Optimisation (suite)

1 Extrema liés

1.1 Une seule contrainte

Il s'agit de trouver les extrema de $f(x\,,\,y\,,\,z)$ lorsque $(x\,,\,y\,,\,z)$ appartient à une surface S définie par $g(x\,,\,y\,,\,z) = C$.

Définition 1.1. *Un point $P = (x_0\,,\,y_0\,,\,z_0)$ est un minimum (resp. maximum) local pour f, lié à la contrainte $g(x\,,\,y\,,\,z) = C$ si :*

(i) $g(P) = C$

(ii) Il existe $\varepsilon > 0$ tel que $f(P) \leqslant f(Q)$ (resp. $f(P) \geqslant f(Q)$) pour tout $Q \in S \cap B(P\,,\,\varepsilon)$.

Théorème 1.2. *(de Lagrange)*
Soit $f(x\,,\,y\,,\,z)$ et $g(x\,,\,y\,,\,z)$ de classe $\mathcal{C}^1$ telle que $\nabla g \neq 0$ sur S.
Alors si f admet un **extrema lié** *en $(x_0\,,\,y_0\,,\,z_0)$ on a : $\nabla f(x_0\,,\,y_0\,,\,z_0) = \lambda\, \nabla g(x_0\,,\,y_0\,,\,z_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.*

Démonstration Ce qui suit n'est qu'une idée de démonstration qu'il faudrait préciser.
Soit $(x(t), y(t), z(t))$ une courbe tracée sur la surface S passant en (x_0, y_0, z_0) en $t = 0$. Pour que f soit maximale en (x_0, y_0, z_0) sur S il faut en particulier que $t \mapsto f(x(t), y(t), z(t)$ soit maximale en $t = 0$. Pour que ce soit le cas il faut que cette fonction ait une dérivée nulle en 0. La dérivée de cette fonction se calcule en appliquant la règle de dérivation en chaîne, $x'(t)\frac{\partial f}{\partial x} + y'(t)\frac{\partial f}{\partial y} + z'(t)\frac{\partial f}{\partial z}$. On obtient la condition

$$x'(0)\frac{\partial f}{\partial x}(x_0\,,\,y_0\,,\,z_0) + y'(0)\frac{\partial f}{\partial y}(x_0\,,\,y_0\,,\,z_0) + z'(0)\frac{\partial f}{\partial z}(x_0\,,\,y_0\,,\,z_0) = 0.$$

Par ailleurs, comme S est définie par $g = C$ le plan tangent à S en (x_0, y_0, z_0) a pour équation :

$$(x - x_0)\frac{\partial g}{\partial x}(x_0\,,\,y_0\,,\,z_0) + (y - y_0)\frac{\partial g}{\partial y}(x_0\,,\,y_0\,,\,z_0) + (z - z_0)\frac{\partial g}{\partial z}(x_0\,,\,y_0\,,\,z_0) = 0.$$

On a aussi

$$x'(0)\frac{\partial g}{\partial x}(x_0\,,\,y_0\,,\,z_0) + y'(0)\frac{\partial g}{\partial y}(x_0\,,\,y_0\,,\,z_0) + z'(0)\frac{\partial g}{\partial z}(x_0\,,\,y_0\,,\,z_0) = 0.$$

Lorsqu'on fait varier $(x(t), y(t), z(t))$ le vecteur $(x'(0), y'(0), z'(0))$ prend toutes les valeurs de la direction du plan tangent à S en $(x_0\,,\,y_0\,,\,z_0)$. Cela signifie que $\nabla f(x_0\,,\,y_0\,,\,z_0)$ est orthogonaux à tous les vecteurs orthogonaux à $\nabla g(x_0\,,\,y_0\,,\,z_0)$ ou encore que $\nabla f(x_0\,,\,y_0\,,\,z_0)$

et $\nabla g(x_0\,,\,y_0\,,\,z_0)$ ont mêmes vecteurs orthogonaux. Cela entraîne que $\nabla f(x_0\,,\,y_0\,,\,z_0)$ et $\nabla g(x_0\,,\,y_0\,,\,z_0)$ sont colinéaires. Soit U, V deux vecteurs orthogonaux à $\nabla g(x_0\,,\,y_0\,,\,z_0)$ tels que $(U, V, \nabla g(x_0\,,\,y_0\,,\,z_0))$ soit une base de $\mathbb{R}^3$. $\square$

Remarque
Si P est un extremum lié, on a $\nabla f(P)$ parallèle à $\nabla g(P)$. La réciproque n'est pas vraie. Nous avons une condition nécessaire mais pas suffisante. C'est l'équivalent de la nullité de la dérivée pour les extrema libres : en un extremum libre la dérivée est nulle mais la dérivée peut être nulle sans que la fonction ait un extremum (penser à $x \mapsto x^3$ en $x = 0$).

Le théorème est encore vrai en dimension plus grande.

Théorème 1.3. *(de Lagrange)*
Soit f et g deux fonctions de $\mathbb{R}^d$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$ telle que $\nabla g \neq 0$ sur l'ensemble S défini par $g = C$ (c'est une hypersurface).
Alors si f admet un **extrema lié** *en x_0 on a : $\nabla f(x_0) = \lambda \nabla g(x_0)$ où $\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.*

1.2 Plusieurs contraintes

On cherche les extrema d'une fonction f sur l'ensemble S défini par $g_1 = g_2 = \ldots = g_k = 0$, toutes les fonctions considérées étant de classe C^1. Pour que les choses marchent bien il faut faire l'hypothèse suivante : en tout point de S les gradients des fonctions g_i sont linéairement indépendants.

Théorème 1.4. *(de Lagrange)*
Soit f et $g_1, \ldots, g_k$ $k+1$ fonctions de $\mathbb{R}^d$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$ telles que les vecteurs $\nabla g_1, \ldots, \nabla g_k$, soit indépendants sur sur l'ensemble S défini par $g_1 = \ldots = g_k = 0$.
Alors si f admet un **extrema lié** *sur S en x_0 le vecteur $\nabla f(x_0)$ est combinaison linéaire des vecteurs $\nabla g_i(x_0)$: il existe $\lambda_1, \ldots, \lambda_k$ tels que $\nabla f(x_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x_0)$. Les nombres $\lambda_1, \ldots, \lambda_k$ sont appelés multiplicateurs de Lagrange.*

<u>Démonstration</u> On montre de la même manière que précédemment $\nabla f(x_0)$ est orthogonal à tous les vecteurs orthogonaux à tous les vecteurs $\nabla g_i(x_0)$ (c'est là qu'on utilise l'hypothèse d'indépendance sur les gradients). Considérons une base de $\mathbb{R}^d$ constituée de $v_1 = \nabla g_1(x_0), \ldots, v_k = \nabla g_k(x_0)$ complétée par une famille $(v_{k+1}, \ldots, v_d)$ de vecteurs orthogonaux entre eux et orthogonaux à tous les $\nabla g_i(x_0)$. Écrivons $\nabla f(x_0)$ dans cette base : $\nabla f(x_0) = \sum_{i=1}^{d} \lambda_i v_i$. Pour tout $j > k$, on a

$$\begin{aligned}
\langle \nabla f(x_0), v_j \rangle &= \langle \sum_{i=1}^{d} \lambda_i v_i, v_j \rangle \\
&= \sum_{i=1}^{d} \lambda_i \langle v_i, v_j \rangle \\
&= \lambda_j \langle v_j, v_j \rangle.
\end{aligned}$$

Mais, par ailleurs, pour $j > k$, on a $\langle \nabla f(x_0), v_j \rangle = 0$. On en déduit donc que, pour $j > k$, λ_j est nul, c'est-à-dire que $\nabla f(x_0)$ est combinaison linéaire des gradients $\nabla g_i(x_0)$. $\square$

2 Extrema de f sur un compact $K \subset \mathbb{R}^d$

La marche à suivre pour étudier les extrema d'une fonction différentiable sur un compact de $\mathbb{R}^d$ est la suivante.

Soit f une fonction différentiable définie sur un compact K de $\mathbb{R}^d$. Comme f est différentiable, elle est continue. Elle est donc bornée sur K et atteint ses bornes.

En pratique (dans les exercices que je vous demanderai de résoudre en particulier) la fonction f sera donnée par une formule valable sur un certain sous-ensemble de $\mathbb{R}^d$ et le compact K sera inclus dans cet ensemble de définition.

On mène l'étude des extrema de f en plusieurs étapes. La première est d'étudier l'existence d'extrema locaux de f à l'intérieur de K. C'est pour cette étude qu'on utilisera le développement de Taylor à l'ordre 2.

Mais cette étude n'est pas suffisante. Il faut aussi regarder ce qui se passe sur le bord de K. Pour cela on peut utiliser les multiplicateurs de Lagrange.

Exemple
Etude de $f(x\,,\,y) = x^2 - y^2$ sur $K = \{(x\,,\,y) \in \mathbb{R}^2 \,/\, x^2 + y^2 \leqslant 1\}$. On procède de la manière suivante :

(i) On cherche les points critiques et les extrema locaux dans $Int(K)$.
On trouve un seul point stationnaire en $(0, 0)$. Mais en $(0, 0)$ f a un point selle. La fonction n'a donc pas d'extremum à l'intérieur de K. Mais comme K est compact et f est continue sur K, f est bornée sur K et atteint ses bornes sur K. Ce sera donc sur le bord de K.

(ii) On analyse f sur ∂K.
• Une possibilité ici est de paramétrer le bord de K : le cercle de rayon 1 centré en $(0, 0)$. On obtient : $f(\cos t, \sin t) = \cos^2 t - \sin^2 t = \cos(2t)$. On peut alors étudier les variations de cette fonction. On obtient qu'elle est maximum égale à 1 lorsque $2t$ est égal à 0 modulo 2π, minimum égale à -1 lorsque $2t$ vaut π modulo 2π. La fonction f atteint donc son maximum 1 aux points $(1, 0)$ et $(-1, 0)$ de K, son minimum -1 aux points $(0, 1)$ et $(0, -1)$.
• On peut aussi utiliser le multiplicateur de Lagrange (qui ne donne qu'une condition nécessaire pour avoir un extremum lié en un point). En un point où f est extrémale sur l'ensemble $\{(x\,,\,y) \in \mathbb{R}^2 \,/\, x^2 + y^2 = 1\}$, le gradient de f est colinéaire à celui de $g(x, y) = x^2 + y^2$. En un tel point on a donc :

$$x^2 + y^2 = 1, (2x, 2y) = \lambda(2x, -2y),$$

pour un certain nombre réel λ. Le point $(0,0)$ n'est pas sur le cercle. Si $x \neq 0$, ces égalités entraînent que λ vaut 1 et $2y = -2y$, donc $y = 0$. Deux points satisfont les conditions : $(1,0)$ et $(-1,0)$. Si $y \neq 0$, alors $\lambda = -1$, $x = 0$, et on obtient deux autres points satisfaisant les conditions : $(0,1)$ et $(0,-1)$. Comme on sait que f atteint ses bornes sur le cercle (par l'argument (i)), il suffit de calculer les valeurs de f en ces quatre points pour aboutir à la même conclusion que celle nous avions obtenue avec le paramétrage.

Intégration des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

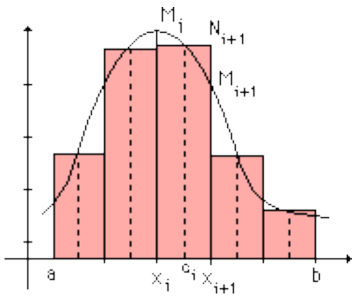
3 Intégration des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

3.1 Intégration des fonctions d'une variable

Soit $f : [a\,,\,b] \to \mathbb{R}$ bornée.

a) Cas où f est en escalier.
Il existe une partition $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ telle que f est constante sur chaque intervalle $\,]t_i\,,\,t_{i+1}[\,$ $(f(x) = C_i)$.
Alors $\displaystyle\int_a^b f(t)\,dt = \sum_{i=0}^{n-1} C_i\,(t_{i+1} - t_i)$.

b) Si f est bornée on l'approche par des fonctions en escalier.



Définition 3.1. *f est* **intégrable** *sur $[a\,,\,b]$ s'il existe un nombre unique I tel que pour toutes fonctions en escalier $u\,,\,v$ sur $[a\,,\,b]$ vérifiant $u(x) \leqslant f(x) \leqslant v(x)$ on a :*

$$\int_a^b u(x)\,dx \quad\leqslant\quad I \quad\leqslant\quad \int_a^b v(x)dx$$

et si pour tout $\varepsilon > 0$ il existe des fonctions en escalier u_ε et v_ε vérifiant :

$$u_\varepsilon(x) \quad\leqslant\quad f(x) \quad\leqslant\quad v_\varepsilon(x)$$

et

$$0 \quad\leqslant\quad \int_a^b v_\varepsilon(x)\,dx - \int_a^b u_\varepsilon(x)\,dx \quad<\quad \varepsilon$$

Notation : I s'appelle l'intégrale de f sur $[a\,,\,b]$ et se note $\displaystyle\int_a^b f(x)\,dx\,.$

Théorème 3.2. *Si f est continue sur $[a,b]$ alors f est intégrable sur $[a,b]$.*

Démonstration : Soit f une fonction continue sur $[a, b]$. Soit ϵ un nombre réel strictement positif. Alors la fonction f est uniformément continue sur $[a, b]$: il existe $\eta > 0$ tel que, si $|x - y| < \eta$ alors $|f(x) - f(y)| < \epsilon$. Prenons n tel que $(b - a)/n < \eta$ et divisons $[a, b]$ en n intervalles de longueur $(b-a)/n < \eta$ dont les extrémités sont les points $x_k = a+k(b-a)/n$, k variant de 0 à n (k prenant $n + 1$ valeurs détermine n intervalles $[x_k, x_{k+1}]$ k variant de 0 à $n - 1$; on a $x_0 = a$ et $x_n = b$). Définissons deux fonctions en escalier

$$f^-(x) = \min\{f(t), t \in [x_k, x_{k+1}]\} \text{ si } x \in [x_k, x_{k+1}]$$

$$f^+(x) = \max\{f(t), t \in [x_k, x_{k+1}]\} \text{ si } x \in [x_k, x_{k+1}]$$

Par définition on a $f^- \leq f \leq f^+$ et on a

$$\begin{aligned}
\int_a^b f^+(x)dx - \int_a^b f^-(x)dx &= \sum_{k=0}^{n-1}(x_{k+1} - x_k)\max\{f(t), t \in [x_k, x_{k+1}]\} \\
&\quad - \sum_{k=0}^{n-1}(x_{k+1} - x_k)\min\{f(t), t \in [x_k, x_{k+1}]\} \\
&= \sum_{k=0}^{n-1}(x_{k+1} - x_k)(\max_{t\in[x_k,x_{k+1}]} f(t) - \min_{t\in[x_k,x_{k+1}]} f(t)) \\
&\leq \sum_{k=0}^{n-1}(x_{k+1} - x_k)\epsilon \\
&= \epsilon\sum_{k=0}^{n-1}(x_{k+1} - x_k) \\
&= \epsilon(b - a)
\end{aligned}$$

□

3.2 Intégration des fonctions de plusieurs variables

Soit $f : [a_1\,,\,b_1] \times [a_2\,,\,b_2] \times \;\ldots\; \times [a_n\,,\,b_n] \to \mathbb{R}$.
Ici on considère le cas $R = [a\,,\,b] \times [c\,,\,d] \subset \mathbb{R}^2$.

a) f est en escalier.
   Il existe une partition de $[a\,,\,b] \times [c\,,\,d]$:
   $a = s_0 < s_1 < s_2 < \;\ldots\; < s_m = b$
   $c = t_0 < t_1 < \;\ldots\; < t_n = d$
   telle que f est constante à l'intérieur de chaque rectangle $]s_i\,,\,s_{i+1}[\times]t_j\,,\,t_{j+1}[$ (où elle vaut C_{ij}).
   On définit $\displaystyle\iint_R f(x\,,\,y)\,dx\,dy = \sum_{i\,,\,j} C_{ij}\,(s_{i+1} - s_i)\,(t_{j+1} - t_j)$.

b) f est bornée sur $R = [a\,,\,b] \times [c\,,\,d]$.
On approche f par des fonctions en escalier.

Définition 3.3. *f est* **intégrable** *sur R s'il existe un nombre unique I tel que pour toutes fonctions en escalier $u(x\,,\,y)$ et $v(x\,,\,y)$, telles que $u(x\,,\,y) \leqslant f(x\,,\,y) \leqslant v(x\,,\,y)$, on a :*

$$\iint_R u(x\,,\,y)\,dx\,dy \quad \leqslant \quad I \quad \leqslant \quad \iint_R v(x\,,\,y)\,dx\,dy$$

et si, pour tout $\varepsilon > 0$, il existe des fonctions en escalier u_ε et v_ε telles que :

$$u_\varepsilon(x\,,\,y) \quad \leqslant \quad f(x\,,\,y) \quad \leqslant \quad v_\varepsilon(x\,,\,y)$$

et

$$0 \quad \leqslant \quad \int_R v_\varepsilon(x\,,\,y)\,dx\,dy - \int_R u_\varepsilon(x\,,\,y)\,dx\,dy \quad < \quad \varepsilon$$

Notation : I s'appelle l'intégrale de f sur R et se note $\displaystyle\iint_R f(x\,,\,y)\,dx\,dy$.

Proposition 3.4. *(Propriétés de l'intégrale double)*

1. *Si f est continue sur R alors f est intégrable.*
2. *Si f est positive sur R alors $\displaystyle\iint_R f(x\,,\,y)\,dx\,dy$ est le volume sous le graphe de f au-dessus de R.*
3. $\displaystyle\iint_R (\alpha f + \beta g)(x\,,\,y)\,dx\,dy = \alpha \iint_R f(x\,,\,y)\,dx\,dy + \beta \iint_R g(x\,,\,y)\,dx\,dy.$
4. *Si $R = R_1 \cup R_2$ avec $R_1 \cap R_2 = \emptyset$ alors :*

$$\iint_R f(x\,,\,y)\,dx\,dy = \iint_{R_1} f(x\,,\,y)\,dx\,dy + \iint_{R_2} f(x\,,\,y)\,dx\,dy$$

Théorème 3.5. *Si f est continue sur R alors $\displaystyle\iint_R f(x\,,\,y)\,dx\,dy$ existe.*

3.3 Calcul des intégrales doubles

Théorème 3.6. *(Fubini) Si f est continue sur $R = [a\,,\,b] \times [c\,,\,d]$ alors :*

$$\iint_R f(x\,,\,y)\,dx\,dy \quad = \quad \int_a^b \left(\int_c^d f(x\,,\,y)\,dy\right) dx \quad = \quad \int_c^d \left(\int_a^b f(x\,,\,y)\,dx\right) dy$$

Exemple

(1) $\displaystyle\iint_R (x^2 + y^2)\,dx\,dy \qquad R = [0\,,\,1] \times [0\,,\,1]$

(2) $\displaystyle\iint_R (1 + x + y)\,dx\,dy \qquad R = [0\,,\,1] \times [0\,,\,1]$