Théorème des extrema liés

Référence : [Gou08] p.317 et p.327.

Définition 0.1 (Extremum local) *Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application, avec U ouvert de $\mathbb{R}^n$. On appelle $a \in U$ un extremum local si il existe V un voisinage de a tel que $\forall x \in V$, $f(x) \leq a$ (ie a est un maximum local) ou tel que $f(x) \geq a$ (ie a est un minimum local).*

Attention ! le fait que U soit ouvert est essentiel !

Définition 0.2 (Point critique) *Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable au voisinage de a, avec U ouvert de $\mathbb{R}^n$. On appelle $a \in U$ un point critique si $Df(a) = 0$.*

Proposition 0.1 *Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une application différentiable au voisinage de a, avec U ouvert de $\mathbb{R}^n$.*
Si a est un extremum local de f, alors a est un point critique de f.

Remarque : La réciproque est fausse, en effet il suffit de considérer l'application $x \longmapsto x^3$, cette application admet un point critique en 0 qui n'est pas un extremum local.

Théorème 0.2 *Soient $f, g_1, \dots, g_r : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ des fonctions de classe C^1, avec U ouvert de $\mathbb{R}^n$.*
On pose $\Gamma = \{x \in U ; g_1(x) = \dots = g_r(x) = 0\}$ (attention a priori Γ n'a aucune raison d'être ouvert !).
Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et si les formes linéaires $(Dg_i(a))_i$ sont linéairement indépendantes (ie la famille est libre), alors il existe des réels $\lambda_1, \lambda_2, \dots, \lambda_r$ (appelés multiplicateurs de Lagrange) tels que :

$$Df(a) = \lambda_1 Dg_1(a) + \dots + \lambda_r Dg_r(a)$$

Attention ! la condition U ouvert est essentielle ! sinon on aurait $Df(a) = 0$ et la conclusion du théorème n'aurait aucun sens !

Démonstration Soit $s = n - r$. On identifie $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$. On écrit donc les éléments de $\mathbb{R}^n$ sous la forme $(x, y) = (x_1, \dots, x_s, y_1, \dots, y_r)$.
Posons $a = (\alpha, \beta) \in \mathbb{R}^s \times \mathbb{R}^r$.
$\rightsquigarrow$ On a nécessairement $r \leq n$, car les formes linéaires $Dg_i(a)$ forment une famille libre et la dimension de l'espace dual de $\mathbb{R}^n$ est n.
Par ailleurs, si $r = n$, le théorème devient évident car les $Dg_i(a)$ forment une base de $(\mathbb{R}^n)^*$.
On peut donc supposer que $r \leq n - 1$, ie $s \geq 1$.
$\rightsquigarrow$ Les formes linéaires $Dg_i(a)$ forment une famille libre, ainsi la matrice :

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

est donc de rang r.
On peut alors en extraire une sous-matrice $r \times r$ inversible. Quitte à changer le nom des variables, on peut supposer que :

$$\det \left(\frac{\partial g_i}{\partial y_j}(a) \right)_{1 \leq i,j \leq r} \neq 0$$

D'après le théorème des fonctions implicites, on peut donc trouver un voisinage U' de α dans $\mathbb{R}^s$ et un voisinage Ω de a dans $\mathbb{R}^n$ et une fonction $\phi = (\phi_1, \ldots, \phi_r) : U' \longrightarrow \mathbb{R}^r$ de classe C^1 tels que (en notant $g = (g_1, \ldots, g_r)$) :

$$g(x, y) = 0 \quad \text{avec } x \in U' \text{ et } (x, y) \in \Omega \quad \Leftrightarrow y = \phi(x)$$

ie sur un voisinage de a, les éléments de Γ s'écrivent $(x, \phi(x))$.
$\rightsquigarrow$ Posons $h(x) = f(x, \phi(x))$. La fonction h admet donc un extremum local en $x = \alpha$ (car $(\alpha, \phi(\alpha)) = a$ et $(x, \phi(x)) \in \Gamma$) ce qui entraîne $\forall 1 \leq i \leq s$:

$$0 = \frac{\partial h}{\partial x_i}(\alpha) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \phi_j}{\partial x_i}(\alpha) \frac{\partial f}{\partial y_j}(a)$$

Par ailleurs, en écrivant les dérivées partielles par rapport aux x_i de $g(x, \phi(x)) = 0$, on trouve $\forall 1 \leq k \leq r$ et $\forall 1 \leq i \leq s$:

$$0 = \frac{\partial g_k}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \phi_j}{\partial x_i}(\alpha) \frac{\partial g_k}{\partial y_j}(a)$$

ie les s premiers vecteurs colonnes de la matrice :

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \ldots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \ldots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \ldots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \ldots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \ldots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \ldots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

s'expriment, d'après les deux formules aux dérivées partielles ci-dessus, linéairement en fonction de ses r derniers vecteurs colonnes, donc $rg(M) \leq r$.
$\rightsquigarrow$ Or le rang des vecteurs lignes est égal au rang des vecteurs colonnes, donc les $r+1$ vecteurs lignes de M forment une famille liée, ce qui entraîne l'existence de réels $\mu_0, \ldots, \mu_r$ non tous nuls tels que :

$$\mu_0 Df(a) + \mu_1 Dg_1(a) + \ldots + \mu_r Dg_r(a) = 0$$

Comme la famille $(Dg_i(a))_i$ est libre, alors $\mu_0 \neq 0$ donc en posant $\lambda_i = -\frac{\mu_i}{\mu_0}$ pour $1 \leq i \leq r$, on obtient :

$$Df(a) = \sum_{j=1}^{r} \lambda_j Dg_j(a)$$

Lemme utilisé

Lemme 0.1 *Le rang d'une matrice A est l'ordre maximal des mineurs non nuls extraits de A, ie $rg(A) = r$ si et seulement si il existe un mineur d'ordre r non nul et tous les mineurs d'ordre $s > r$ sont nuls.*

Démonstration Supposons que le rang de A est r. S'il existe un mineur d'ordre $s > r$ non nul, on pourrait extraire des colonnes de A une famille libre formée de $s > r$ vecteurs (d'après le lemme qui suit). Or ceci est impossible car les vecteurs colonnes de A engendrent un espace de dimension r.
La réciproque découle d'un lemme qui suit.

Lemme 0.2 *Soient $\{v_1, \ldots, v_r\}$ r vecteurs d'un espace vectoriel E de dimension n $(r \leq n)$ et A la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_r$ dans une base quelconque de E. Alors la famille $\{v_1, \ldots, v_r\}$ est libre si et seulement si on peut extraire de A un mineur non nul.*

Démonstration Montrons d'abord que la condition est suffisante. Soit :

$$\begin{pmatrix} \boxed{\begin{matrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{matrix}} \\ \begin{matrix} \vdots & \quad & \vdots \\ a_{n1} & & a_{nr} \end{matrix} \end{pmatrix}$$

Quitte à permuter les lignes (ce qui n'affecte pas le rang de la matrice), on peut supposer que le mineur encadré est non nul.
Considérons l'application :

$$\begin{aligned} p : \mathbb{K}^n &\longrightarrow \mathbb{K}^r \\ (b_1, \dots, b_r, \dots, b_n) &\longmapsto (b_1, \dots, b_r) \end{aligned}$$

Il s'agit d'une application linéaire. On a, en notant $\| \; p(v_1), \dots, p(v_r) \; \|$ la matrice ayant pour les colonnes les vecteurs $p(v_i)$:

$$\| \; p(v_1), \dots, p(v_r) \; \| = \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}$$

Donc $\det \| \; p(v_1), \dots, p(v_r) \; \| \neq 0$.
Par conséquent $\{p(v_1), \dots, p(v_r)\}$ est une base de $\mathbb{K}^r$, donc une famille libre. On en déduit que $\{v_1, \dots, v_r\}$ est libre. En effet si $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$, en appliquant p et le fait que la famille des $p(v_i)$ est libre, on obtient que les λ_i sont tous nuls.
Réciproquement, supposons la famille $\{v_1, \dots, v_r\}$ libre. On peut alors la compléter par $n - r$ vecteurs de la base canonique de $\mathbb{K}^n$ de manière à former une base (c'est le théorème de la base incomplète, qui nous dit qu'on peut compléter une famille libre en une base en prenant n'importe quels éléments d'une famille génératrice). Quitte à échanger la numérotation de la base canonique, on peut supposer que $\{v_1, \dots, v_r, e_{r+1}, \dots, e_n\}$ est une base. Soit :

$$B = \| \; v_1, \dots, v_r, e_{r+1}, \dots, e_n \; \|_{e_i} = \begin{pmatrix} a_{11} & \dots & a_{1r} & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ a_{r1} & \dots & a_{rr} & 0 & \dots & \dots & 0 \\ a_{r+1,1} & \dots & a_{r+1,r} & 1 & 0 & \dots & 0 \\ \vdots & & \vdots & 0 & 1 & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & \dots & a_{nr} & 0 & 0 & \dots & 1 \end{pmatrix}$$

Comme B est la matrice formée de vecteurs colonnes formant une base de $\mathbb{K}^n$, alors $\det(B) \neq 0$.
Et en développant $\det(B)$ selon la dernière colonne et en répétant successivement cette opération, on trouve que :

$$\det(B) = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix}$$

On peut donc extraire de A un mineur d'ordre r non nul.

Lemme 0.3 *Soient $\{v_1, \dots, v_r\}$ r vecteurs linéairement indépendants. Soit A la matrice dont les colonnes sont les composantes de ces vecteurs dans une base quelconque. Soit δ un mineur d'ordre r non nul, extrait de A.*
Alors pour qu'un vecteur w appartienne à l'espace $Vect(v_1, \dots, v_r)$ il faut et il suffit que tous les bordants de δ dans la matrice $B = \| \; v_1, \dots, v_r, w \; \|$ soient nuls.

Démonstration Soient :

$$A = \left(\begin{array}{cccccc}
| & - & - & - & - & | \\
| & a_{11} & \dots & \dots & a_{1r} & | \\
| & \vdots & & & \vdots & | \\
| & a_{r1} & \dots & \dots & a_{rr} & | \\
| & - & - & - & - & | \\
& a_{r+1,1} & \dots & \dots & a_{r+1,r} & \\
& \vdots & & & \vdots & \\
& a_{n1} & & & a_{nn} &
\end{array}\right)
\qquad
B = \left(\begin{array}{ccccccc}
| & - & - & - & - & | & b_1 \\
| & a_{11} & \dots & \dots & a_{1r} & | & \vdots \\
| & \vdots & & & \vdots & | & \vdots \\
| & a_{r1} & \dots & \dots & a_{rr} & | & b_r \\
| & - & - & - & - & | & \\
& a_{r+1,1} & \dots & \dots & a_{r+1,r} & & b_{r+1} \\
& \vdots & & & \vdots & & \vdots \\
& a_{n1} & \dots & \dots & a_{nn} & & b_n
\end{array}\right)$$

Quitte à changer l'ordre des lignes et des colonnes, on peut supposer que le mineur δ non nul est le mineur encadré.
Supposons que $w \in Vect(v_1, \dots v_r)$. Si l'un des bordants était non nul, la famille $\{v_1, \dots, v_r, w\}$ serait libre d'après le lemme précédent, ce qui est exclu car w appartient à l'espace engendré par $v_1, \dots, v_r$. Donc tous les bordants de δ dans B sont nuls.
Réciproquement, supposons que tous les bordants de δ sont nuls. En regardant les vecteurs lignes de la matrice B, on voit que les r premiers sont indépendants, car $\delta \neq 0$ (d'après le lemme précédent) et chacun des autres est liés aux r premiers. Ainsi la matrice B a pour rang r ; Donc les vecteurs colonnes de B, $v_1, \dots, v_r, w$ forment une famille de rang r et comme $\{v_1, \dots, v_r\}$ est une famille libre, $w \in Vect(v_1, \dots, v_r)$.

Lemme 0.4 *Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. Alors le rang de A est r si et seulement si on peut extraire de A un mineur δ d'ordre r non nul tel que tous les bordants (ie mineurs d'ordre $r+1$) de δ dans A sont nuls.*

Démonstration Les conditions sont suffisantes. Soit, en effet :

$$A = \| v_1, \dots, v_n \| = \left(\begin{array}{ccccccc}
| & - & - & - & | & & \\
| & a_{11} & \dots & a_{1r} & | & \dots & a_{1n} \\
| & \vdots & & \vdots & | & & \vdots \\
| & a_{r1} & \dots & a_{rr} & | & \dots & a_{rn} \\
| & - & - & - & | & & \\
& \vdots & & \vdots & & & \vdots \\
& a_{p1} & \dots & a_{pr} & & \dots & a_{pn}
\end{array}\right)$$

et soit $\delta \neq 0$ le mineur encadré.
Comme $\delta \neq 0$, les vecteurs $\{v_1, \dots, v_r\}$ sont indépendants (d'après un lemme précédent). Puisque tous les bordants de δ sont nuls, chaque vecteur v_s ($s \geq r+1$) appartient à l'espace $Vect(v_1, \dots, v_r)$ d'après le lemme précédent. Donc les vecteurs colonnes engendrent un espace de dimension r et par conséquent $rg(A) = r$.
Réciproquement, soit $r = rg(A)$ et supposons quitte à changer l'ordre des colonnes que $\{v_1, \dots, v_r\}$ est une famille libre. On peut alors extraire de la matrice formée par les r premières colonnes de A, un mineur δ d'ordre r non nul (d'après un lemme précédent). Quitte à changer la numérotation des coordonnées, on peut supposer que δ est le mineur formé par les r premières lignes et les r premières colonnes de A. Or $rg(A) = r$, donc $v_s \in Vect(v_1, \dots, v_r)$ pour $s \geq r+1$; d'après le lemme précédent, tous les bordants de δ dans A sont nuls.

Références

[Gou08] Xavier Gourdon. *Analyse.* Ellipses, 2008.