

THÉORÈME DES EXTREMA LIÉS ET APPLICATION

1 Résultat et idée de la preuve

THÉORÈME 1. [ROU, PAGE 372] (THÉORÈME DES EXTREMA LIÉS) Soient $U \subset \mathbb{R}^n$ un ouvert et $f, g_1, \dots, g_p : U \rightarrow \mathbb{R}$ des fonctions de classe C^1 . Soit

$$X = \{x \in U, \quad \forall 1 \leq i \leq p, \quad g_i(x) = 0\}.$$

Si $f|_X$ admet un extremum local en $a \in X$ et si les formes linéaires $d_{g_1}(a), \dots, d_{g_p}(a)$ sont indépendantes, alors $d_f(a)$ est combinaison linéaire des $d_{g_i}(a)$:

$$\exists ! \lambda_1, \dots, \lambda_p \in \mathbb{R}, \quad d_f(a) = \sum_{i=1}^p \lambda_i d_{g_i}(a).$$

Les λ_i sont appelés *multiplicateurs de Lagrange*.

Pour la preuve, on admet le lemme suivant :

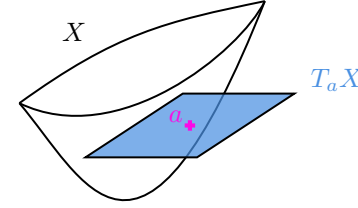
LEMME 2. $T_a X = \bigcap_{i=1}^p \ker(d_{g_i}(a)) = \ker d_g(a)$ en notant $g = (g_1, \dots, g_p) : U \rightarrow \mathbb{R}^p$.

Donnons d'abord l'idée de la preuve.

Comme les g_i sont C^1 et les $d_{g_i}(a)$ indépendantes, X est une sous-variété de $\mathbb{R}^n$ de dimension $n - p$ au point a . Notons $T_a X$ l'espace tangent à X en a : il s'agit d'un sous-espace vectoriel de dimension $n - p$.

Dire qu'une forme linéaire ℓ est combinaison linéaire de formes linéaires g_i revient à dire que $\bigcap_{i=1}^p \ker h_i \subset \ker \ell$. Dans notre cas, on a $h_i = d_{g_i}(a)$, $\bigcap_{i=1}^p \ker(d_{g_i}(a)) = T_a X$, et $\ell = d_f(a)$. On cherche donc à montrer que $T_a X \subset \ker d_f(a)$, c'est-à-dire que $d_f(a)$ s'annule sur $T_a X$.

Et c'est logique : localement, X ressemble à $\mathbb{R}^{n-p}$, donc on a une fonction $f : \mathbb{R}^{n-p} \rightarrow \mathbb{R}$. Cette fonction a un extremum local en a , donc $d_f(a) = 0$. Ainsi, pour tout $v \in \mathbb{R}^{n-p}$, $d_f(a)(v) = 0$. Mais en réalité, l'espace des v n'est rien d'autre que l'espace tangent !



2 $d_f(a)$ est nulle sur $T_a X$

Preuve. Soit γ un arc dérivable tracé sur X tel que $\gamma(0) = a$. Alors, la fonction $f \circ \gamma$ admet un extremum en 0, donc $0 = (f \circ \gamma)'(0) = d_f(a)(\gamma'(0))$. Ainsi, on a $d_f(a)(v) = 0$ pour tout vecteur $v \in T_a X$. La forme linéaire $d_f(a)$ est donc identiquement nulle sur $T_a X$. $\square$

3 Fin de la preuve

LEMME 3. [AVEZ, PAGE 103] Soient $\ell, h_1, \dots, h_p$ des formes linéaires sur un espace vectoriel E de dimension n . On suppose que les h_i sont linéairement indépendantes, et que

$$\bigcap_{i=1}^p \ker h_i \subset \ker \ell.$$

Alors, ℓ s'écrit comme combinaison linéaire des h_i .

Preuve. Complétons les h_i en une base $(h_1, \dots, h_n)$ de E^* , et soit $(e_1, \dots, e_n)$ sa base antéduale. Il existe des constantes c_i telles que $\ell = \sum_{i=1}^n c_i h_i$. Cependant, pour tout

$r \geq p + 1$ on a $e_r \in \bigcap_{i=1}^p \ker h_i \subset \ker \ell$, donc $0 = \ell(e_r) = c_r$. Ainsi, $\ell = \sum_{i=1}^p c_i h_i$. $\square$

On peut maintenant terminer la preuve du théorème. On sait que $d_f(a)$ est identi-

quement nulle sur $T_a X = \bigcap_{i=1}^p \ker d_{g_i}(a)$: autrement dit,

$$\bigcap_{i=1}^p \ker d_{g_i}(a) \subset \ker d_f(a).$$

Comme les $d_{g_i}(a)$ sont indépendantes, $d_f(a)$ est combinaison linéaire des $d_{g_i}(a)$ d'après le lemme précédent.

4 Application : directions principales d'une quadrique

Soit Q une forme quadratique définie positive sur $\mathbb{R}^n$, et soit A sa matrice dans la base canonique.

Considérons la quadrique

$$\begin{aligned} S &= \{x \in \mathbb{R}^n, \quad Q(x) = 1\} \\ &= \{x \in \mathbb{R}^n, \quad {}^t x A x = 1\}. \end{aligned}$$

On souhaite déterminer les extrema globaux de $f : x \mapsto \|x\|^2$ sur S .

Soit $a \in S$ un extremum de f sur S . On a $\nabla_Q(a) = 2Aa \neq 0$, donc le théorème des extrema liés s'applique, et fournit l'existence d'un $\mu \in \mathbb{R}$ tel que

$$d_f(a) = \mu d_Q(a).$$

Alors, $a = \lambda Aa$. En multipliant à gauche par ${}^t a$ et en utilisant le fait que $Q(a) = 1$, cela donne $\mu = \|a\|^2 > 0$. Finalement,

$$Aa = \frac{1}{\mu} a,$$

donc a est un vecteur propre de A de norme $\sqrt{\mu}$, associé à la valeur propre $\frac{1}{\mu}$.

Réciproquement, notons $0 < \lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres de A . D'après le théorème spectral, il existe une base orthonormée $(e_1, \dots, e_n)$ telle que, si $x =$

$$\sum_{i=1}^n x_i e_i,$$

$$Q(x) = \sum_{i=1}^n \lambda_i x_i^2,$$

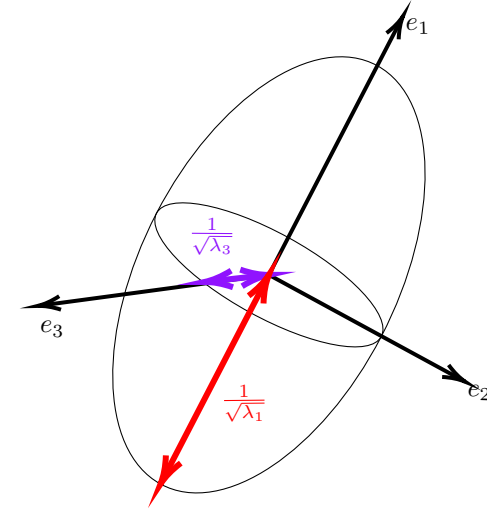
et $f(x) = \sum_{i=1}^n x_i^2$ car la base est orthonormée. On a alors, pour tout $x \in \mathbb{R}^n$,

$$\lambda_1 f(x) \leq Q(x) \leq \lambda_n f(x),$$

donc pour tout $a \in S$,

$$\frac{1}{\lambda_n} \leq f(a) \leq \frac{1}{\lambda_1}.$$

Chacune des inégalités est atteinte respectivement pour les vecteurs $\frac{e_n}{\sqrt{\lambda_n}}$ et $\frac{e_1}{\sqrt{\lambda_1}}$: on a donc déterminé les extrema globaux de f sur S . Pour les maxima, ce sont les $e/\sqrt{\lambda_1}$ avec e vecteur propre unitaire associé à λ_1 , idem pour les minima avec λ_n .



Remarques

- Convient pour les leçons 159, 206, 215, 219, 267.
- Peut (je dirais même doit) être énoncé dans la leçon 214.
- Il faut connaître un contre-exemple avec les gradients liés : cf. Beck, Malick, Peyré, *Objectif agrégation*.
- Le théorème des extrema liés a de nombreuses applications : directions principales d'une quadrique, mais aussi inégalité de Hadamard, théorème spectral, inégalité arithmético-géométrique... il y a l'embarras du choix, et il est bon de connaître ces résultats.
- Chercher les extrema de la norme euclidienne sur la quadrique S équivaut à chercher les extrema de $x \mapsto \frac{Q(x)}{\|x\|^2}$, donc les meilleures constantes d'équivalence pour les normes $\sqrt{Q}$ et $\|\cdot\|$.
- En fait, les extrema globaux de la norme euclidienne sur la quadrique sont les seuls extrema locaux : cf. [Rou] (si a est un éventuel candidat pour être un extremum local non global, en se plaçant dans les plans $\text{Vect}\{a, e_1\}$ et $\text{Vect}\{a, e_n\}$, on montre que a est un "point selle").

Preuve du lemme admis

REMARQUE 4. Cette preuve est souvent faite par inclusion et égalité des dimensions, mais je n'ai trouvé de preuve directe nulle part : en voici donc une ci-dessous.

- Montrons l'inclusion directe : soit $v \in T_a X$. Il existe une courbe γ dérivable tracée sur X telle que $\gamma(0) = a$ et $\gamma'(0) = v$. Pour tout $t \in I$, on a $\gamma(t) \in X$, donc

$$\forall t \in I, \quad \forall 1 \leq i \leq p, \quad g_i(\gamma(t)) = 0.$$

En différentiant ces égalités, on obtient

$$\forall t \in I, \quad \forall 1 \leq i \leq p, \quad d_{g_i}(\gamma(t))(\gamma'(t)) = 0.$$

Alors, $t = 0$ donne $v \in \bigcap_{i=1}^p \ker d_{g_i}(a)$, ce qui montre que $T_a X \subset \bigcap_{i=1}^p \ker d_{g_i}(a)$.

- Montrons l'inclusion réciproque. Soit $v \in \ker d_g(a)$: on cherche une courbe γ tracée sur X telle que $\gamma(0) = a$ et $\gamma'(0) = v$. Puisque les $d_{g_i}(a)$ sont libres, $d_g(a)$ est surjective (résultat classique de dualité), *i.e.* de rang p . On peut par exemple supposer, quitte à permuter les g_i , que les p dernières colonnes de sa jacobienne sont libres. Notons alors $x = (x_1, x_2) \in \mathbb{R}^{n-p} \times \mathbb{R}^p$: l'application

$$g : U \subset \mathbb{R}^{n-p} \times \mathbb{R}^p \rightarrow \mathbb{R}^p$$

est de classe C^1 , vérifie $g(a) = 0$ et $D_2 g(a)$ est inversible par hypothèse. Par le théorème des fonctions implicites, il existe V, W voisinages respectifs de a_1 dans $\mathbb{R}^{n-p}$ et de a_2 dans $\mathbb{R}^p$, et

$$\varphi : V \rightarrow W \text{ de classe } C^1$$

tels que

$$\forall (x, y) \in V \times W, (x, y) \in X \iff g(x, y) = 0 \iff y = \varphi(x).$$

On peut alors chercher γ sous la forme

$$t \mapsto (x(t), \underbrace{\varphi(x(t))}_{=y(t)})$$

(si $\gamma(0) = a$, alors par continuité, pour t assez petit, on aura $\gamma(t) \in V \times W$). On veut $x(0) = a_1$, $x'(0) = v_1$: $x(t) = a_1 + v_1 t$ convient !

On a $y(0) = \varphi(a_1) = a_2$, et il reste à voir que $y'(0) = v_2$. Pour ce faire, on utilise l'hypothèse $d_g(a)(v) = 0$. Cela donne

$$D_1 g(a)(v_1) + D_2 g(a)(v_2) = 0,$$

puis comme $D_2 g(a)$ est inversible,

$$v_2 = -D_2 g(a)^{-1} \circ D_1 g(a)(v_1).$$

Or,

1. d'une part, en différentiant l'expression $g(x, \varphi(x)) = 0$ et en évaluant en $x = a_1$, on montre que

$$d_\varphi(a_1) = -D_2 g(a)^{-1} \circ D_1 g(a).$$

En évaluant en v_1 , cela donne

$$d_\varphi(a_1)(v_1) = -D_2 g(a)^{-1} \circ D_1 g(a)(v_1).$$

2. D'autre part, en dérivant l'expression $\varphi(x(t)) = y(t)$ et en évaluant en $t = 0$, on obtient

$$y'(0) = d_\varphi(a_1)(v_1).$$

En définitive, $v_2 = y'(0)$.

Références

[Rou] François ROUVIÈRE, *Petit guide de calcul différentiel*, Cassini, 2009.

[Avez] André AVEZ, *Calcul différentiel*, Masson, 1983.