en x, que l'on écrit $d(df)_x$ est une application linéaire continue de \mathbb{R}^p dans $\mathbb{E}(\mathbb{R}^p; \mathbb{R}^q)$. Autrement dit, on a

$$df: U \to \mathcal{L}(\mathbb{R}^p; \mathbb{R}^q).$$

et

$$d(df)_x: U \to \mathcal{L}(\mathbb{R}^p, \mathcal{L}(\mathbb{R}^p; \mathbb{R}^q)).$$

Mais elle s'identifie naturellement avec une application linéaire continue sur $\mathbb{R}^p \times \mathbb{R}^p$ (c'est à dire une application bilinéaire continue sur \mathbb{R}^p) grâce à la proposition suivante.

Proposition 1 (HORS-PROGRAMME : ESPACES ISOMETRIQUES).

Les espaces $L(\mathbb{R}^p; L(\mathbb{R}^q; \mathbb{R}^n))$ et $L(\mathbb{R}^p, \mathbb{R}^q; \mathbb{R}^n)$ munis des normes usuelles sont isométriques.

Preuve. Pas faite en cours.

Définition 3 (DIFFERENTIELLE SECONDE).

La différentielle seconde d'une fonction $f:U\subset\mathbb{R}^p\to\mathbb{R}^q$ deux fois différentiable est l'application

$$d^2f: U \to \mathbb{L}(\mathbb{R}^p, \mathbb{R}^p; \mathbb{R}^q)$$
$$x \mapsto d^2f_x$$

définie par

$$d^2 f_x(h,k) = d(df)_x(h)(k)$$
 pour tout $(h,k) \in \mathbb{R}^p \times \mathbb{R}^p$.

Remarque.

On peut interpréter cette définition de la façon suivante (qu'on utilise en pratique pour calculer d^2f). Si f est deux fois différentiable sur U, alors, quel que soit $k \in \mathbb{R}^p$, l'application

$$g: U \to \mathbb{R}^q$$
$$x \mapsto df_x(k)$$

est différentiable et

$$dg_x(h) = d^2 f_x(h, k).$$

Théorème 3 (THEOREME DE SCHWARZ).

Si $f:U\subset\mathbb{R}^p\to\mathbb{R}^q$ est deux fois différentiable en x alors d^2f_x est une application bilinéaire SYMETRIQUE. Autrement dit, pour tout $(h, k) \in \mathbb{R}^p \times \mathbb{R}^p$, on a

$$d^2f_x(h,k) = d^2f_x(k,h).$$

Preuve. Pas faite en cours.

2.3. Exemples de différentielles d'ordre 2

Donnons ici quelques différentielles d'ordre 2 pour deux types de fonctions classiques : les applications affines et les applications quadratiques.

- 1. Une application affine $f: x \mapsto l(x) + b$ avec $l \in \mathbb{R}^p$; \mathbb{R}^q) et $b \in \mathbb{R}^q$ est deux fois différentiable et sa différentielle seconde est identiquement nulle.
- 2. Une application quadratique $f: x \mapsto \phi(x, x)$ avec $\phi \in \mathbb{E}(\mathbb{R}^p, \mathbb{R}^p; \mathbb{R}^q)$ est deux fois différentiable et sa différentielle seconde est constante, et même égale à 2ϕ si ϕ est symétrique.

2.4. Matrice Hessienne

Définition 4 (MATRICE HESSIENNE).

Soit $f:U\subset\mathbb{R}^p\to\mathbb{R}$ et soit $(e_1,...,e_p)$ la base canonique de \mathbb{R}^p . Si f est deux fois différentiable sur l'ouvert U alors pour tout $x \in E$, pour tous $i, j \in \{1, ..., p\}$

$$d^2 f_x(e_i, e_j) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x).$$

Alors la matrice

$$d^{2}f_{x} := Hess f_{x} := \begin{pmatrix} \frac{\partial^{2}f}{\partial x_{1}^{2}}(x) & \dots & \frac{\partial^{2}f}{\partial x_{1}\partial x_{p}}(x) \\ \vdots & & \vdots \\ \frac{\partial^{2}f}{\partial x_{p}\partial x_{1}}(x) & \dots & \frac{\partial^{2}f}{\partial x_{p}^{2}}(x) \end{pmatrix}$$

est appelée matrice hessienne de f en x.

Le théorème de Schwarz montre que les dérivées partielles croisées sont égales, c'est à dire

$$\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}$$

pour tous $i, j \in \{1, ..., p\}$. Et donc la matrice hessienne est symétrique. Ces dérivées sont en général notées

$$\frac{\partial^2 f}{\partial x_i \partial x_i}.$$

Par bilinéarité, si h et k sont deux vecteurs de \mathbb{R}^p de composantes $(h_1,...,h_p)$ et $(k_1,...,k_p)$ respectivement, alors

$$d_x^2 f(h,k) = {}^t h.Hess f_a.k = \sum_{i=1}^p \sum_{j=1}^p h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

Autrement dit, $Hess\ f_a$ est la matrice de la forme bilinéaire d^2f_a par rapport à la base canonique de \mathbb{R}^p . L'égalité de Schwarz assure de plus que la matrice hessienne est symétrique.

2.5. Fichou: En dimension 1

Donnons-nous une fonction d'une variable f définie sur \mathbb{R} et de classe C^2 .

Théorème 4.

(Formule de Taylor à l'ordre 2) Si f est une fonction trois fois continument dérivable sur \mathbb{R} , il existe une fonction ϵ tendant vers 0 en 0 telle que

$$f(x+h) = f(x) + hf'(x) + f''(x)h^{2}/2 + |h|^{2}\varepsilon(h)$$

Choisissons pour x un point x = a tel que $f''(a) \neq 0$. Alors pour h assez petit, le terme $f''(a)h^2/2 + |h|^2 \varepsilon(h)$ est du même signe que f''(a). Si par exemple f''(a) > 0, on en déduit que f a un minimum local en x.

La recherche pratique des extrema locaux pour une fonction d'une variable se passe donc ainsi :

- (i) On recherche les points critiques (f'(x) = 0).
- (ii) On étudie la dérivée seconde f'' si a est un point critique et si
 - f''(a) > 0 il y a un minimum local,
 - f''(a) < 0 il y a un maximum local,
 - f''(a) = 0 il faut approfondir l'étude.

Lorsque f n'est plus définie sur \mathbb{R} entier, ou sur un intervalle ouvert, il faudra de plus étudier le comportement de f sur les bords du domaine de définition. si l'ensemble de départ est compact, on a la garantie de l'existence d'extrema globaux.

- 3.
- 3.1.
- 3.2.
- 3.3.
- 3.4.

Mini-exercices.

4. Extremum et points critiques

Définition 5.

Soit $f: U \subset \mathbb{R}^2 \to \mathbb{R}$, où U est un ouvert. On dit que f admet un maximum (resp. minimum) local en $A \in U$, s'il existe une boule ouverte $B \subset U$, centrée en A, telle que :

$$\forall M \in B, f(M) \leq f(A), (\text{resp. } f(M) \geq f(A).$$

On dit que f admet un extremum local en A, si elle y admet un maximum local ou un minimum local.

Proposition 2.

Soit $f: U \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe \mathscr{C}^1 sur un ouvert U et $A \in U$. Si f possède un extremum local en un point A, alors le gradient de f au point A est nul : $\frac{\partial f}{\partial x}(A) = \frac{\partial f}{\partial y}(A) = 0$.

Démonstration. Soit $\vec{v} \in \mathbb{R}^2$. La fonction $\varphi: t \mapsto f(A + t\vec{v})$, définie pour t assez petit, est dérivable au voisinage de 0 et admet un extremum en 0. Donc $\varphi'(0) = 0$ et puis

$$0 = \varphi'(0) = d_A f(\vec{v}) = \langle \nabla f(A), \vec{v} \rangle.$$

Or ceci est vrai pour tout vecteur \vec{v} , donc $\nabla f(A) = 0$.

Les points de U où le gradient de f s'annule sont appelés points critiques de f. Le résultat précédent dit que les extremums d'une fonction ne peuvent se produire qu'en un point critique. La réciproque est fausse.

Théorème 5.

Soit U un ouvert de \mathbb{R}^2 , f une fonction de classe \mathscr{C}^2 sur U et $(a,b) \in U$ un point critique de f.

1. Si $H_f(a,b)$ a toutes ses valeurs propres strictement positives, alors f présente un minimum en (a,b).

- 2. Si $H_f(a, b)$ a toutes ses valeurs propres strictement négatives, alors f présente un maximum en (a, b).
- 3. Si $H_f(a,b)$ a deux valeurs propres de signes opposés, alors f ne présente pas d'extremum en (a,b). On dit que (a,b) est un point selle.
- 4. Dans les autres cas, on ne peut rien dire (tout peut arriver).

Démonstration. Pour h et k, assez petits, non nuls, la différence f(a+h,b+k)-f(a,b) est du signe de

$$q(h,k) = \frac{\partial^2 f}{\partial x^2}(a,b)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a,b)hk + \frac{\partial^2 f}{\partial y^2}(a,b)k^2 = \begin{pmatrix} h & k \end{pmatrix} H_f(a,b) \begin{pmatrix} h \\ k \end{pmatrix}.$$

Il se trouve que, comme pour toute matrice symétrique réelle, $H_f(a, b)$ est diagonalisable : il existe une matrice orthogonale $P \in \mathcal{M}_2(\mathbb{R})$, $P^{-1} = {}^t P$, et deux réels λ et μ tels que

$$H_f(a,b) = P \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} P^{-1}.$$

Posons

$$\begin{pmatrix} h^* \\ k^* \end{pmatrix} = {}^t P \begin{pmatrix} h \\ k \end{pmatrix} \iff \begin{pmatrix} h^* & k^* \end{pmatrix} = \begin{pmatrix} h & k \end{pmatrix} P.$$

La quantité q(h,k) s'écrit

$$q(h,k) = \begin{pmatrix} h^* & k^* \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} h^* \\ k^* \end{pmatrix} = \lambda (h^*)^2 + \mu (k^*)^2.$$

Le signe de q(h, k) dépend donc du signe des valeurs propres λ et μ .

- Si $\lambda > 0$ et $\mu > 0$, alors q(h, k) > 0. On a un minimum local.
- Si $\lambda < 0$ et $\mu < 0$, alors q(h, k) < 0. On a un maximum local.
- Si $\lambda > 0$ et $\mu < 0$, alors q(h,k) > 0 dans la direction $(h^*,0)P$, et q(h,k) < 0 dans la direction $(0,k^*)P$. On a un point selle.

Rappel. Les valeurs propres de $H_f(a, b)$ sont les racines du polynôme caractéristique de $H_f(a, b)$ qui est défini par $\chi(X) = \det[H_f(a, b) - XI_2]$.

Exemple 1. Soit $f(x, y) = x^2 + y^2$. Le point (0, 0) est l'unique point critique de f et $H_f(0, 0) = 2I_2$.

Les valeurs propres de la hessienne de f au point (0,0) sont toutes les deux strictement positives. Donc f admet un minimum local en (0,0).

Exemple 2. Soit $f(x,y) = x^2 - y^2$. On trouve un seul point critique : (0,0). La hessienne de f au point (0,0), $H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, admet une valeur propre strictement positive et une valeur propre strictement négative. Donc (0,0) n'est ni un maximum ni un minimum; c'est un point selle (cf. Figure 1 : selle de cheval).

Exemple 3. Soit $f(x,y) = x^3 - 3xy^2$. On trouve un seul point critique : (0,0). La hessienne de f au point (0,0), $H_f(0,0) = (0)$, a deux valeur propre nulles. On ne peut pas conclure. Or,

 $f_0: x \mapsto f(x,0) = x^3$ possède un point d'inflexion au point 0. Donc (0,0) est un point selle (cf. Figure 2 : selle de singe).

Exemple 4. Soit $f(x,y) = x^4 + y^4 - 2x^2$. On trouve trois points critiques (-1,0), (1,0) et (0,0). Par ailleurs.

$$H_f(-1,0) = H_f(1,0) = \begin{pmatrix} 8 & 0 \\ 0 & 0 \end{pmatrix}$$
 et $H_f(0,0) = \begin{pmatrix} -4 & 0 \\ 0 & 0 \end{pmatrix}$.

On ne peut pas conclure. Or

$$f(x,y) = (x^2 - 1)^2 + y^4 - 1 \ge -1 = f(\pm 1, 0).$$

Donc f admet un minimum aux points ($\pm 1,0$). D'autre part, pour $|x| \leq 1$,

$$f(x,0) = x^4 - 2x^2 = x^2(x^2 - 2) \le 0 = f(0,0)$$
 et $f(0,y) = y^4 \ge 0 = f(0,0)$.

Donc (0,0) n'est ni un maximum ni un minimum; c'est un point selle.

Exemple 5. Soit $f(x, y) = 2x^3 - y^4 - 3x^2$. On trouve deux points critiques (0,0) et (1,0). Par ailleurs,

$$\mathbf{H}_f(1,0) = \left(\begin{array}{cc} 6 & 0 \\ 0 & 0 \end{array} \right) \quad \text{et} \quad \mathbf{H}_f(0,0) = \left(\begin{array}{cc} -6 & 0 \\ 0 & 0 \end{array} \right).$$

On ne peut pas conclure. Or, pour $|x-1| \le 1$,

$$f(1,y) = -1 - y^4 \le -1 = f(1,0)$$
 et $f(x,0) = (x-1)^2(2x+1) - 1 \ge -1 = f(1,0)$.

Donc (1,0) n'est ni un maximum ni un minimum; c'est un point selle. D'autre part, pour $|x| \le 1$,

$$f(x,y) = x^2(2x-3) - y^4 \le 0 = f(0,0).$$

Donc f admet un maximum local au point (0,0).

Ce chapitre est consacré à l'étude de l'existence de deux types d'extrema : les extrema libre et les extrema liés. Les seconds correspondent au cas où les extrema sont justement "liés" à des contraintes. Dans les deux cas, nous pourrons définir ce que l'on appelle extrema locaux (ou relatifs) et les extrema globaux (ou absolus). Il se pourra donc que des minima locaux par exemple ne soit pas globaux si l'on étend le domaine de définition de la fonction f.

Nous ne nous intéresserons également qu'à l'étude de minima (pour simplifier le chapitre) étant donné que les maxima des fonctions f peuvent être vus comme les minima des fonction -f.

Enfin, nous ne nous intéresserons qu'aux fonctions scalaires (autrement dit les fonctions f définies sur \mathbb{R}^p à valeurs dans \mathbb{R}).

EXTREMUMS 5. EXTREMA LIBRES 9

5. Extrema libres

Définition 6 (MINIMUM LOCAL ET GLOBAL).

Si f est une fonction définie sur une partie U de \mathbb{R}^p et à valeurs réelles, un point $a \in U$ est un minimum local (ou relatif) de f s'il existe un voisinage V_a de a ouvert dans U tel que

$$f(x) \ge f(a)$$
 pour tout $x \in V_a$.

On dira que a est un minimum global (ou absolu) de f si

$$f(x) \ge f(a)$$
 pour tout $x \in U$.

Un minimum est dit strict si l'inégalité est stricte, c'est à dire f(x) > f(a), pour tout $x \neq a$ (que ce soit local ou global).

Définition 7 (POINT CRITIQUE).

Soient U un ouvert non vide de \mathbb{R}^p et $f: U \to \mathbb{R}$ un application différentiable en $a \in U$. On dit que $a \in U$ est un point critique de f si $df_a = 0$.

Remarque.

- 1. Noter que si $f: I \subset \mathbb{R} \to \mathbb{R}$, $a \in U$ est un point critique de f si f'a) = 0.
- 2. Nous allons voir dans la proposition suivant qu'un extremum est un point critique, mais attention, la réciproque est fausse. Ce n'est pas parce que l'on a un point critique que c'est un extremum. Par exemple, 0 est un point critique de $f: x \mapsto x^3$ et pourtant ce n'est pas un extremum sur \mathbb{R} .

Dans toute la suite, nous allons donner des critères nécessaires (et on l'espère quelques fois suffisants) pour trouver l'existence de ces extrema locaux selon le degré de différentiabilité de la fonction f.

5.1. Condictions nécessaires du premier ordre

Rappelons ici un résultat pour les applications de \mathbb{R} à valeurs dans \mathbb{R} .

Proposition 3 (RAPPEL: EXTREMA ET POINT CRITIQUE).

Soit I un intervalle ouvert non vide de \mathbb{R} . Si $f:I\subset\mathbb{R}\to\mathbb{R}$ dérivable en $a\in I$ et si f admet un extremum local en a alors f'(a)=0.

Si l'on passe maintenant dans un cadre plus général des fonctions de $\mathbb{R}^P \to \mathbb{R}$. Nous obtenons le résultat suivant.

Proposition 4 (REGLE DE FERMAT).

Soient U un ouvert non vide de \mathbb{R}^p et $f:U\to\mathbb{R}$ une application différentiable en $a\in U$. Si f est un extremum local en a alors $df_a=0$ (autrement dit, a est un point critique), ce qui revient à dire que pour tout $i=1,...,p, \frac{\partial f}{\partial x_i}(a)=0$, ou encore le gradient de f en a est nul.

Remarque.

Attention:

1. La condition précédente n'est pas suffisante car *a* peut être un point critique de *f* sans pour autant que *f* possède d'extremum local.

Extremums 5. Extrema libres 10

2. La condition U ouvert est importante. En effet, il se peut que f admette un extremum sur un domaine U inclus dans \mathbb{R}^p alors que ses dérivées partielles ne s'annuleront pas sur U. Ce qui nous ramène au rappel suivant (voir la remarque de la section $\ref{eq:posterior}$).

Proposition 5 (EXTREMA SUR UN COMPACT).

Soit $f: K \subset \mathbb{R}^p \to \mathbb{R}$, où K est un compact de \mathbb{R}^p , une fonction continue alors f atteint ses extrema dans K.

Preuve. Pas faite en cours.

5.2. Conditions du second ordre

Dans cette section nous allons non seulement donner des conditions nécessaires mais également suffisantes qui permettront d'identifier des minima locaux dnas le cas où la fonction admet des différentielles secondes.

Commençons comme dans la section précédente par les fonctions de $\mathbb{R} \to \mathbb{R}$.

Proposition 6 (RAPPEL: EXTREMA ET POINT CRITIQUE).

Soit I un intervalle ouvert non vide de \mathbb{R} . Si $f: I \subset \mathbb{R} \to \mathbb{R}$ dérivable en $a \in I$ et si f admet un MINIMUM local en a alors f'(a) = 0 (comme on l'a vu dans la section suivante), mais si de plus f est deux fois dérivable en a alors $f''(a) \geqslant 0$.

Inversement : Si $b \in I$ est tel que f'(b) = 0 et f''(b) > 0 alors b est un minimum local de f.

Preuve. Pas faite en cours.

Remarque.

Attention!

- 1. Les conditions f'(a) = 0 et $f''(a) \ge 0$ ne sont pas suffisantes! On a besoin d'avoir f''(a) > 0. On pourrait par exemple le voir avec $f: x \mapsto x^3$ en x = 0.
- 2. D'autre part, la condition f''(a) > 0 n'est pas nécessaire (autrement dit on peut avoir $f''(a) \ge 0$. On pourrait par exemple le voir avec $f: x \mapsto x^4$ en x = 0.

Passons ensuite aux fonctions de $\mathbb{R}^p \to \mathbb{R}$.

Proposition 7 (MINIMA LOCAUX ORDRE 2).

Soient U un ouvert non vide de \mathbb{R}^p et $f: U \to \mathbb{R}$ une application différentiable en $a \in U$. Si f est un extremum local en a alors $df_a = 0$ (comme on l'a vu dans la règle de fermat) mais si de plus f est deux fois différentiable en a alors $d^2f_a(a)(h,h) \geqslant 0$ pour tout $h \in \mathbb{R}^p$.

Inversement, si $b \in U$ est telle que $df_b = 0$ et s'il existe C > 0 tel que $d^2f_b(h,h) \geqslant C||h||^2$, pour tout $h \in \mathbb{R}^p$ alors b est un minimum local de f

Remarque.

- 1. On retrouve ici la notion de coercivité donnée dans le rappel tout au début de ce chapitre puisque d^2f_a est une application bilinéaire symétrique (d'après le théorème de Schwarz).
- 2. Comme nous sommes en dimension finie, la condition "s'il existe C>0 tel que $d^2f_b(h,h)\geqslant C\|h\|^2$, pour tout $h\in\mathbb{R}^p$ " revient à dire "s'il existe C>0 tel que $d^2f_b(h,h)>0$, pour tout $h\in\mathbb{R}^p\setminus 0_{\mathbb{R}^p}$ " (car la coercivité est équivalent à dire que la fonction bilinéaire symétrique est définie positive en dimension finie).

Extremums 5. Extrema libres 11

3. Ce résultat peut s'interpréter facilement avec les matrices Hessiennes. C'est ce que nous allons voir dans la section suivante.

5.3. Critères avec les matrices Hessiennes

Théorème 6 (HESSIENNE ET EXTREMA LOCAUX).

Soient $f: U \subset \mathbb{R}^p \to \mathbb{R}$ une fonction deux fois différentiable sur U ouvert de \mathbb{R}^p , et $a \in U$ un point critique de f. Alors la Hessienne de f, Hess $(f) \in \mathcal{M}_n(\mathbb{R})$ est une matrice symétrique et

- 1. $si\ Hessf_a$ est DEFINIE POSITIVE alors a est un minimum local strict de f sur U,
- 2. $si\ Hess f_a$ est SEMI-DEFINIE POSITIVE alors a est un minimum local de f sur U,
- 3. si $Hess f_a$ est DEFINIE NEGATIVE alors a est un maximum local strict de f sur U,
- 4. $si\ Hess f_a\ est\ SEMI-DEFINIE\ POSITIVE\ alors\ a\ est\ un\ maximum\ local\ de\ f\ sur\ U,$
- 5. S'il existe un l tel que $\Delta_{2l} < 0$, ou s'il existe m tel que $\Delta_{1}.\Delta_{2.m} < 0$, alors A est indéfinie.

Preuve. Pas faite en cours.

Remarque.

Pour appliquer ces résultats, il est utile d'utiliser les rappels d'algèbre du début de ce chapitre parmi lesquels :

déterminer si $Hess\ f_a$ est définie négative ou positive revient à déterminer les valeurs propres de $Hess\ f_a$.

- 1. Si toutes les valeurs propres sont > 0, $Hess f_a$ est définie positive.
- 2. Si toutes les valeurs propres sont < 0, $Hess f_a$ est définie négative.
- 3. Si les valeurs propres de $Hess\ f_a$ sont non nulles mais de signes différents, on dit que a est un point col (ou un point selle).

Nous pouvons également donner des résultats sur les extrema globaux :

Théorème 7 (EXTREMA GLOBAUX ET HESSIENNE).

Soit $f: U \subset \mathbb{R}^p \to \mathbb{R}$, où U est un ouvert de \mathbb{R}^p , et f est deux fois différentiable sur U, soit $a \in U$ un point critique de f, alors si

$$(x-a)^{T}$$
. $Hess f(a) \cdot (x-a) \ge 0$ (resp. ≤ 0),

pour tout $x \in U$, alors a est un minimum global de f sur U (resp. maximum global), et si les inégalités sont strictes les extrema sont stricts.

Preuve. Pas faite en cours.

5.4. Cas particulier où $f: \mathbb{R}^2 \to \mathbb{R}$

Dans le cas où p=2 ($f: \mathbb{R}^2 \to \mathbb{R}$, ce que l'on utilisera le plus souvent en exercice), nous utiliserons la notation de Monge qui nous sera bien utile, et plus rapide en général que les critères de section précédentes.

Théorème 8 (NOTATION DE MONGE).

Soient $f: U \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction deux fois différentiable sur U ouvert de \mathbb{R}^2 , et $(a, b) \in U$ un

point critique de f. On pose

$$r = \frac{\partial^2 f}{\partial x^2}(a, b), \ s = \frac{\partial^2 f}{\partial x \partial y}(a, b), \ \text{et } t = \frac{\partial^2 f}{\partial y^2}(a, b).$$

- si rt s² > 0 et r > 0, (a, b) est un minimum local de f sur U,
 si rt s² > 0 et r < 0, (a, b) est un maximum local de f sur U,
 si rt s² < 0 la fonction n'admet pas d'extremum local, on dit alors que (a, b) est un point SELLE ou COL.

Preuve. Pas faite en cours.

5.5. Fichou : Extrema locaux de $f: \mathbb{R}^2 \to \mathbb{R}$

On suppose f de classe \mathscr{C}^2 , c'est-à-dire que ses dérivées partielles jusqu'à l'ordre 2 existent et sont continues.

Définition 8.

f a un **minimum** (resp. **maximum**) local en (x_0, y_0) s'il existe $\varepsilon > 0$ tel que : $\forall (x, y) \in$ $B((x_0, y_0), \varepsilon)$ alors $f(x, y) \ge f(x_0, y_0)$ (resp. $f(x, y) \le f(x_0, y_0)$).

Exemple

$$f(x, y) = -x^{2} - y^{2}$$

$$f(x, y) = x^{2} + y^{2}$$

$$f(x, y) = x^{2} - y^{2}$$

Proposition 8.

Si f admet un **extremum local** en (x_0, y_0) alors $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$.

La fonction de une variable $x \mapsto f(x, y_0)$ admet un extremum local en x_0 , donc sa dérivée $\frac{\partial f}{\partial x}(x_0, y_0)$ en x_0 est nulle. On fait de même avec $y \mapsto f(x_0, y)$.

Définition 9.

Si en (x_0, y_0) on a $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$ on dit que (x_0, y_0) est **un point critique** de f ou un point stationnaire de f.

Remarque

Un extremum local est un point critique mais la réciproque n'est pas vraie.

5.6. Formule de Taylor

Définition 10.

Soit f(x, y) une fonction de classe \mathscr{C}^2 . La **matrice hessienne** de f en (x_0, y_0) est la matrice Hess $(x_0, y_0) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$ où $A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)$, $B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$, $C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$.

où
$$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)$$
, $B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$, $C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$.

Extremums 5. Extrema libres 13

Exemple

Calculer la matrice Hessienne de $f(x, y) = 4xy - x^4 - y^4$.

Théorème 9.

(Formule de Taylor à l'ordre 2, en $X = (x_0, y_0)$)

Si f une fonction de classe \mathscr{C}^2 sur \mathbb{R}^2 , il existe une fonction ϵ tendant vers 0 en (0,0) telle que

$$f(X+H) = f(X) + \langle \nabla f(X), H \rangle + \frac{1}{2}H^{t} \operatorname{Hess}(x_{0}, y_{0}) H + ||H||^{2} \varepsilon(H)$$

Démonstration dans le cas où la hessienne est diagonale.

Théorème 10.

Soient f(x, y) de classe \mathscr{C}^2 et (x_0, y_0) un point critique. Soit $\Delta = AC - B^2 = \det(\operatorname{Hess}(x_0, y_0))$. Alors:

 $si \triangle > 0$ et A > 0, f a un minimum local en (x_0, y_0)

si $\triangle > 0$ et A < 0, f a un maximum local en (x_0, y_0)

 $si \triangle < 0$, f n'a ni maximum ni minimum, elle a un point selle

si $\triangle = 0$ on ne peut conclure (avec le seul développement à l'ordre 2).

Exemple

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3 + y^3 - 3xy$. Les points critiques de f sont (0,0) et (1,1). Le premier est un point col, le second un minimum local (non global).

5.7. Extrema de f sur un compact $K \subset \mathbb{R}^2$

La marche à suivre pour étudier les extrema d'une fonction différentiable sur un compact de \mathbb{R}^2 est la suivante.

Soit f une fonction différentiable définie sur un compact K de \mathbb{R}^2 . Comme f est différentiable, elle est continue. Elle est donc bornée sur K et atteint ses bornes.

En pratique (dans les exercices que je vous demanderai de résoudre en particulier) la fonction f sera donnée par une formule valable sur un certain sous-ensemble de \mathbb{R}^2 et le compact K sera inclus dans cet ensemble de définition.

On mène l'étude des extrema de f en plusieurs étapes. La première est d'étudier l'existence d'extrema locaux de f à l'intérieur de K. C'est pour cette étude qu'on utilisera le développement de Taylor à l'ordre 2 donné ci-dessus.

Mais cette étude n'est pas suffisante. Il faut aussi regarder ce qui se passe sur le bord de K. Pour cela on procède autrement.

Exemple

Etude de $f(x, y) = x^2 - y^2$ sur $K = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \le 1\}$. On procède de la manière suivante :

- (i) On cherche les points critiques et les extrema locaux dans Int(K).

 On trouve un seul point stationnaire en (0,0). Mais en (0,0) f a un point selle. La fonction n'a donc pas d'extremum à l'intérieur de K. Mais comme K est compact et f est continue sur K, f est bornée sur K et atteint ses bornes sur K. Ce sera donc sur le bord de K.
- (ii) On analyse f sur ∂K .

Une possibilité ici est de paramétrer le bord de K: le cercle de rayon 1 centré en (0,0). On obtient : $f(\cos t, \sin t) = \cos^2 t - \sin^2 t = \cos(2t)$. On peut alors étudier les variations de cette fonction. On obtient qu'elle est maximum égale à 1 lorsque 2t est égal à 0 modulo 2π , minimum

Extremums 7. Extrema liés 14

égale à -1 lorsque 2t vaut π modulo 2π . La fonction f atteint donc son maximum 1 aux points (1,0) et (-1,0) de K, son minimum -1 aux points (0,1) et (0,-1).

6. Extremums liés

- 6.1.
- 6.2.
- 6.3.
- 6.4.

Mini-exercices.

1.

7. Extrema liés

Dans cette section, nous allons nous intéresser à la recherche des extrema sous certaines contraintes. Il paraît normal donc de définir ce que sont des contraintes.

7.1. Contraintes

Définition 11 (CONTRAINTES).

Si f et $g_1,...,g_k$ sont des fonctions définies sur un ouvert $U \subset \mathbb{R}^p$ à valeurs dans \mathbb{R} , un point $a \in U$ tel que $g_1(a) = 0,...,g_k(a) = 0$ est un minimum local de f sous les contraintes $g_1,...,g_k$ s'il existe un voisinage V_a de a tel que

$$f(x) \geqslant f(a)$$

pour tout $x \in V_a$ et $g_1(x) = 0, ..., g_k(x) = 0$.

7.2. Extrema liés avec une seule contrainte

Commençons par les fonctions d'un espace \mathbb{R}^2 avec une seule contrainte :

Théorème 11 (EXTREMA LIES ET GRADIENTS).

Soient $f, g: U \in \mathbb{R}^2 \to \mathbb{R}$ de classe \mathscr{C}^1 , sur un ouvert U de \mathbb{R}^2 , soit $(a, b) \in U$ tels que

- 1. f soumise à la contrainte g(x, y) = 0 admette un extremum au point (a, b),
- 2. $\nabla g(a,b) \neq 0$, alors il existe un nombre réel $\lambda \neq 0$ tel que $\nabla f(a,b) = \lambda \nabla g(a,b)$. Autrement

Extremums 7. Extrema liés 15

dit, on a

$$\begin{cases} \frac{\partial f}{\partial x}(a,b) - \lambda \frac{\partial g}{\partial x} &= 0, \\ \frac{\partial f}{\partial y}(a,b) - \lambda \frac{\partial g}{\partial y} &= 0, \\ g(a,b) &= 0. \end{cases}$$

Preuve. Pas faite en cours.

7.3. Extrema liés avec plusieurs contraintes

En généralisant le cas précédent, on peut donner les résultats sous les formes suivantes pour k contraintes.

Définition 12 (Γ REGULIER).

Soit $g: U \in \mathbb{R}^p \to \mathbb{R}^k$ une fonction de classe \mathscr{C}^1 et $\Gamma = g^{-1}(\{0\})$. On dit que Γ est régulier (ou encore qu'il satisfait à la condition de qualification non dégénérée) si pour tout $a \in \Gamma$, $dg_a: \mathbb{R}^p \to \mathbb{R}^k$ est surjective.

Remarque.

Si $g: U \in \mathbb{R}^p \to \mathbb{R}$ (k = 1) (cas où l'on a une seule contrainte) la condition signifie seulement que pour tout $a \in \Gamma$, $dg_a \neq 0$.

On retrouve sans surprise le résultat de la section précédente à une seule contrainte :

Théorème 12 (EXTREMA LIES (UNE CONTRAINTE)).

Soient $f, g: U \in \mathbb{R}^n \to \mathbb{R}$ de classe \mathscr{C}^1 , soit $\Gamma = g^{-1}(\{0\})$ régulière. Si $a \in \Gamma$ est un extremum local de $f_{|_{\Gamma}}$, alors il existe un unique $\lambda \in \mathbb{R}$ tel que

$$df a + \lambda dg_a = 0.$$

Preuve. Pas faite en cours.

N.B. : Le réel λ est appelé multiplicateur de Lagrange.

Regradons maintenant ce qu'il se passe quand on a *k* contraintes.

Théorème 13 (EXTREMA LIES (k CONTRAINTES)).

Soient $f: U \in \mathbb{R}^p \to \mathbb{R}$ de classe \mathscr{C}^1 , $g: U \in \mathbb{R}^p \to \mathbb{R}^k$, et $\Gamma = g^{-1}(\{0\})$ régulière. Si $a \in \Gamma$ est un extremum local de $f_{|r|}$, alors il existe un unique $\lambda = (\lambda_1, ..., \lambda_k) \in \mathbb{R}^k$ tel que

$$df a + \sum_{i=1}^{p} \lambda_i (dg_i)_a = 0.$$

Si jamais on est mal à l'aise avec la notion de Γ régulière, il suffit juste de voir le résultat précédent de la façon suivante :

1. on suppose que les fonctions f et $g_1,...,g_k$ définies de $\mathbb{R}^p \to \mathbb{R}$ sont continûment différentiables.

2. on dit que les contraintes $g_1,...,g_k$ sont indépendantes au point $a \in U$ si la famille de formes linéaires continues $\{(dg_1)_a,...,(dg_k)_a\}$ est libre (ce qui revient exactement à dire que g est Γ régulière. Alors on a le résultat suivant équivalement au théorème précédent :

Extremums 7. Extrema liés 16

Théorème 14 (CONDITION NECESSAIRE MINIMUM SOUS k CONTRAINTES).

Soient f et $g_1,...,g_k$ sont des fonctions de classe \mathscr{C}^1 définies sur un ouvert $U \subset \mathbb{R}^p$ d'un à valeurs dans \mathbb{R} . Soit $a \in U$ tel que $g_1(a) = 0,...,g_k(a) = 0$ et les contraintes $g_1,...,g_k$ sont indépendantes au point a. Si a est un minimum local de f sous les contraintes $g_1,...,g_k$, alors il existe des réels $\lambda_1,...,\lambda_k$ tels que

$$df_a = \lambda_1 (dg_1)_a + ... + \lambda_k (dg_k)_a$$
.

Preuve. Pas faite en cours.

Les résultats présentés dans ce paragraphe sont liés à des problèmes d'extremum essentiellement sur des ouverts : et il est à noter que les conditions nécessaires d'extremum local sont fausses lorsque U n'est pas un ouvert.

Nous allons dans la section suivante considérer des problèmes d'extremum sur des sous-ensemble convexes de E.

7.4. Fichou : Extrema liés (multiplicateur de Lagrange)

Une seule contrainte

Il s'agit de trouver les extrema de f(x, y, z) lorsque (x, y, z) appartient à une surface S définie par g(x, y, z) = c.

Exemple

Maximiser $x^2 y^2 z^2$ lorsque $x^2 + y^2 + z^2 = 1$.

Définition 13.

Un point $X_0 = (x_0, y_0, z_0)$ est un minimum (resp. maximum) local pour f, lié à la contrainte g(x, y, z) = c si :

- (i) $g(X_0) = c$
- (ii) Il existe r > 0 tel que $f(X_0) \le f(X)$ (resp. $f(X_0) \ge f(X)$) pour tout $X \in S \cap B(X_0, r)$.

Théorème 15.

(de Lagrange)

Soit f(x, y, z) et g(x, y, z) de classe \mathscr{C}^1 telle que $\nabla g \neq 0$ sur S.

Alors si f admet un **extrema lié** en (x_0, y_0, z_0) on $a : \nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.

Démonstration. Soit $\gamma: \mathbb{R} \to S$ passant par X_0 en $t=t_0$. Alors $f\circ \gamma$ a un extremum local en t_0 donc sa dérivée s'annule, c'est-à-dire

$$<\nabla f(X_0), \gamma'(t_0)>=0.$$

Ceci étant valable pour toute courbe γ passant par X_0 , on en déduit que le gradient de f en X_0 est orthogonal au plan tangent à S en X_0 , donc est colinéaire au gradient de g en X_0 .

Remarque

Si P est un extremum lié, on a $\nabla f(P)$ parallèle à $\nabla g(P)$. La réciproque n'est pas vraie. Nous avons une condition nécessaire mais pas suffisante. C'est l'équivalent de la nullité de la dérivée pour les extrema libres : en un extremum libre la dérivée est nulle mais la dérivée peut être nulle sans que la fonction ait un extremum (penser à $x \mapsto x^3$ en x = 0).

Extremums 8. 17

Exemple

Sur l'exemple précédent on montre la méthode de résolution.

Autre exemple

Maximum de la distance de l'ellipsoide donné par $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ à l'origine. Le théorème est encore vrai en dimension plus grande.

Théorème 16.

(de Lagrange)

Soit f et g deux fonctions de \mathbb{R}^d dans \mathbb{R} de classe \mathscr{C}^1 telle que $\nabla g \neq 0$ sur l'ensemble S défini par g = C (c'est une hypersurface).

Alors si f admet un **extrema lié** en x_0 on a: $\nabla f(x_0) = \lambda \nabla g(x_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.

Plusieurs contraintes

On cherche les extrema d'une fonction f sur l'ensemble S défini par $g_1 = g_2 = \ldots = g_k = 0$, toutes les fonctions considérées étant de classe C^1 . Pour que les choses marchent bien il faut faire l'hypothèse suivante : en tout point de S les gradients des fonctions g_i sont linéairement indépendants.

Théorème 17.

(de Lagrange)

Soit f et g_1, \ldots, g_k k+1 fonctions de \mathbb{R}^d dans \mathbb{R} de classe \mathscr{C}^1 telles que les vecteurs $\nabla g_1, \ldots, \nabla g_k$, soit indépendants sur sur l'ensemble S défini par $g_1 = \ldots = g_k = 0$.

Alors si f admet un **extrema lié** sur S en x_0 le vecteur $\nabla f(x_0)$ est combinaison linéaire des vecteurs $\nabla g_i(x_0)$: il existe $\lambda_1, \ldots, \lambda_k$ tels que $\nabla f(x_0) = \sum_{i=1}^k \lambda_i \nabla g_i(x_0)$. Les nombres $\lambda_1, \ldots, \lambda_k$ sont appelés multiplicateurs de Lagrange.

- 8.
- 8.1.
- 8.2.
- 8.3.
- 8.4.

Mini-exercices.

1.

Auteurs du chapitre

D'après un cours de ...

Revu et augmenté par Arnaud Bodin.

Relu par Stéphanie Bodin et Vianney Combet.