Théorème des extrema liés

Références : Gourdon, *Les maths en tête - Analyse*, p 317 et p 327
Beck, Malick, Peyré, *Objectif agrégation*, p20 (pour l'interprétation géométrique)

Théorème (Théorème des extrema liés)**.**

Soient f, g_1, ..., g_r des fonctions de classe $\mathcal{C}^1$ d'un ouvert $U \subset \mathbb{R}^n$ dans $\mathbb{R}$. On pose $\Gamma = \{x \in U, g_1(x) = \cdots = g_r(x) = 0\}$.
Si $f_{|\Gamma}$ admet un extremum relatif en $a \in \Gamma$, et si les formes linéaires $Dg_1(a), \ldots, Dg_r(a)$ sont linéairement indépendantes, alors il existe des réels $\lambda_1, \cdots, \lambda_r$, appelés multiplicateurs de Lagrange, tels que

$$Df(a) = \sum_{i=1}^{r} \lambda_i Dg_i(a).$$

Démonstration. • Soit $s = n - r$. On peut identifier $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$. On écrit donc les éléments de $\mathbb{R}^n$ comme (x, y) où $x = (x_1, \ldots, x_s)$ et $y = (y_1, \ldots, y_r)$.
Posons $a = (\alpha, \beta) \in \mathbb{R}^s \times \mathbb{R}^r$.
Déjà, $r \leqslant n$ car les formes linéaires $Dg_i(a)$ forment une famille libre et la dimension de l'espace dual de $\mathbb{R}^n$ est n. De plus, si $r = n$, le théorème devient évident car les $Dg_i(a)$ forment une base de $(\mathbb{R}^n)^*$.
On peut donc supposer que $r < n$, ie $s \geqslant 1$.

• Les formes linéaires $(Dg_i(a))_{1 \leqslant i \leqslant r}$ forment une famille libre, ainsi la matrice :

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

est de rang r [1].
On peut donc en extraire une sous-matrice $r \times r$ inversible. Quitte a changer le nom des variables, on peut supposer que :

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

Ce qui peut se reformuler, en posant $g := (g_1, \ldots, g_r)$ par : $D_y g(a)$ est inversible.

• D'après le théorème des fonctions implicites, on peut donc trouver un voisinage ouvert U' de α dans $\mathbb{R}^s$, un voisinage ouvert Ω de $a = (\alpha, \beta)$ dans $\mathbb{R}^n$ et une fonction $\varphi = (\varphi_1, \ldots, \varphi_r) : U' \to \mathbb{R}^r$ de classe $\mathcal{C}^1$ tels que :

$$(x \in U', (x, y) \in \Omega \text{ et } g(x, y) = 0) \Leftrightarrow (y = \varphi(x)).$$

En d'autre termes, sur un voisinage de a, les éléments de Γ s'écrivent $(x, \varphi(x))$. Comme $a \in \Gamma$, on a $\beta = \varphi(\alpha)$.

Posons $\psi := (\psi_1, \ldots, \psi_n) : x \in U' \subset \mathbb{R}^s \mapsto (x, \varphi(x))$. Alors $\psi(x) \in \Gamma$ par le TFI. Posons également, sur U', $h = f \circ \psi$.
Comme $h(\alpha) = f(a)$, h admet un extremum local en α (car f admet un extremum local sur Γ en a).
Ainsi, pour tout $i \in [\![1, s]\!]$,

$$0 = \frac{\partial h}{\partial x_i}(\alpha) = \sum_{j=1}^{s} \frac{\partial f}{\partial x_j}(\psi(\alpha)) \frac{\partial \psi_j}{\partial x_i}(\alpha) + \sum_{j=1}^{r} \frac{\partial f}{\partial y_j}(\psi(\alpha)) \frac{\partial \psi_{s+j}}{\partial x_i}(\alpha)$$

1. Il y a r lignes donc le rang est au plus r. Puis les r lignes sont indépendantes donc le rang est au moins r.

En remarquant que $\forall j \in [\![1,s]\!]$, $\dfrac{\partial \psi_j}{\partial x_i} = \delta_{i,j}$ et que $\forall j \in [\![1,r]\!]$ $\dfrac{\partial \psi_{s+j}}{\partial x_i} = \dfrac{\partial \varphi_j}{\partial x_i}$. Et comme $a = \psi(\alpha)$, on obtient :

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

De plus, $g \circ \psi$ est nulle sur U' donc pour tout $k \in [\![1,r]\!]$ c'est également le cas pour $g_k \circ \psi$. Donc, par un calcul similaire à celui du dessus, pour $i \in [\![1,s]\!]$:

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

Si on considère donc la matrice :

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

Les s premiers vecteurs colonnes de M s'expriment, d'après les deux formules aux dérivées partielles ci-dessus, linéairement en fonction de ses r derniers vecteurs colonnes, donc $\mathrm{rg}(M) \leqslant r$. Ainsi les $r+1$ lignes de M forment une famille liée. [2]
Ceci entraine l'existence de réels $\mu_0, \dots, \mu_r$ non tous nuls tels que :

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

Comme la famille $(Dg_i(a))_i$ est libre, $\mu_0 \neq 0$ donc en posant $\lambda_i = -\dfrac{\mu_i}{\mu_0}$ pour $i \in [\![1,r]\!]$, on obtient

$$Df(a) = \sum_{i=1}^{r} \lambda_i Dg_i(a).$$

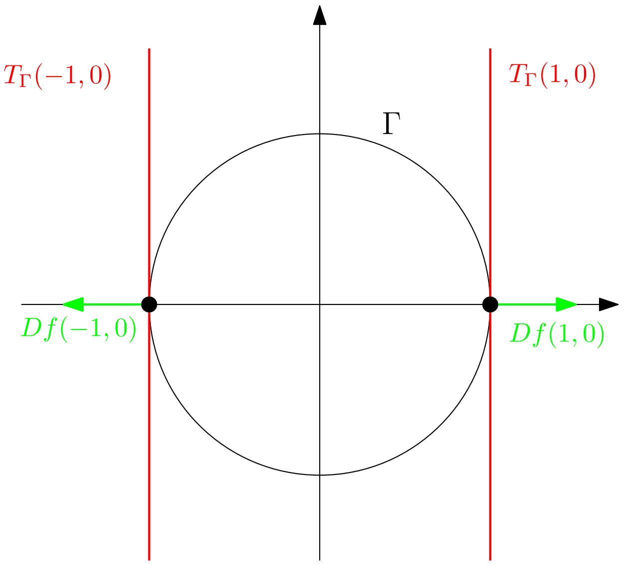
□

Interprétation géométrique du théorème : (d'après Objectif Agrégation p 20)
$\Gamma \cap V_a$ est une sous-variété pour V_a un certain voisinage de a.
Pour montrer cela, on commence par enlever des coordonnées comme dans la preuve pour obtenir la matrice (Dg_i) de taille $r \times r$ de déterminant non nul. Par continuité du déterminant, il existe un voisinage de a où le déterminant est non nul, donc où les Dg_i sont linéairement indépendants.
Puis on montre si les $Dg_i(x)$ sont linéairement indépendants, alors la matrice $DG(x)$ des $Dg_i(x)$ est surjective (donc on a la submersion voulue). En effet, $\mathrm{Ker}(DG(x)) = \bigcap \mathrm{Ker}(Dg_i(x)) = \bigcap (\nabla g_i(x))^\perp = \mathrm{Vect}(\nabla g_i(x))^\perp$ et par indépendance, $\mathrm{Vect}(\nabla g_i(x))$ est de dimension r, d'où $\mathrm{Vect}(\nabla g_i(x))^\perp$ est de dimension $n-r$. Donc la dimension de l'image est égale à r, ce qui montre la surjectivité.

L'égalité $Df(a) = \sum_{i=1}^{r} \lambda_i Dg_i(a)$ nous donne $\bigcap \mathrm{Ker}\, Dg_i(a) \subset \mathrm{Ker}\, Df(a)$. C'est même une équivalence. [3]
La caractérisation de l'espace tangent pour les sous variétés implicites est $\{h \in \mathbb{R}^n, Dg_i(a)h = 0, \forall i\}$. Le théorème des extrema liés donne donc que $Df(a)$ est nulle sur l'espace tangent à Γ en a. Cela s'écrit aussi $\forall h \in \Gamma, < Df(a), h >= 0$, donc le vecteur $Df(a)$ est orthogonal à l'espace tangent en a.

→ En exemple, on prend $f(x,y) = x^2$ sur $\Gamma = \{(x,y) \in \mathbb{R}^2, g(x,y) = x^2 + y^2 - 1 = 0\}$. Les points extrémaux sont $(\pm 1, 0)$ et $(0, \pm 1)$ et les vecteurs correspondants sont $Df(\pm 1, 0) = \begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}$ et le vecteur nul pour les deux autres. On a la figure suivante (où sont tracés les espaces tangents **affines**) qui représente l'orthogonalité de Df par rapport à l'espace tangent aux points où Df est non nulle.

2. Plus précisément, ça vient de $\mathrm{rg}({}^tM) = \mathrm{rg}(M)$, le rang des vecteurs lignes est égal au rang des vecteurs colonnes de M. Cette propriété vient d'elle même en faisant un pivot de Gauss pour se ramener à une matrice avec une diagonale de 1 et des 0 partout.
3. On le montre en utilisant qu'en dimension finie $(F^0)^\perp = F$, cf FGNal1 ex 6.27



En fait, l'exercice sur les directions principales d'une quadrique (Rouvière, ex 129) est la généralisation en dimension 3 de cette vision. On pose $f(x) = \|x\|^2$ et Q la forme quadratique définissant l'ellipsoïde. Alors les points extrémaux sont atteints quand le vecteur x est orthogonal à l'ellipsoïde. Un rapide dessin montre que cela arrive seulement pour les directions principales.

Remarques : • Il existe une version sous-variété de ce théorème présente dans le Avez de calcul différentiel. Elle trivialise la démonstration mais ne donne pas les multiplicateurs de Lagrange. Il y a du travail après.
• Une application du théorème des extrema liés est la preuve de l'inégalité arithmético-géométrique. (Gourdon, p319)
Idée : on pose $f(x_1, \dots, x_n) = x_1 \dots x_n$ et $g(x_1, \dots, x_n) = x_1 + \dots + x_n - s$. On applique le TEL et on trouve $\lambda = \dfrac{f(a)}{a_i}$, donc les a_i sont constants (à $\dfrac{s}{n}$) et $f(x) \leqslant \left(\dfrac{s}{n}\right)^n$ sur l'ensemble $\Gamma = \{x \in (\mathbb{R}^+)^n, \sum x_i = s\}$.
• Une autre application du même type est la mise en boîte à peu de frais de l'exercice 128 du Rouvière.
• Une dernière application est la preuve du théorème spectral. On a juste à maximiser $x \mapsto (u(x), x)$ sur la sphère unité.

Adapté du travail de Laura Gay