CM-S4 : Extrema et extrema liés

Vincent Borrelli

Université de Lyon

Extremum

Soit $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ une fonction. On rappelle que f admet en $a \in \mathcal{U}$

- un MAXIMUM LOCAL s'il existe un voisinage $V(a)$ tel que $\forall x \in V(a), \quad f(x) \leq f(a)$.
- un MINIMUM LOCAL s'il existe un voisinage $V(a)$ tel que $\forall x \in V(a), \quad f(x) \geq f(a)$.
- un EXTREMUM LOCAL si f admet en ce point un maximum ou un minimum local.
- Dans ces trois définitions, on rajoute le terme STRICT quand le cas d'égalité n'a lieu que pour $x = a$ (dans le voisinage $V(a)$).

Extremum

• Si f est C^1, sa différentielle permet de définir le champ de vecteurs GRADIENT au moyen de la relation

$$\langle \mathrm{grad}\, f(x), V \rangle = df_x(V).$$

On note aussi $(\nabla f)(x)$ pour grad $f(x)$.

Définition.– On dit que a est un POINT CRITIQUE de f si $df_a = 0$ (ou grad $f(a) = 0$). L'ensemble de tous les points critiques de f est appelé le LIEU CRITIQUE et noté

$$Crit(f) := \{a \in \mathcal{U} \mid df_a = 0\}.$$

Théorème.– *Si a est un extremum local de f alors a est un point critique de f.*

Extremum

Démonstration.– Supposons que a soit un extremum local et que a ne soit pas un point critique. Alors, il existe V tel que $df_a(V) \neq 0$, par exemple $df_a(V) > 0$. Puisque

$$\lim_{h \to 0} \frac{\|f(a+hV) - f(a) - df_a(hV)\|}{\|hV\|} = 0$$

on en déduit

$$\lim_{h \to 0} \frac{f(a+hV) - f(a)}{h} = df_a(V) > 0$$

Donc en prenant h suffisamment petit, on obtient

$$f(a+hv) \geq f(a) \quad \text{si} \quad h > 0$$

et

$$f(a+hv) \leq f(a) \quad \text{si} \quad h < 0.$$

Contradiction. $\square$

• Pour connaître f au voisinage d'un point critique, on étudie son développement limité à l'ordre 2.

Différentielle seconde

Rappel sur la différentielle seconde.– Soit

$$\begin{array}{rccc} df : & \mathcal{U} & \longrightarrow & \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \\ & x & \longmapsto & df_x \end{array}$$

la différentielle première de f. Si l'application df est différentiable, on note

$$d^2 f : \quad \mathcal{U} \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$$

sa différentielle et on l'appelle la DIFFÉRENTIELLE SECONDE. Si cette différentielle $d^2 f$ est continue, on dit que f est de classe C^2.

- Soit $(V, W) \in \mathbb{R}^n \times \mathbb{R}^n$, par définition

$$d^2 f_x(V) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \quad \text{et} \quad d^2 f_x(V)(W) \in \mathbb{R}.$$

Différentielle seconde

• L'application

$$(V, W) \longmapsto d^2 f_x(V)(W)$$

étant bilinéaire, on peut voir $d^2 f$ comme étant à valeurs dans les formes bilinéaires

$$d^2 f : \mathcal{U} \longrightarrow \mathcal{L}_2(\mathbb{R}^n, \mathbb{R}) \simeq \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})).$$

• Le *théorème de Schwarz* affirme que pour une fonction de classe C^2 l'ordre dans lequel on écrit les vecteurs V et W n'importe pas :

$$\forall (V, W) \in \mathbb{R}^n \times \mathbb{R}^n, \quad d^2 f_x(V)(W) = d^2 f_x(W)(V).$$

• On peut donc voir la différentielle seconde comme une application à valeurs dans les formes bilinéaires symétriques

$$d^2 f : \mathcal{U} \longrightarrow S_2(\mathbb{R}^n).$$

Différentielle seconde

• Une base étant donnée, la matrice carrée symétrique représentant la forme bilinéaire symétrique $d^2f(x)$ dans cette base est appelée la MATRICE HESSIENNE de f en x et notée $\operatorname{hess} f(x)$.

• Dans la base standard de $\mathbb{R}^n$, cette matrice pend la forme

$$\operatorname{hess} f(x) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

Exercice.– Montrer que la matrice Jacobienne du gradient $\operatorname{grad} f : \mathcal{U} \to \mathbb{R}^n$ est la matrice hessienne de f.

Différentielle seconde

Définition.– On dit qu'un point critique a est NON DÉGÉNÉRÉ si le rang de la hessienne est maximum i.e.

$$\text{dét hess} f(a) \neq 0.$$

Proposition.– *Si a est un point critique non dégénéré alors il est isolé dans le lieu critique : il existe une boule ouverte centrée en a dans laquelle a est le seul point critique.*

Démonstration.– On doit résoudre l'équation

$$df_x = 0$$

au voisinage de a. Pour cela on applique le théorème d'inversion locale à l'application

$$df : \mathcal{U} \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}).$$

Différentielle seconde

• Montrons que l'hypothèse du TIL est vérifiée à savoir que $d(df)_a$ est inversible. Nous avons

$$d(df)_a := d^2f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})).$$

Or, au travers de l'isomorphisme

$$\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) \simeq \mathcal{L}_2(\mathbb{R}^n, \mathbb{R})$$

et du choix d'une base, $d^2f(a)$ s'identifie à la matrice hessienne hess $f(a)$. Le point critique étant non dégénérée, cette dernière est inversible. Donc $d(df)_a$ est inversible.

• Le TIL affirme qu'il en est alors de même pour df sur un voisinage $V(a)$. Ceci montre que a est le seul point critique de f dans ce voisinage. $\square$

Différentielle seconde

Définition.– On appelle INDICE DU POINT CRITIQUE *a* l'indice de sa hessienne (=le nombre de valeurs propres strictement négatives).

Proposition.– *Soient $\phi : \mathcal{U}' \to \mathcal{U}$ un difféomorphisme de classe C^2 et $a' \in \mathcal{U}'$, $a \in \mathcal{U}$ tels que $a = \phi(a')$. Alors a' est un point critique de $g = f \circ \phi$ ssi a est un point critique de f. De plus, a' est non dégénéré ssi a est non dégénéré et les indices de a et de a' coïncident.*

Différentielle seconde

• La preuve de la proposition nécessite de connaître la règle de la chaîne pour la différentielle seconde.

Règle de la chaîne pour la différentielle seconde.– *Soit $\phi : \mathcal{U}' \to \mathcal{U}$ un difféomorphisme de classe C^2 et $g = f \circ \phi$ alors pour tout $(V, W) \in \mathbb{R}^n \times \mathbb{R}^n$ on a*

$$d^2 g_{x'}(V, W) = d^2 f_x(d\phi_{x'}(V), d\phi_{x'}(W)) + df_x(d^2\phi_{x'}(V, W))$$

où $x' \in \mathcal{U}'$ et $x = \phi(x')$.

• La démonstration de cette formule est laissée en exercice.

Différentielle seconde

Démonstration de la proposition.– La règle de la chaîne pour la différentielle première s'écrit

$$dg_{a'} = df_a \circ d\phi_{a'}$$

et puisque ϕ est un difféomorphisme, $d\phi_{a'}$ est un isomorphisme linéaire ainsi

$$dg_{a'} = 0 \quad \Longleftrightarrow \quad df_a = 0.$$

• En un point critique, la règle de la chaîne pour la différentielle seconde s'écrit

$$d^2g_{a'}(V, W) = d^2f_a(d\phi_{a'}(V), d\phi_{a'}(W))$$

car $df_a = 0$.

Différentielle seconde

Notons P la matrice de $d\phi_{a'}$ dans la base standard de $\mathbb{R}^n$. L'égalité précédente, valable pour tout $(V, W) \in \mathbb{R}^n \times \mathbb{R}$ implique donc que les deux matrices hessiennes sont congruentes i.e.

$$\text{hess}\, g(a') = {}^t P \,\text{hess}\, f(a) P$$

et en particulier, elles ont le même rang. Ceci prouve que a' est non dégénéré ssi a est non dégénéré.

• Rappelons la *loi d'inertie de Sylvester* : Si A est une matrice symétrique réelle, alors il existe une matrice inversible Q telle que

$${}^t QAQ = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

où les entiers p et q ne dépendent que de A et pas de la matrice Q choisie pour diagonaliser A.

Différentielle seconde

- Si Q est telle que

$$^{t}Q\,\mathrm{hess}\,g(a')Q = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

alors PQ est telle que

$$^{t}(PQ)\,\mathrm{hess}\,f(a)PQ = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

ce qui montre que les points critiques a et a' ont le même indice. $\square$

Différentielle seconde

Théorème.– *Soit a un point critique non dégénéré.*

- *Si l'indice de f en a est nul alors a est un minimum local strict.*
- *Si l'indice de f en a est égal à n alors a est un maximum local strict.*
- *Dans les autres cas, a n'est pas un extremum local.*

Démonstration.– Écrire le développement limité à l'ordre 2 de f en a. □

Définition.– Un point critique non dégénéré dont l'indice n'est ni égal à 0, ni égal à n est appelé POINT COL ou POINT SELLE.

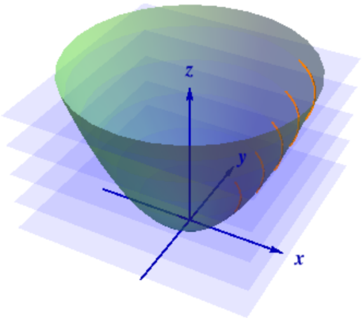
Le lemme de Morse

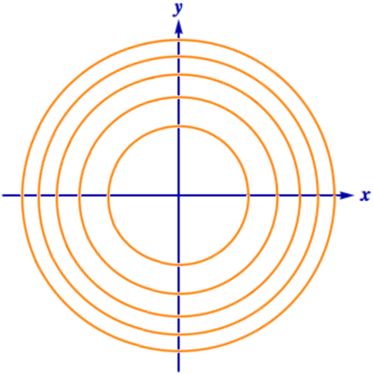
Lemme de Morse (admis).– *Soit* $f : V(0) \subset \mathbb{R}^n \to \mathbb{R}$ *une fonction* C^k, $k \geq 3$. *On suppose que* $0 \in \mathbb{R}^n$ *est un point critique non dégénéré alors il existe un difféomorphisme* $\phi : \mathcal{U}' \to \mathcal{U}$ *de classe* C^{k-2} *tel que pour tout* $x \in \mathcal{U}'$

$$f \circ \phi(x) = f(0) + \frac{1}{2} d^2 f_0(x, x).$$

• Ce lemme indique que dans des coordonnées convenables, les lignes de niveaux de f sont celles d'une forme quadratique. Il en dit donc beaucoup plus sur l'allure locale de f au voisinage d'un point critique non dégénéré que le théorème précédent.

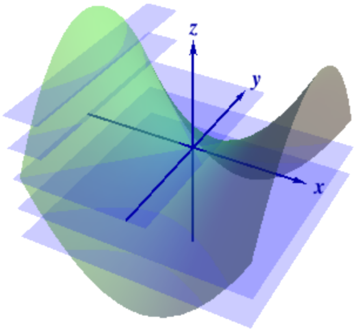
Point critique d'indice 0, $n = 2$

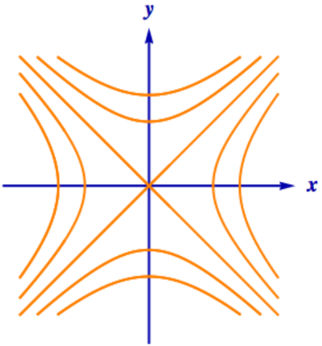




Graphe de f au voisinage d'un point critique d'indice 0

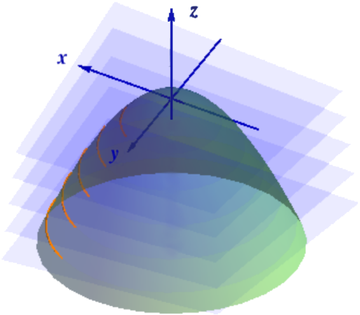
Point critique d'indice 1, $n = 2$

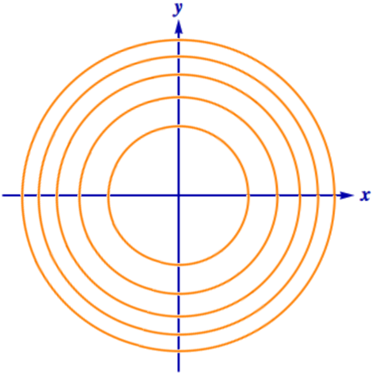




Graphe de f au voisinage d'un point critique d'indice 1

Point critique d'indice 2, $n = 2$





Graphe de f au voisinage d'un point critique d'indice 2

Le lemme de Morse

• L'accent mis sur le point 0 est purement conventionnel. Le lemme fonctionne évidemment en n'importe quel point critique non dégénéré.

• Le coefficient $\frac{1}{2}$ qui apparaît devant la dérivée seconde est tout aussi conventionnel. On peut choisir à sa place n'importe quel nombre non nul.

• Ce lemme joue un rôle crucial en géométrie. Il est la pierre fondamental de la THÉORIE DE MORSE qui permet de comprendre la topologie des sous-variétés/variétés en analysant les lignes de niveaux d'une fonction définie sur cette sous-variété.

Marston Morse (1892-1977)

Marston Morse en 1925 (33 ans)

Marston Morse (1892-1977)

• Mathématicien américain connu pour avoir bâti la théorie qui porte son nom et qui est depuis l'une des branches majeures de la *topologie différentielle*.

• Il intègre l'IAS en 1935 quelques années après la formation de cet institut. Il y côtoie -entre autres - Albert Einstein, Robert Oppenheimer, Hermann Weyl, John von Neumann et Kurt Gödel.

• Francophile, il rentre dans le corps expéditionnaire américain en 1914. Il est intégré dans le service ambulancier et décoré de la Croix de guerre pour « bravoure au feu ».

• Il est fait Chevalier de la Légion d'Honneur en 1952 et devient membre correspondant de l'Académie Française des Sciences en 1956.

Marston Morse (1892-1977)

- Il donne également son nom à la *suite de Thue–Morse* :

$$T_0 = 0,\ T_1 = 01,\ T_2 = 0110,\ T_3 = 01101001,$$

$$T_4 = 0110100110010110, \ldots$$

Cette suite peut aussi être définie par l'égalité

$$\prod_{i=1}^{\infty}(1 - x^{2^i}) = \sum_{j=0}^{\infty}(-1)^{t_j} x^j$$

où t_j est le $(j+1)$-ème chiffre de T_j. La suite de Thue-Morse a été découverte indépendamment par plusieurs personnes dont un champion d'échec, Max Euwe. Le premier a l'avoir étudiée est Eugène Prouhet en 1851 (donc 41 ans avant la naissance de Marston Morse).

Marston Morse (1892-1977)

• Il ne donne pas son nom au fameux *théorème de Morse-Sard*. Il s'agit d'un homonyme : Anthony Morse.

• Très croyant, il représente le Vatican dans la conférence *Atoms for Peace* organisée par les Nations Unies en 1952.

• Il divorce de sa première femme Céleste Phelps en 1930. Celle-ci se remarie avec un de ses collègues 28 ans plus âgé que lui : William Osgood.

• L'affaire fait grand bruit dans la communauté mathématique américaine et se transforme en scandale. William Osgood est conduit à démissionner. Il s'exile à Pékin où il poursuivra une certaine activité mathématique.

Marston Morse par Raoul Bott

James Alexander, Albert Einstein, Frank Aydelotte (dir. IAS), Oswald Veblen et Marston Morse

« I suppose that a young prospective knight approaching King Arthur's table for the first time must have felt as I did »

Raoul Bott (23 ans à l'époque), lors de sa première rencontre avec Marston Morse et les habitants de l'IAS.

IAS vs Kaamelott

Une citation de Marston Morse

« *Forever the foundation and never the Cathedral* »

Marston Morse.

Extrema liés

• On se donne une fonction $g : U \subset \mathbb{R}^n \to \mathbb{R}$ et une application $h : U \to \mathbb{R}^q$. Le lieu d'annulation de cette application définit un sous-ensemble $S = h^{-1}(0)$ de $\mathbb{R}^n$ appelé la *contrainte* et on cherche les extrema de g sous cette contrainte, c'est-à-dire les extrema de $g_{|S}$.

• Nous allons voir que sous des hypothèses de régularités raisonnables et à condition que les formes linéaires

$$(dh_1)_x, ..., (dh_q)_x$$

soient linéairement indépendantes en tout point $x \in S$ alors le problème peut se traiter comme une recherche d'extremum au sens précédent (la contrainte devient transparente).

Multiplicateurs de Lagrange

Proposition.– *Soient $\ell_1, ..., \ell_q$ des formes linéaires sur $\mathbb{R}^n$ et*

$$E = \bigcap_{i=1}^{q} \ker \ell_i.$$

Si $\Lambda \in (\mathbb{R}^n)^*$ est une forme linéaire telle que $E \subset \ker \Lambda$ alors il existe des scalaires $\lambda_1, ..., \lambda_q$ tels que

$$\Lambda = \sum_{i=1}^{q} \lambda_i \ell_i.$$

Démonstration.– Notons $L = Vect(\ell_1, ..., \ell_q) \subset (\mathbb{R}^n)^*$. Quitte à permuter les indices, on peut supposer que $(\ell_1, ..., \ell_r)$ avec $r \leq q$ est une base de L.

Multiplicateurs de Lagrange

• On introduit $n-r$ formes linéaires $\tau_{r+1},\ldots,\tau_n$ pour la compléter en une base de $(\mathbb{R}^n)^*$. La forme linéaire Λ s'écrit dans cette base

$$\Lambda=\sum_{i}^{r}\lambda_i\ell_i+\sum_{i=r+1}^{n}\mu_i\tau_i$$

où $\lambda_1,\ldots,\lambda_r,\mu_{r+1},\ldots,\mu_n$ sont des scalaires.

• Soient $(e_1,\ldots,e_n)$ la base duale. Puisque pour tout $i\in\{1,\ldots,r\}$ on a

$$\ell_i(e_j)=0\quad\text{si}\quad j\in\{r+1,\ldots,n\}.$$

On en déduit que

$$Vect(e_{r+1},\ldots,e_n)\subset E=\bigcap_{i=1}^{q}\ker\,\ell_i.$$

Multiplicateurs de Lagrange

• Puisque Λ s'annule sur E, on doit donc avoir

$$\Lambda(e_j) = 0 \quad \text{pour tout} \quad j \in \{r+1, ..., n\}.$$

• Or $\Lambda(e_j) = \mu_j$ par définition de la base duale. Tous les scalaires μ_j doivent être nuls. Au final

$$\Lambda = \sum_i^r \lambda_i \ell_i.$$

$\square$

Définition.– Les scalaires $\lambda_1, ..., \lambda_q$ s'appellent des MULTIPLICATEURS DE LAGRANGE.

Multiplicateurs de Lagrange

Théorème des multiplicateurs de Lagrange.– *Soient* $g : U \subset \mathbb{R}^n \to \mathbb{R}$ *et* $h : U \to \mathbb{R}^q$ *deux applications de classe* C^1 *et* $S = h^{-1}(0)$*. Si* $g_{|S}$ *admet un extremum en* $a \in S$ *et si les formes linéaires*

$$(dh_1)_a, \ldots, (dh_q)_a$$

sont linéairement indépendantes alors il existe des scalaires $\lambda_1, \ldots, \lambda_q$ *tels que*

$$dg_a = \sum_{i=1}^{q} \lambda_i (dh_i)_a.$$

Démonstration.– On écrit $\mathbb{R}^n$ comme le produit cartésien $\mathbb{R}^{n-q} \times \mathbb{R}^q$ et chaque élément de $\mathbb{R}^n$ comme un couple (x, y) avec $x \in \mathbb{R}^{n-q}$ et $y \in \mathbb{R}^q$. Ainsi $a = (a_x, a_y) \in \mathbb{R}^{n-q} \times \mathbb{R}^q$.

Multiplicateurs de Lagrange

• Puisque

$$S = \{(x, y) \in \mathbb{R}^{n-q} \times \mathbb{R}^q \mid h(x, y) = 0\}$$

on invoque le théorème de la fonction implicite pour affirmer l'existence d'une application $\varphi : V(a_x) \to V(a_y)$ telle que

$$\forall (x, y) \in V(a_x) \times V(a_y), \quad h(x, y) = 0 \iff y = \varphi(x).$$

• Pour cela, il nous faut vérifier que

$$\frac{\partial h}{\partial y}(a) = \begin{pmatrix} \frac{\partial h_1}{\partial y_1}(a) & \cdots & \frac{\partial h_1}{\partial y_q}(a) \\ \vdots & & \vdots \\ \frac{\partial h_q}{\partial y_1}(a) & \cdots & \frac{\partial h_q}{\partial y_q}(a) \end{pmatrix}$$

est inversible.

Multiplicateurs de Lagrange

• C'est ici que l'hypothèse d'indépendance linéaire des $(dh_i)_a$ rentre en jeu. Celle-ci signifie que la matrice

$$\begin{pmatrix} \dfrac{\partial h_1}{\partial x_1}(a) & \cdots & \dfrac{\partial h_1}{\partial x_{n-q}}(a) & \dfrac{\partial h_1}{\partial y_1}(a) & \cdots & \dfrac{\partial h_1}{\partial y_q}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \dfrac{\partial h_q}{\partial x_1}(a) & \cdots & \dfrac{\partial h_q}{\partial x_{n-q}}(a) & \dfrac{\partial h_q}{\partial y_1}(a) & \cdots & \dfrac{\partial h_q}{\partial y_q}(a) \end{pmatrix}$$

est de rang q. On peut donc extraire une matrice carré $q \times q$ qui est inversible. Quitte à effectuer une permutation des indices, on peut toujours supposer que la matrice constituée des q dernières colonnes convient. On montre ainsi l'inversibilité de

$$\frac{\partial h}{\partial y}(a).$$

Multiplicateurs de Lagrange

• Par définition de φ on a, pour tout $x \in V(a_x)$,

$$h(x, \varphi(x)) = 0$$

ou encore, en posant $\Phi(x) = (x, \varphi(x))$,

$$(h \circ \Phi)(x) = 0.$$

• La règle de la chaîne impose que pour tout $x \in V(a_x)$,

$$\mathrm{Im}\, d\Phi_x \subset \ker dh_{\Phi(x)}.$$

• Le rang de $d\Phi_x$ est $n - q$ parce que le premier facteur de Φ est l'identité.

• En $x = a_x$, la dimension du noyau de $dh_{\Phi(a_x)} = dh_a$ est $n - q$ car les formes linéaires $(dh_i)_a$ sont indépendantes.

Multiplicateurs de Lagrange

• On a donc : $\operatorname{Im} d\Phi_{a_x} = \ker dh_a$ et aussi de manière évidente

$$\ker dh_a = \bigcap_{i=1}^{q} \ker\,(dh_i)_a.$$

• Dire que a est un extremum de $g_{|S}$ c'est dire que a est un extremum de $f(x) = g(x, \varphi(x))$. Or

$$\begin{aligned} df_a = 0 \quad &\Longleftrightarrow \quad dg_{\Phi(a_x)} \circ d\Phi_{a_x} = 0 \\ &\Longleftrightarrow \quad \operatorname{Im} d\Phi_{a_x} \subset \ker dg_a. \end{aligned}$$

• Ceci montre que

$$\bigcap_{i=1}^{q} \ker\,(dh_i)_a \subset \ker dg_a$$

et la proposition précédente permet de conclure $\square$

Fonction de Lagrange

Définition.– Étant données $g : U \subset \mathbb{R}^n \to \mathbb{R}$ et $h : U \to \mathbb{R}^q$ deux applications de classe C^1, la FONCTION DE LAGRANGE associée au couple (g, h) est la fonction $L : U \times \mathbb{R}^q \to \mathbb{R}$ définie par

$$L(x, \lambda_1, ..., \lambda_q) = g(x) + \sum_{i=1}^{q} \lambda_i h_i(x).$$

Théorème.– *Soient $g : U \subset \mathbb{R}^n \to \mathbb{R}$ et $h : U \to \mathbb{R}^q$ deux applications de classe C^1 et $S = h^{-1}(0)$. On suppose que les formes linéaires*

$$(dh_1)_a, ..., (dh_q)_a$$

sont linéairement indépendantes. Si $g_{|S}$ admet un extremum en $a \in S$ alors il existe $\lambda^* = (\lambda_1^*, ..., \lambda_q^*)$ tel que (a, λ^*) soit un point critique de L.

Fonction de Lagrange

Démonstration.– La différentielle de L est donnée par

$$dL_{(x,\lambda)} = dg_x + \sum_{i=1}^{q} \lambda_i (dh_i)_x + \sum_{i=1}^{q} h_i(x) d\lambda_i.$$

• Ainsi $dL_{(x,\lambda)} = 0$ si et seulement si

$$\begin{cases} dg_x + \sum_{i=1}^{q} \lambda_i (dh_i)_x & = & 0 \\ h_1(x) = \ldots = h_q(x) & = & 0 \end{cases}$$

autrement dit si $x \in S$ et s'il existe des scalaires $\lambda_1, \ldots, \lambda_q$ tels que

$$dg_a = \sum_{i=1}^{q} \lambda_i (dh_i)_a.$$

Application 1 : le théorème spectral

• À titre d'application, nous allons établir le théorème spectral à partir du théorème des multiplicateurs de Lagrange.

• Soient A une matrice carré $n \times n$ symétrique réelle et $g, h : \mathbb{R}^n \to \mathbb{R}$ les fonctions données par

$$g(x) = \langle Ax, x \rangle \quad \text{et} \quad h(x) = \|x\|^2 - 1.$$

• On considère les extrema de g sous la contrainte $S = h^{-1}(0)$, c'est-à-dire la sphère unité $\mathbb{S}^{n-1}$ de $\mathbb{R}^n$.

• Puisque S est compact et g continue (et même C^∞) de tels extrema existent. Soit $a \in S$ l'un d'entre eux.

Application 1 : le théorème spectral

• Le théorème des multiplicateurs de Lagrange assure l'existence d'un scalaire λ tel que

$$dg_a = \lambda dh_a$$

• Or

$$\begin{aligned} dh_a(V) &= \langle V, a\rangle + \langle a, V\rangle \\ &= 2\langle a, V\rangle \end{aligned}$$

et puisque ${}^tA = A$

$$\begin{aligned} dg_a(V) &= \langle AV, a\rangle + \langle Aa, V\rangle \\ &= 2\langle Aa, V\rangle. \end{aligned}$$

• Il s'en suit que $Aa = \lambda a$ autrement dit qu'il existe un couple propre (a, λ) pour toute matrice symétrique A.

Application 1 : le théorème spectral

• Le résultat que nous venons d'obtenir permet de déduire le théorème spectral

Théorème spectral.– *Si A est une matrice $n \times n$ symétrique réelle alors elle est diagonalisable dans une base orthonormée formée de vecteurs propres.*

Démonstration.– On effectue une récurrence. Le théorème est trivialement vrai pour $n = 1$.

• Supposons qu'il soit vrai pour la dimension $n - 1$. Soit A une matrice symétrique $n \times n$ et soit a le vecteur obtenu avec le théorème des multiplicateurs de Lagrange. On considère l'orthogonal de $Vect(a)$ que l'on note E.

Application 1 : le théorème spectral

• Montrons que l'espace E est stable par A. Pour cela on considère $y \in E$, donc $\langle a, y\rangle = 0$ et on constate que

$$\langle a, Ay\rangle = \langle Aa, y\rangle = \lambda\langle a, y\rangle = 0$$

donc $Ay \in E$.

• On applique ensuite l'hypothèse de récurrence à $A_{|E}$. $\square$

Application 2 : l'inégalité arithmético-géométrique

• Voici une seconde application du théorème des multiplicateurs de Lagrange :

Inégalité arithmético-géométrique.– *Soient $x_1, .., x_n$ des réels positifs, on a*

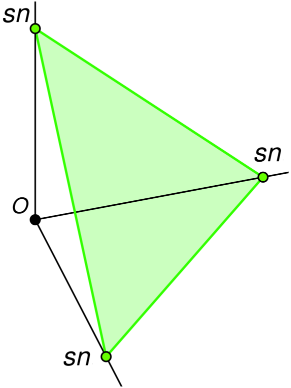
$$(x_1 \ldots x_n)^{\frac{1}{n}} \leq \frac{x_1 + \cdots + x_n}{n}.$$

Démonstration.– Soit $s > 0$. On note g et h les fonctions suivantes

$$\begin{array}{rccc} g : & (\mathbb{R}_{\geq 0})^n & \longrightarrow & \mathbb{R}^+ \\ & x = (x_1, ..., x_n) & \longmapsto & \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \end{array}$$

$$\begin{array}{rccc} h : & (\mathbb{R}_{\geq 0})^n & \longrightarrow & \mathbb{R} \\ & x = (x_1, ..., x_n) & \longmapsto & \frac{1}{n}\left(\sum_{i=1}^n x_i\right) - s. \end{array}$$

Application 2 : l'inégalité arithmético-géométrique



- On considère le sous-espace Δ^{n-1} de $\mathbb{R}^n$ défini par

$$\Delta_s = \{x \in (\mathbb{R}_{\geq 0})^n \mid \frac{1}{n}\sum_{i=1}^{n} x_i = s\} = h^{-1}(0).$$

Application 2 : l'inégalité arithmético-géométrique

• Le sous-espace Δ_s est fermé comme intersection de deux fermés. On a également

$$x \in \Delta_s \quad \Longrightarrow \quad \forall i \in \{1, ..., n\}, \quad 0 \leq x_i \leq sn.$$

Ainsi, Δ_s est borné, c'est donc un compact de $\mathbb{R}^n$.

• Par conséquent la fonction g, qui est évidemment continue, atteint son maximum en au moins un point de Δ_s.

• Notons $a = (a_1, ..., a_n)$ ce point. Les coordonnées de a sont toutes strictement positives car si une seule des coordonnées est zéro alors la valeur $g(a)$ vaut zéro et ne peut être la valeur du maximum. Ainsi $a \in \, Int\, \Delta_s$.

Application 2 : l'inégalité arithmético-géométrique

• D'après le théorème des multiplicateurs de Lagrange, il existe un scalaire λ tel que $dg_a = \lambda dh_a$ or

$$dg_x = \frac{g(x)}{n}\sum_{i=1}^{n}\frac{dx_i}{x_i} \quad \text{et} \quad dh_x = \frac{1}{n}\sum_{i=1}^{n} dx_i$$

ce qui entraîne

$$\forall i \in \{1, ..., n\}, \quad a_i = \frac{g(a)}{\lambda}.$$

• En particulier, tous les a_i sont égaux entre eux. Puisque $h(a) = 0$ cela signifie que a est le point

$$a = (s, \cdots, s)\,.$$

Application 2 : l'inégalité arithmético-géométrique

• Ceci entraîne que

$$g(a) = g(s, \cdots, s) = \left(\prod_{i=1}^{n} s \right)^{\frac{1}{n}} = s.$$

• Ainsi, pour tout $x \in \Delta_s$ on a

$$g(x_1, ..., x_n) \leq g(a) = s = \frac{x_1 + ... + x_n}{n}.$$

On a ainsi obtenu l'inégalité arithmético-géométrique pour les x de Δ_s.

• Puisque le raisonnement ne dépend pas du $s > 0$ choisi et que $(\mathbb{R}_{\geq 0})^n = \cup_{s>0} \Delta_s$, on a en réalité obtenu cette inégalité pour tous les x de $(\mathbb{R}_{\geq 0})^n$. $\square$

Joseph-Louis Lagrange (1736-1813)

Joseph-Louis Lagrange (1736-1813)

• Mathématicien et astronome d'origine italienne. Son nom est *Giuseppe Luigi Lagrangia*.

• À trente ans, il rejoint Berlin sur l'invitation de Frédéric II : « *Mon désir est que le plus grand roi d'Europe puisse compter parmi sa Cour le plus grand mathématicien d'Europe* »

• Il quitte Berlin à la mort de Frédéric II après plus de 20 ans d'une activité mathématique très féconde. Il est invité à l'Académie des Sciences à Paris et y passera les 26 dernières années de sa vie.

• Il devient « président du premier conseil » de l'École Polytechnique. En raison de sa notoriété, il est surnommé « l'illustrissime » par les étudiants.

Joseph-Louis Lagrange (1736-1813)

• Pour autant, il n'aime pas donner des cours. Sa voix faible et son accent italien en font un enseignant peu apprécié (jugement rapporté par Joseph Fourier).

• Il est avec Euler, le fondateur du calcul variationnel et de la théorie des formes quadratiques.

• Il résout la conjecture de Bachet : tout entier positif est somme de quatre carrés.

• En algèbre, le théorème de Lagrange affirme que le cardinal de tout sous-groupe H d'un groupe fini G divise le cardinal de G.

• En physique, en développant le principe de moindre action, il invente le *lagrangien* et les *équations de Lagrange* qui lui sont associées.

Joseph-Louis Lagrange (1736-1813)

• À la mort de son épouse Vittoria Conti en 1783, il plonge dans une profonde dépression. Il cesse d'écrire et de publier pendant plusieurs années.

• Il subit une trentaine de saignées : il était persuadé que sa mélancolie le prédisposait à accumuler les humeurs propices aux varices et aux hémorroïdes

• Il ne sort de sa mélancolie qu'en 1792 lorsqu'il rencontre Renée-Françoise-Adélaïde Le Monnier, sa seconde épouse.

Lagrange & Lavoisier

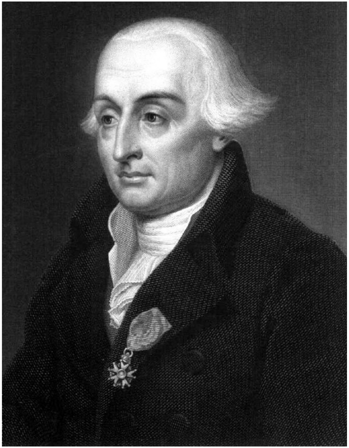
• Lagrange et Lavoisier élaborent le système métrique pendant la Révolution. Ils se lient d'amitié.

• En pleine Terreur, une loi ordonne l'arrestation et la confiscation des biens de tous les étrangers nés dans un pays ennemi. Lavoisier intervient et réussit à faire protéger Lagrange.

• Quatre mois plus tard, c'est Lavoisier qui est arrêté avec vingt-sept autres fermiers généraux comme « traître à la nation ». Il est condamné à mort par le tribunal révolutionnaire. Son président, Jean-Baptiste Coffinhal, aurait déclaré

« *La République n'a pas besoin de savants, ni de chimistes ; le cours de la justice ne peut être suspendu* »

Lagrange & Lavoisier

« *Il a fallu un instant pour couper sa tête, et un siècle ne suffira pas pour en produire une si bien faite.* »

Lagrange au sujet de l'exécution de Lavoisier pendant la Terreur.

Interprétation géométrique

• L'hypothèse faite sur h dans le théorème des multiplicateurs de Lagrange signifie que h est une submersion en tout point $a \in h^{-1}(0)$. Cela implique que $S = h^{-1}(0)$ est une sous-variété de dimension $n - q$ de $\mathbb{R}^n$.

Définition.– Soit $S \subset \mathbb{R}^n$ une sous-variété et $f : S \to \mathbb{R}$. On dit que $a \in S$ est un POINT CRITIQUE si $df_a : \overrightarrow{T_aS} \to \mathbb{R}$ est la forme linéaire nulle.

Reformulation du problème des extrema liés.– Chercher les points critiques de $g : U \subset \mathbb{R}^n \to \mathbb{R}$ sous la contrainte $S = h^{-1}(0)$ c'est donc chercher les points critiques de la fonction $f = g_{|S} : S \to \mathbb{R}$.

Interprétation géométrique

Théorème.– *Le point $a \in S$ est un point critique de f si et seulement si* grad $g(a)$ *est orthogonal à* $\overrightarrow{T_aS}$.

Démonstration.– ($\Longrightarrow$) Le théorème des multiplicateurs de Lagrange assure l'existence de $\lambda_1, \ldots, \lambda_p$ tels que

$$dg_a = \sum_{i=1}^{p} \lambda_i (dh_i)_a.$$

• Par dualité, ceci signifie que

$$\operatorname{grad} g(a) = \sum_{i=1}^{p} \lambda_i \operatorname{grad} h_i(a).$$

• Or

$$\overrightarrow{T_aS} = \ker dh_a = Vect^{\perp}(\operatorname{grad} h_1(a), \cdots, \operatorname{grad} h_q(a)),$$

le vecteur grad $g(a)$ est donc orthogonal à $\overrightarrow{T_aS}$.

Interprétation géométrique

• ($\Longleftarrow$) Réciproquement, si grad $g(a)$ est donc orthogonal à $\overrightarrow{T_aS}$ alors dg_a s'annule sur $\overrightarrow{T_aS}$ ce qui entraîne que $df_a = 0$. Par conséquent, a est point critique de f. $\square$

Perspective.– La théorie de l'optimisation sous contraintes va au delà de la recherche de points critiques. La considération de la *matrice hessienne bordée* permet de donner des conditions suffisantes pour qu'un point critique soit un extremum.

• Une mise en garde : il n'est en général pas possible de définir la différentielle seconde d'une fonction $f : S \to \mathbb{R}$ en tout point de la sous-variété S et ceci à cause de la présence d'un terme en df dans la règle de la chaîne :

$$d^2 g_{x'}(V, W) = d^2 f_x(d\phi_{x'}(V), d\phi_{x'}(W)) + df_x(d^2\phi_{x'}(V, W)).$$

Interprétation géométrique

• En revanche, une telle définition existe en un point critique $a \in S$. Voyons pourquoi.

Définition.– Soit $a \in S$ un point critique de $f : S \to \mathbb{R}$. On définit la DIFFÉRENTIELLE SECONDE $d^2 f_a$ DE f EN LE POINT CRITIQUE a comme étant

$$d^2 f_a := d^2 (f \circ \varphi)_0$$

où φ est une paramétrisation (locale) de S telle que $\varphi(0) = a$.

• Cette définition est consistante car puisque l'on est en un point critique, $df_a = 0$ et la règle de la chaîne montre alors que la définition de $d^2 f_a$ est indépendante de la carte choisie.