# Théorème des extrema liés

Pierron Théo

9 juin 2014

## 1 Version pour dimension et TFI/TIL

<u>Théorème</u> *Soit $U$ un ouvert de $\mathbb{R}^n$, $f, g_1, \dots, g_r : U \to \mathbb{R}$ de classe $C^1$.*

*Soit $\Gamma = \{x \in U, \forall i \in [\![1, r]\!], g_i(x) = 0\}$. Si $f|_\Gamma$ admet un extremum relatif en $a \in \Gamma$ et si $D_a g_1, \dots, D_a g_r$ sont des formes linéaires indépendantes, alors il existe $\lambda_1, \dots, \lambda_r$ des réels tels que*

$$D_a f = \sum_{i=1}^{r} \lambda_i D_a g_i$$

*Démonstration.* On identifie $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$ avec $s = n - r$. Notons $a = (\alpha, \beta)$.

Comme la famille $(D_a g_i)_i$ est libre, 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 $r$, donc il existe une sous-matrice $r \times r$ inversible, c'est-à-dire que quitte à permuter les coordonnées, $\det(\frac{\partial g_i}{\partial y_j}(a))_{1 \leqslant i,j \leqslant r} \neq 0$.

Utilisons maintenant le théorème des fonctions implicites pour $g = (g_1, \dots, g_r)$ au voisinage de $\alpha$. Il existe un voisinage $V$ de $\alpha$ dans $\mathbb{R}^s$ et un voisinage $W$ de $\beta$ dans $\mathbb{R}^r$ tel que $V \times W \subset U$ et $\varphi_i \in C^1(V, W)$ tel que

$$(x, y) \in V \times W \text{ et } g_i(x, y) = 0 \quad \text{ssi} \quad x \in V \text{ et } y = \varphi_i(x)$$

Ainsi, au voisinage de $a$, les éléments de $\Gamma$ s'écrivent comme des $(x, \varphi_1(x), \dots, \varphi_r(x))$ pour un certain $x \in V$.

Posons $h(x) = f(x, \varphi(x))$ où $\varphi = (\varphi_1, \dots, \varphi_r)$. $h$ admet un extremum local en $\alpha$ par hypothèse donc

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

En différentiant la relation $g_k(x, \varphi(x)) = 0$ par rapport à $x_i$, on trouve que

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

On a donc montré que les $s$ premiers vecteurs colonnes de la matrice

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

sont combinaison linéaire des $r$ derniers, donc $\mathrm{rg}(M) \leqslant r$.

Par définition du rang, les vecteurs ligne de $M$ forment une famille de rang au plus $r$. Comme il y en a $r+1$, ils forment une famille liée. Il existe donc $\mu_0, \ldots, \mu_r$ non tous nuls tels que

$$\mu_0 D_a f + \sum_{i=1}^{r} \mu_i D_a g_i = 0$$

Comme la famille $(D_a g_i)_i$ est libre, $\mu_0 \neq 0$, et en prenant $\lambda_i = -\frac{\mu_i}{\mu_0}$, on retrouve bien

$$D_a f = \sum_{i=1}^{r} \lambda_i D_a g_i \qquad \blacksquare$$

*Remarque* *Géométriquement la relation* $D_a f = \sum_{i=1}^{r} \lambda_i D_a g_i$ *signifie que*

$$\bigcap_{i=1}^{r} \mathrm{Ker}(D_a g_i) \subset \mathrm{Ker}(D_a f)$$

*Donc* $D_a f$ *est nulle sur l'intersection des noyaux des formes linéaires* $D_a g_i$*.*

*Or le plan tangent* $a$ *à la sous-variété* $\Gamma$ *est*

$$T_a\Gamma = \{h \in \mathbb{R}^n, \forall i \in [\![1, r]\!], D_a g_i(h) = 0\}$$

*Autrement dit,* $D_a f$ *est nulle sur* $T_a\Gamma$*.*

**Exemple** Soit $E$ un espace vectoriel de dimension finie muni d'une structure euclidienne et $u \in L(E)$ symétrique.

Soit $f : x \mapsto \langle u(x), x \rangle$ et $g : x \mapsto \|x\|^2$.

La sphère unité $S$ de $E$ est compacte donc $f|_S$ admet un maximum en $x_0$. Comme $S = \{x \in E, g(x) = 1\}$, on a, par le théorème des extrema liés, l'existence d'un $\lambda \neq 0$ tel que $D_{x_0} f = \lambda D_{x_0} g$.

Or, comme $u$ est symétrique,

$$f(x+h) = \langle u(x+h), x+h \rangle = f(x) + \langle u(h), x \rangle + \langle u(x), h \rangle + o(\|h\|) = f(x) + 2\langle u(x), h \rangle + o(\|h\|)$$

et $D_x f(h) = 2\langle u(x), h \rangle$. Pour $u = \mathrm{Id}$, on retrouve l'expression de $D_x g(h) = 2\langle x, h \rangle$.

Ainsi, pour tout $h \in E$, $2\langle u(x_0), h \rangle = 2\lambda \langle x_0, h \rangle$ donc $u(x_0) = \lambda x_0$.

Autrement dit, tout endomorphisme symétrique admet un vecteur propre, ce qui constitue un premier pas vers la diagonalisation.

## 2 Version sympathique

**Lemme**
Soient $\varphi, \varphi_1, \ldots, \varphi_r$ des formes linéaires. Alors

$$\varphi \in \mathrm{Vect}\{\varphi_1, \ldots, \varphi_r\} \quad \text{ssi} \quad \bigcap_{i=1}^{r} \mathrm{Ker}\, \varphi_i \subset \mathrm{Ker}\, \varphi$$

*Démonstration.* Le sens $\Rightarrow$ est clair. Pour obtenir l'autre sens on dualise :

$$\sum_{i=1}^{r} \mathrm{Ker}(\varphi_i)^\perp \supset \mathrm{Ker}(\varphi)^\perp$$

On constate alors que $\mathrm{Ker}(\varphi_i)^\perp$ est la droite $\mathrm{Vect}\left\{\varphi_i\right\}$ (deux formes linéaires de même noyau sont colinéaires). Donc finalement

$$\mathrm{Vect}\left\{\varphi\right\} \subset \sum_{i=1}^{r} \mathrm{Vect}\left\{\varphi_i\right\}$$

ce qui conclut. ■

<u>THÉORÈME 1</u> *Soit $U$ un ouvert de $\mathbb{R}^n$, $f, g_1, \ldots, g_r : U \to \mathbb{R}$ de classe $C^1$.*

*Soit $\Gamma = \{x \in U, \forall i \in [\![1, r]\!], g_i(x) = 0\}$. Si $f|_\Gamma$ admet un extremum relatif en $a \in \Gamma$ et si $D_a g_1, \ldots, D_a g_r$ sont des formes linéaires indépendantes, alors il existe $\lambda_1, \ldots, \lambda_r$ des réels tels que*

$$D_a f = \sum_{i=1}^{r} \lambda_i D_a g_i$$

*Démonstration.* Par hypothèse, $\Gamma$ est lisse en $a$. Soit $v \in T_a\Gamma$. Il existe alors une courbe $\gamma : I \to \Gamma$ telle que $\gamma(0) = a$ et $\gamma'(0) = v$. On différentie $f \circ \gamma$ en 0 :

$$D_0(f \circ \gamma) = D_{\gamma(0)}(f)(\gamma'(0)) = D_a f(v)$$

$f|_\Gamma$ admet un extremum en $a$ donc $f \circ \gamma$ admet un extremum en 0 donc $D_a f(v) = 0$.

Ainsi, $D_a f$ est nulle sur $T_A\Gamma = \bigcap_{i=1}^{r} \mathrm{Ker}(D_a g_i)$. On utilise alors le lemme pour conclure. ■

<u>THÉORÈME 2</u> *Pour toute famille de vecteurs de $\mathbb{R}^n$ $(v_1, \ldots, v_n)$, on a*

$$|\det(v_1, \ldots, v_n)| \leqslant \prod_{i=1}^{n} \|v_i\|$$

*Démonstration.* Cherchons à maximiser det sur $S := \{(v_1, \ldots, v_n), \|v_i\|^2 = 1\}$.

Soit $(v_1, \ldots, v_n)$ un tel extremum. Il est clair que $(v_1, \ldots, v_n)$ doit être libre (sinon le déterminant est nul et ça ne peut être un extremum vu qu'on a des familles de déterminant 1 et $-1$). Par les extrema liés, il existe $\lambda_1, \ldots, \lambda_n$ tel que

$$\sum_{i=1}^{n} \det(v_1, \ldots, v_{i-1}, h_i, v_{i+1}, \ldots, v_n) = \sum_{i=1}^{n} 2\lambda_i \langle v_i, h_i\rangle$$

Posons pour tout $i, j$, $H_{i,j} = (h_1, \ldots, h_n) = (0, \ldots, 0, v_j, 0, \ldots, 0)$ où $v_j$ est à la position $i$. Alors en évaluant en $h_{i,i}$ la formule précédente, on obtient

$$\det v = 2\lambda_i \|v_i\|^2 = 2\lambda_i$$

donc $\lambda_i = \frac{\det v}{2}$. On évalue ensuite en $h_{i,j}$ pour $i \neq j$ :

$$0 = 2\lambda_i \langle v_i, v_j\rangle$$

Donc, comme $\lambda_i \neq 0$, on obtient que $(v_1, \ldots, v_n)$ est une base orthonormale.

Alors $\det(v) = \pm 1$. Ainsi, le minimum de det sur $S$ est $-1$ et son maximum est 1 donc pour tout famille $(u_1, \ldots, u_n)$ de vecteurs unitaires, on a

$$|\det(u_1, \ldots, u_n)| \leqslant 1$$

Soit $u_1, \ldots, u_n$ une famille quelconque de vecteurs de $\mathbb{R}^n$. Si un des $u_i$ est nul, on a évidemment le résultat. Sinon, on a

$$\left|\det\left(\frac{u_1}{\|u_1\|}, \ldots, \frac{u_n}{\|u_n\|}\right)\right| \leqslant 1$$

Et youpi le résultat tombe en sortant les normes. ■