Chapitre 22

Extrema sous contrainte

Compétences attendues.

✓ Identifier la nature d'une contrainte (contrainte non critique ou d'égalités linéaires).

✓ Déterminer les points critiques de f sous une contrainte non critique.

✓ Déterminer les extrema locaux de f sous une contrainte d'égalités linéaires.

Mathieu Mansuy - Professeur de Mathématiques en ECS2 au Lycée Louis Pergaud (Besançon)
mansuy.mathieu@hotmail.fr

Exemple introductif

Un cultivateur souhaite exporter sa production de petits pois. Il veut pour cela réduire ses coûts d'emballage au maximum. Il opte pour un emballage métallique de forme cylindrique. À surface fixée égale à $s\pi$, quel cylindre a un volume maximal en vue d'optimiser les coûts d'exportation ?

Notons $r > 0$ le rayon du cylindre, et $h > 0$ sa hauteur. La surface et le volume du cylindre sont donnés respectivement par les expressions :

$$S(r,h) = \pi(2r^2 + 2hr) \quad \text{et} \quad V(r,h) = \pi r^2 h.$$

Il s'agit donc de trouver le maximum de la fonction V, non pas sur $\mathbb{R}^2$ tout entier, mais seulement sur le sous-ensemble :

$$\mathscr{C} = \{(r,h) \in (\mathbb{R}_+^*)^2,\ S(r,h) = s\pi\}.$$

Dans ce chapitre, nous développons des méthodes afin de déterminer les extrema (locaux ou globaux) d'une telle fonction sur une partie (en général fermée) $\mathscr{C}$ d'un ouvert Ω. On parle alors d'optimisation sous la contrainte $\mathscr{C}$.

1 Optimisation sous une contrainte non critique

Soit f une fonction définie sur un ouvert Ω de $\mathbb{R}^n$ avec $n \geq 1$. Dans tout ce qui suit, nous travaillerons avec des contraintes d'égalité : φ désignera une fonction de classe $\mathscr{C}^1$ sur Ω, et $c \in \mathbb{R}$ une constante. On notera alors la *contrainte* :

$$\mathscr{C} = \{x \in \Omega,\ \varphi(x) = c\}.$$

Nous cherchons les extrema[1] de f sur $\mathscr{C}$.

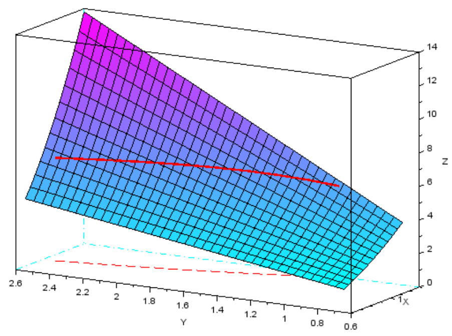
1.1 Définitions

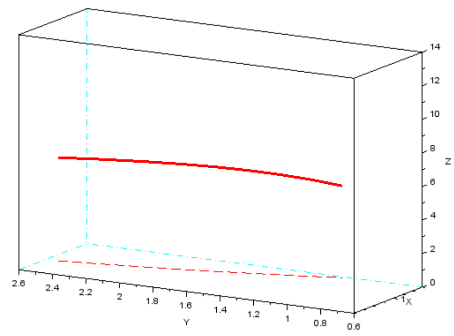
Définition.

On dit que f *admet un extremum en a sous la contrainte $\mathscr{C}$* lorsque f admet un extremum en a sur l'ensemble $\mathscr{C} \cap \Omega$.

Exemple. Recherche d'extrema sous contrainte par substitution.
Cherchons les extrema de $V(r,h) = \pi r^2 h$ sous la contrainte $\mathscr{C} = \{(r,h) \in (\mathbb{R}_+^*)^2,\ S(r,h) = s\pi\}$.

[1] Notons que $\mathscr{C}$ étant un fermé, il n'est pas possible d'utiliser les outils de calcul différentiel introduits jusqu'à maintenant, qui nécessitent de travailler sur un ouvert.





Les extrema de V sous la contrainte $\mathscr{C}$ sont les extrema de la courbe rouge (avec $s = 6$).

Méthode.
Si la contrainte $\mathscr{C}$ permet d'exprimer l'une des variables en fonction des autres (ici h en fonction de r), on peut par substitution se ramener à la recherche d'extrema pour une fonction à strictement moins de variables sur un ouvert.

1.2 Théorème des extrema liés

Rappel. Soit f une fonction de classe $\mathscr{C}^1$ sur un **ouvert** Ω de $\mathbb{R}^n$. Nous avons la condition **nécessaire** d'extremum sur Ω suivante :

$$f \text{ admet un extremum local en } a \quad \Rightarrow \quad \nabla f(a) = (0, \ldots, 0).$$

Dans cette section, on étudie l'analogue de ce résultat dans le cas d'un extremum sous contrainte.

Définition.

On dit que $\mathscr{C} = \{x \in \Omega,\ \varphi(x) = c\}$ est une *contrainte non critique* si on a :

$$\forall x \in \mathscr{C}, \quad \nabla\varphi(x) \neq (0, \ldots, 0).$$

Exercice. Montrer que la contrainte $\mathscr{C} = \{(r, h) \in (\mathbb{R}_+^*)^2,\ S(r, h) = s\pi\}$ est non critique.

Théorème 1 (*des extrema liés*)

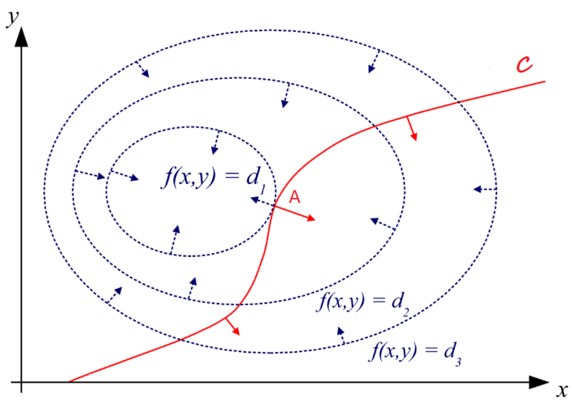
Soit f une fonction de classe $\mathscr{C}^1$ sur un ouvert Ω de $\mathbb{R}^n$, et soit $\mathscr{C} = \{x \in \Omega,\ \varphi(x) = c\}$ une contrainte non critique.

Si f admet un extremum local en a sous la contrainte $\mathscr{C}$, **alors** il existe un réel λ tel que :

$$\begin{cases} \varphi(a) = c \\ \nabla f(a) = \lambda \nabla\varphi(a) \end{cases} \qquad (*)$$

Un point a vérifiant ces conditions est appelé *point critique sous la contrainte $\mathscr{C}$*.

Idée de la preuve. Plaçons nous dans le cas où $n = 2$. On cherche les extrema de la fonction f sous la contrainte $\mathscr{C}$. Cela signifie qu'on cherche les extrema de f lorsque le point (x, y) appartient à la courbe de niveau $\varphi(x, y) = c$. Représentons cette courbe, ainsi que plusieurs lignes de niveau de f, d'équations $f(x, y) = d_i$ pour $i = 1, 2, 3$.



Optimiser f sous la contrainte $\varphi(x, y) = c$ signifie chercher la plus grande (ou plus petite) valeur de d telle que la ligne de niveau $f(x, y) = d$ croise la courbe $\varphi(x, y) = c$. Or ceci ne peut se produire que si **ces deux courbes sont tangentes au point A d'intersection** : en effet une ligne de niveau délimite des points d'altitude strictement plus haute d'un côté, et strictement plus basse de l'autre. Ainsi, si la courbe $\mathscr{C}$ traverse la ligne de niveau $f(x, y) = d$, alors il y aurait des points au voisinage de A qui prendraient des valeurs strictement supérieures et strictement inférieures à $f(A)$.

D'autre part, rappelons nous que le gradient est orthogonal aux tangentes aux lignes de niveau, et que cela vaut aussi bien pour f que pour φ. Donc $\nabla f(A)$ et $\nabla \varphi(A)$ sont tous les deux des vecteurs orthogonaux à une même droite, la tangente commune aux deux lignes de niveau : ils ont donc même direction, et sont donc colinéaires.

$\square$

Mise en garde.

De même que dans le cas d'une recherche d'extrema sur un ouvert, la condition $(*)$ est une condition nécessaire d'extremum sous la contrainte $\mathscr{C}$, mais pas suffisante. En d'autres termes, on a :

$$a \text{ extremum sous la contrainte } \mathscr{C} \quad \Rightarrow \quad a \text{ point critique sous la contrainte } \mathscr{C}$$

mais la réciproque est fausse. Aussi, si a est un point critique sous la contrainte $\mathscr{C}$, il faudra faire une étude plus poussée pour déterminer si oui ou non c'est un extremum local sous la contrainte $\mathscr{C}$.

Exercice. Déterminer les points critiques de V sous la contrainte $\mathscr{C} = \{(x, y),\ S(r, h) = s\pi\}$.

Méthode.

Pour déterminer les points critiques de $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ sous une contrainte non critique $\mathscr{C}$, on procèdera comme suit :

(i) on justifie que f et φ sont de classe $\mathscr{C}^1$ et on calcule leur gradient.

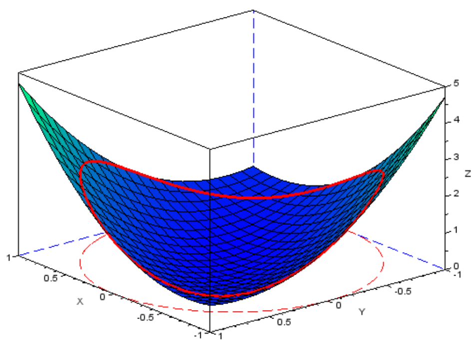
(ii) on vérifie que la contrainte $\mathscr{C}$ est non critique, c'est-à-dire que $\nabla\varphi(x) \neq (0, \ldots, 0)$ pour tout $x \in \mathscr{C}$.

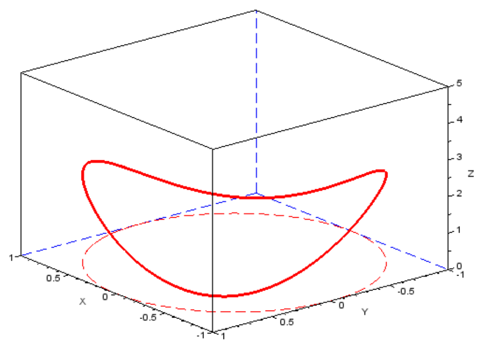
(iii) on résout le système :

$$\begin{cases} \varphi(x_1, \ldots, x_n) = c \\ \partial_1 f(x_1, \ldots, x_n) = \lambda \partial_1 \varphi(x_1, \ldots, x_n) \\ \qquad \vdots \\ \partial_n f(x_1, \ldots, x_n) = \lambda \partial_n \varphi(x_1, \ldots, x_n) \end{cases}$$

à $n+1$ équations et $n+1$ inconnues $(x_1, \ldots, x_n, \lambda)$. Les solutions sont les points critiques de f sous la contrainte $\mathscr{C}$.

Exercice. Déterminer les points critiques de $f(x, y) = \sqrt{1 + x^2 + y^2} + (x + y)^2 - 1$ sous la contrainte $\mathscr{C} = \{(x, y),\ x^2 + y^2 = 3\}$.





Méthode.

Pour obtenir la nature (locale ou globale) d'un point critique a, on pourra selon les cas :

- *étudier le signe de $f(a+h) - f(a)$ pour $a + h \in \Omega \cap \mathscr{C}$,*
- *utiliser un argument topologique (si $\mathscr{C}$ est un fermé borné),*

ou toute autre indication fournie par l'énoncé.

1.3 Application à la recherche d'extrema sur un fermé borné

On considère une fonction f de classe $\mathscr{C}^1$ sur un ensemble fermé borné de la forme $\mathscr{D} = \{x \in \mathbb{R}^n,\ \varphi(x) \leq c\}$ où φ est supposée $\mathscr{C}^1$. On sait alors que f admet un maximum global et un minimum global sur $\mathscr{D}$, mais on ne sait pas en quels points ils sont atteints.

Méthode.

Pour déterminer les extrema globaux de f sur $\mathscr{D}$, on procèdera ainsi :

(i) on justifie que f admet un minimum global et un maximum global sur $\mathscr{D}$ en montrant que

- *f est continue sur $\mathscr{D}$,*
- *$\mathscr{D}$ est un fermé borné.*

(ii) sur l'intérieur $\mathscr{D}_0 = \{x \in \mathbb{R}^n,\ \varphi(x) < c\}$ de $\mathscr{D}$: $\mathscr{D}_0$ est ouvert, et on peut donc appliquer les théorèmes usuels de calcul différentiel pour déterminer les points critiques de f.

(iii) sur le bord $\mathscr{C} = \{x \in \mathbb{R}^n,\ \varphi(x) = c\}$: on peut utiliser le théorème des extrema liés pour déterminer les points critiques de f sous la contrainte $\mathscr{C}$.

(iv) reste alors à comparer les valeurs de f en chacun de ses points : f atteint un maximum global et un minimum global en l'un de ces points.

Exercice. Déterminer les extrema globaux de $f(x,y) = \sqrt{1+x^2+y^2} + (x+y)^2 - 1$ sur $\mathscr{D} = \{(x,y),\ x^2+y^2 \leq 3\}$.

2 Optimisation sous contrainte d'égalités linéaires

Dans cette partie, on suppose que la contrainte $\mathscr{C}$ est l'ensemble des solutions du système linéaire :

$$\mathscr{C} : \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \quad\vdots \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n = b_p \end{cases}$$

2.1 Systèmes linéaires

Pour tout $1 \leq i \leq p$, notons $g_i : \mathbb{R}^n \to \mathbb{R}$ la forme linéaire définie par :

$$g_i(x_1, \ldots, x_n) = a_{i,1}x_1 + \cdots + a_{i,p}x_p,$$

de sorte que $\mathscr{C}$ se réécrit :

$$\mathscr{C} : \begin{cases} g_1(x) = b_1 \\ g_2(x) = b_2 \\ \quad\vdots \\ g_p(x) = b_p \end{cases}$$

On note $\mathscr{H}$ l'ensemble des solutions du système homogène associé, c'est-à-dire de :

$$\mathscr{C}_0 : \begin{cases} g_1(x) = 0 \\ g_2(x) = 0 \\ \quad\vdots \\ g_p(x) = 0 \end{cases}$$

Propriété 2 (Structure des solutions d'un système linéaire)

- L'ensemble $\mathscr{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$.
- Si $x_0 \in \mathbb{R}^n$ est solution de $\mathscr{C}$, alors on a :

$$x \text{ solution de } \mathscr{C} \quad \Leftrightarrow \quad x - x_0 \in \mathscr{H}.$$

  Ainsi, toute solution x de $\mathscr{C}$ est la somme d'une solution particulière x_0 de $\mathscr{C}$ et d'une solution $h \in \mathscr{H}$ du système homogène associé.

Preuve.

- On a $\mathscr{H} = \bigcap_{i=1}^{p} \mathrm{Ker}(g_i)$, donc $\mathscr{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$ en tant qu'intersection des sous-espaces $\mathrm{Ker}(g_i)$ (les g_i sont linéaires).
- Soit $x_0 \in \mathbb{R}^n$ solution de $\mathscr{C}$. On a :

$$x \text{ solution de } \mathscr{C} \Leftrightarrow \begin{cases} g_1(x) = b_1 = g_1(x_0) \\ \quad\vdots \\ g_p(x) = b_p = g_p(x_0) \end{cases} \Leftrightarrow \begin{cases} g_1(x - x_0) = g_1(x) - g_1(x_0) = 0 \\ \quad\vdots \\ g_p(x - x_0) = g_p(x) - g_p(x_0) = 0 \end{cases} \Leftrightarrow x - x_0 \in \mathscr{H}$$

  Ainsi x est solution de $\mathscr{C}$ si et seulement si il existe $h \in \mathscr{H}$ tel que $x = x_0 + h$.

$\square$

Exercice. Déterminer l'ensemble des solutions du système linéaire $\begin{cases} x+y+z=3 \\ y-z=0 \end{cases}$.

Pour tout $1 \leq i \leq p$, la fonction g_i est polynomiale, donc de classe $\mathscr{C}^1$. $\mathscr{C}$ est fermé comme intersection des fermés $\{x \in \mathbb{R}^n,\ g_i(x) = b_i\}$. Calculons le gradient de g_i en $x \in \mathbb{R}^n$:

$$\nabla g_i(x) = (\partial_1 g_i(x), \dots, \partial_n g_i(x)) = (a_{i,1}, \dots, a_{i,n}).$$

Ce gradient étant constant, indépendant de $x \in \mathbb{R}^n$, on le notera simplement ∇g_i dans la suite. On peut alors réécrire le système homogène $\mathscr{C}_0$ sous la forme :

$$\mathscr{C}_0 : \left\{ \begin{array}{l} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = 0 \\ \qquad \vdots \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \langle \nabla g_1, x \rangle = 0 \\ \langle \nabla g_2, x \rangle = 0 \\ \qquad \vdots \\ \langle \nabla g_p, x \rangle = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} x \perp \nabla g_1 \\ x \perp \nabla g_2 \\ \quad \vdots \\ x \perp \nabla g_p \end{array} \right.$$

On obtient la description suivante de $\mathscr{H}$.

Propriété 3

On a : $\quad \mathscr{H} = \mathrm{Vect}(\nabla g_1, \dots, \nabla g_p)^{\perp}$.

2.2 Extrema sous contrainte d'égalités linéaires

Théorème 4

Soit f une fonction de classe $\mathscr{C}^1$ sur Ω ouvert, et $\mathscr{C}$ la contrainte d'égalités linéaires définie précédemment.

Si f admet un extremum local en a sous la contrainte $\mathscr{C}$, **alors** $\nabla f(a) \in \mathscr{H}^{\perp}$ et il existe des réels $\lambda_1, \dots, \lambda_p$ tels que :

$$\left\{ \begin{array}{l} a \text{ est solution de } \mathscr{C} \\ \nabla f(a) = \lambda_1 \nabla g_1 + \cdots + \lambda_p \nabla g_p. \end{array} \right. \tag{$**$}$$

Un point a satisfaisant $(**)$ est appelé *point critique sous la contrainte* $\mathscr{C}$, et les réels λ_i sont les *multiplicateurs de Lagrange.*

Preuve. Supposons que f admette un maximum local en a sous la contrainte $\mathscr{C}$:

$$\exists r > 0, \quad \forall x \in \mathscr{B}_o(a,r) \cap \mathscr{C}, \quad f(x) \leq f(a).$$

Fixons $h \in \mathscr{H}$, et posons $g : t \in]-\eta, \eta[\mapsto f(a+th)$, où $\eta = \dfrac{r}{\|h\|}$ est choisi de sorte que $a + th \in \mathscr{B}_o(a,r) \cap \mathscr{C}$ pour tout $t \in]-\eta, \eta[$. Puisque f est de classe $\mathscr{C}^1$ sur Ω, on sait que g est aussi de classe $\mathscr{C}^1$ sur $]-\eta, \eta[$ et on a $g'(t) = \langle \nabla f(a+th), h \rangle$.
D'autre part pour tout $t \in]-\eta, \eta[$, $a + th$ appartient à $\mathscr{C} \cap \mathscr{B}_o(a,r)$. On a donc :

$$f(a+th) \leq f(a), \qquad \text{soit encore} \qquad g(t) \leq g(0).$$

Ceci étant vrai pour tout $t \in]-\eta, \eta[$, la fonction g admet un maximum local en 0. On a donc :

$$g'(0) = 0 \quad \Rightarrow \quad \langle \nabla f(a), h \rangle = 0.$$

Cette égalité étant valable pour tout $h \in \mathscr{H}$, on en déduit que :

$$\nabla f(a) \in \mathscr{H}^\perp = (\mathrm{Vect}(\nabla g_1, \dots, \nabla g_p)^\perp)^\perp = \mathrm{Vect}(\nabla g_1, \dots, \nabla g_p).$$

D'où l'existence de réels $\lambda_1, \dots, \lambda_p$ tels que $\nabla f(a) = \lambda_1 \nabla g_1 + \cdots + \lambda_p \nabla g_p$. □

Mise en garde.

Noter encore une fois qu'il s'agit d'une condition nécessaire d'extremum sous la contrainte $\mathscr{C}$, mais pas suffisante.

Remarque. Comme $\langle \nabla f(a), h \rangle = f'_h(a)$, on en déduit qu'en un point critique a sous la contrainte $\mathscr{C}$, toutes les dérivées directionnelles de f dans une direction appartenant à $\mathscr{H}$ sont nulles.

Méthode.

*Pour déterminer les points critiques de $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ sous une contrainte $\mathscr{C}$ d'**égalités linéaires**, on procèdera comme suit :*

(i) on justifie que f est $\mathscr{C}^1$ et on calcule $\nabla f(x)$.

(ii) on calcule les gradients des fonctions g_i (qui sont des vecteurs constants).

(iii) on résout le système :

$$\begin{cases} \mathscr{C} \\ \partial_1 f(x_1, \dots, x_n) = \lambda_1 \partial_1 g_1(x_1, \dots, x_n) + \cdots + \lambda_p \partial_1 g_p(x_1, \dots, x_p) \\ \quad \vdots \\ \partial_n f(x_1, \dots, x_n) = \lambda_1 \partial_n g_1(x_1, \dots, x_n) + \cdots + \lambda_p \partial_n g_p(x_1, \dots, x_p) \end{cases}$$

à $n+p$ équations et $n+p$ inconnues $(x_1, \dots, x_n, \lambda_1, \dots, \lambda_p)$. Les solutions sont les points critiques de f sous la contrainte linéaire $\mathscr{C}$.

Exercice. Déterminer les points critiques de $f(x_1, \dots, x_n) = \sum_{k=1}^{n} x_k^2$ sous la contrainte $\mathscr{C} : x_1 + \cdots + x_n = n$.

Méthode.

Pour déterminer la nature **locale** *d'un point critique a de f sous une contrainte $\mathscr{C}$ d'égalités linéaires, on pourra procéder comme suit :*

- *on écrit le développement limité de f à l'ordre 2 au voisinage de a. On écrira cette formule pour $x \in \mathscr{C}$, c'est-à-dire de la forme $a+h$ avec $h \in \mathscr{H}$. On notera en particulier que $\langle \nabla f(a), h\rangle = 0$ car $\nabla f(a) \in \mathscr{H}^{\perp}$.*
- *On étudie le signe de $q_a(h)$ pour $h \in \mathscr{H}$.*
  - *Si q_a est de signe constant sur $\mathbb{R}^n$, il l'est aussi sur $\mathscr{H}$ et on conclut à un extremum.*
  - *Sinon, il faut faire une étude plus poussée pour savoir si elle est ou non de signe constant sur $\mathscr{H}$. L'énoncé nous guidera alors.*

Exercice. Déterminer la nature locale des points critiques de $f(x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k^2$ sous la contrainte $\mathscr{C} : x_1 + \cdots + x_n = n$.

Méthode.

Pour étudier la nature globale d'un point critique a de f sous une contrainte $\mathscr{C}$ d'égalités linéaires, on pourra selon les cas :

- *Étudier le signe de $f(a_1+h_1, \ldots, a_n+h_n) - f(a_1, \ldots, a_n)$ pour $(h_1, \ldots, h_n) \in \mathscr{H}$.*
- *Pour $h \in \mathscr{H}$, étudier la fonction d'une variable $g : t \mapsto f(a+th)$. Cette méthode est particulièrement efficace lorsqu'en tout point x de $\mathscr{C}$, la forme quadratique q_x est de signe constant. L'énoncé nous guidera dans ce cas.*

Exercice. Déterminer la nature globale des points critiques de $f(x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k^2$ sous la contrainte $\mathscr{C} : x_1 + \cdots + x_n = n$.