Chapitre 22

Extrema sous contrainte

Compétences attendues.

✓ Reconnaitre la nature topologique d'un ensemble (fermé, borné).

✓ Déterminer les extrema globaux de f sur un fermé borné.

✓ Déterminer les points critiques de f sous une contrainte d'égalités linéaires.

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

Objectif. Dans ce chapitre, nous étudions les extrema d'une fonction sur une partie $\mathscr{C}$ d'un ouvert Ω. Cette partie $\mathscr{C}$, appelée *contrainte*, sera définie par des inégalités larges ou par un système linéaire. Elle ne sera pas ouverte, et nous ne pourrons donc pas appliquer les résultats de calcul différentiel vu dans les chapitres précédents. Nous allons dans ce chapitre développer d'autres méthodes pour l'optimisation sous la contrainte $\mathscr{C}$.

Mathieu Mansuy

Professeur en ECG deuxième année spécialité mathématiques approfondies au Lycée Louis Pergaud (Besançon)

Page personnelle : `mathieu-mansuy.fr/`

E-mail : `mansuy.mathieu@hotmail.fr`

1 Optimisation sous contrainte d'inégalités

1.1 Motivations

Dans cette section, on suppose que la contrainte $\mathscr{C}$ peut se définir à partir d'ensembles de la forme :

$$\{x \in \mathbb{R}^n,\ \varphi(x) \leq a\}$$

où $\varphi : \mathbb{R}^n \to \mathbb{R}$ est une fonction continue sur $\mathbb{R}^n$ et $a \in \mathbb{R}$. On étudie les extrema globaux d'une fonction f définie sur $\mathscr{C}$.

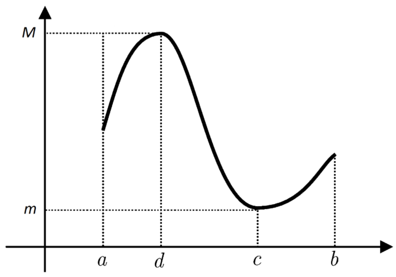
Exemple. On étudie les extrema globaux de la fonction $f : (x,y) \mapsto x^2 + y^2$ sous la contrainte $\mathscr{C} = \{(x,y) \in \mathbb{R}^2,\ x^2 + y^2 \leq 1\}$.

La recherche d'extremum d'une fonction définie sur $\mathscr{C}$ dépend de la nature topologique de l'ensemble $\mathscr{C}$. On peut déjà l'observer lorsque $n = 1$. Prenons par exemple la fonction f définie sur $\mathscr{C} = [-2, 5]$ par $f(x) = x^2$. Sans avoir besoin de faire son tableau de variations, nous pouvons garantir l'existence d'un maximum et d'un minimum de f sur $[-2, 5]$ car f est continue sur $[-2, 5]$ qui est un **segment**. Rappelons en effet un résultat vu en première année.

Théorème 1 (des bornes atteintes)

Soit $f : [a, b] \to \mathbb{R}$ une fonction **continue sur un segment** de $\mathbb{R}$. Alors f est bornée et atteint ses bornes. En d'autres termes, f admet un maximum global et un minimum global sur le segment $[a, b]$, de sorte qu'il existe $c, d \in [a, b]$ tels que :

$$\forall x \in [a, b], \quad f(c) \leq f(x) \leq f(d).$$



Peut-on généraliser ce résultat aux fonctions de plusieurs variables en remplaçant « segment » par autre chose ?

Cela nous amène à introduire la notion de sous-ensemble fermé et borné de $\mathbb{R}^n$.

1.2 Partie fermée

Définition.

Une partie A de $\mathbb{R}^n$ est un *fermé de* $\mathbb{R}^n$ si son complémentaire A^c est un ouvert de $\mathbb{R}^n$.

Exemples.

- $\mathbb{R}^n$ et $\emptyset$ sont des fermés de $\mathbb{R}^n$ (et des ouverts comme on l'a vu).
- Parmi les intervalles de $\mathbb{R}$, **ceux qui sont fermés sont exactement ceux de la forme** $[a, b]$**,** $[a, +\infty[$ **ou** $]-\infty, b]$ **avec** $-\infty < a < b < +\infty$ car :

Remarques.

- Graphiquement, une partie fermée est une partie qui contient son bord.
- Par définition, on a aussi que le complémentaire d'un ouvert est un fermé.

Mise en garde.

Fermé n'est pas le contraire d'ouvert : certaines parties de $\mathbb{R}^n$ peuvent être ni ouvertes ni fermées. Par exemple, l'intervalle $]-1,1]$ n'est ni ouvert ni fermé dans $\mathbb{R}$.

Propriété 2

Soit $\varphi : \mathbb{R}^n \to \mathbb{R}$ une fonction continue. Pour tout $a, b \in \mathbb{R}$, les ensembles

$$\{x \in \mathbb{R}^n,\ \varphi(x) \leq a\}, \quad \{x \in \mathbb{R}^n,\ \varphi(x) \geq a\},$$

$$\{x \in \mathbb{R}^n,\ a \leq \varphi(x) \leq b\} \quad \text{et} \quad \{x \in \mathbb{R}^n,\ \varphi(x) = a\}$$

sont des fermés de $\mathbb{R}^n$.

Remarque. Il faut retenir qu'un ensemble défini par une fonction continue sur $\mathbb{R}^n$ des inégalités larges ou des égalités est un fermé.

Exemple.

- Une boule fermée $\mathscr{B}_f(a,r)$ est une partie fermée de $\mathbb{R}^n$.

  En particulier, pour tout $a \in \mathbb{R}^n$, le singleton $\{a\} = \mathscr{B}_f(a,0)$ est un fermé de $\mathbb{R}^n$.
- Une ligne de niveau λ d'une fonction continue $f : \mathbb{R}^n \to \mathbb{R}$ est un fermé.

Propriété 3

- Une union **finie** de fermés[a] de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.
- Une intersection (finie ou infinie) de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.
- Un produit cartésien de fermés est fermé.

[a] Attention, c'est faux en général pour une union infinie de fermés.

Exemples.

- $[0,1] \cup \{2\}$ est un fermé de $\mathbb{R}$.
- $(\mathbb{R}_+)^2 = \mathbb{R}_+ \times \mathbb{R}_+$ est un fermé de $\mathbb{R}^2$, de même que $[-1,1] \times]-\infty, 3]$.

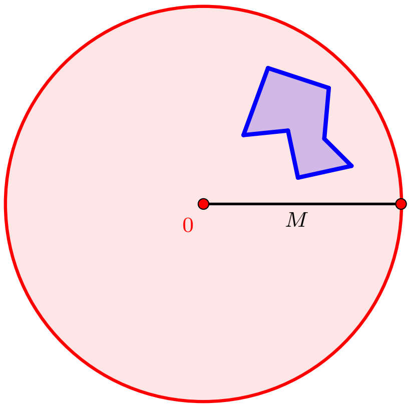
Exercice. Montrer que l'ensemble $A = \{(x, y) \in \mathbb{R}^2,\ x \geq 0,\ y \geq 0,\ x^2 + y^2 \leq 1\}$ est un fermé de $\mathbb{R}^2$.

1.3 Partie bornée

Définition.

Une partie A de $\mathbb{R}^n$ est *bornée* si A est incluse dans une boule fermée de centre 0, c'est-à-dire si :

$$\exists M > 0, \quad A \subset \mathscr{B}_f(0, M).$$



En bleu, une partie bornée de $\mathbb{R}^2$.

Exemples.

- Une boule de $\mathbb{R}^n$, ouverte ou fermée, est bornée.
- Un intervalle de $\mathbb{R}$ est borné si et seulement s'il est majoré et minoré, soit encore si ses bornes sont finies. Ainsi l'intervalle $]-5, 3]$ est une partie bornée de $\mathbb{R}$, mais pas l'intervalle $]0, +\infty[$.
- Si A et B sont des parties de $\mathbb{R}^n$ telles que $A \subset B$ et B bornée, alors A est bornée.

Pour aller plus loin.

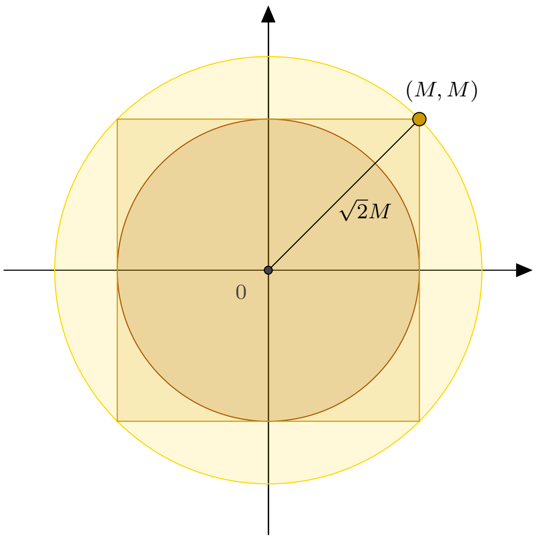
Une partie A de $\mathbb{R}^n$ n'est pas bornée si et seulement s'il existe une suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de A telle que $\lim\limits_{k\to+\infty} \|a_k\| = +\infty$.

Propriété 4

Une partie A de $\mathbb{R}^n$ est bornée si et seulement si A est incluse dans un « cube », c'est-à-dire si :

$$\exists M > 0, \quad A \subset [-M, M]^n.$$

Preuve.



Il suffit de noter que :

$$\mathscr{B}_f(0,M) \subset [-M,M]^n \subset \mathscr{B}_f(0,\sqrt{n}M).$$

Montrons la deuxième inclusion. Pour tout $x \in [-M,M]^n$, on a :

$$\|x\|^2 = \sum_{k=1}^{n} x_k^2 \leq \sum_{k=1}^{n} M^2 = n \times M^2$$

Ainsi si A est inclus dans une boule, il est aussi inclus dans un « cube », et réciproquement.

$\square$

Exercice. Les parties suivantes sont-elles bornées :

- $\mathscr{E} = \{(x,y) \in \mathbb{R}^2,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1\}$.

- $\mathscr{C} = \{(x,y) \in \mathbb{R}^2,\ ye^x = 1\}$.

1.4 Fonction continue sur un fermé borné

On peut maintenant étendre le résultat vu en première année au cas des fonctions de plusieurs variables comme suit.

Théorème 5 (*des bornes atteintes*)

Soit f une fonction **continue sur un fermé borné** $A \subset \mathbb{R}^n$, et à valeurs dans $\mathbb{R}$. Alors f admet un maximum global et un minimum global sur A.

Remarque. Ce théorème garantit l'existence des extrema, mais ne dit pas comment les trouver !

Exercice. Montrer que la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^2 - 2y^2 - 5xy$ admet un maximum global et un minimum global sur $\mathscr{D} = \{(x,y) \in \mathbb{R}^2,\ x \geq 0,\ y \geq 0,\ x^2 + y^2 \leq 1\}$.

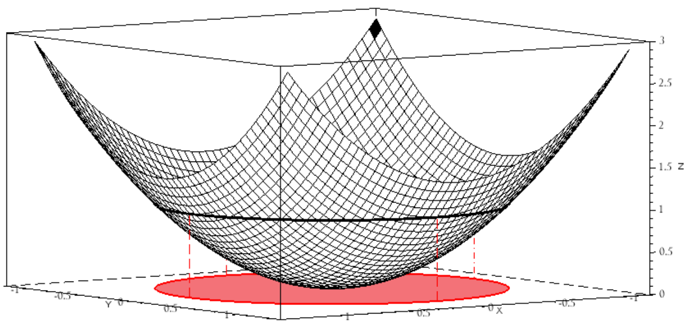
1.5 Recherche d'extrema sur un fermé borné

Méthode. Extrema globaux sur un fermé borné.
Pour étudier les extrema globaux d'une fonction f continue sur un fermé borné $A \subset \mathbb{R}^n$, on procèdera comme suit :

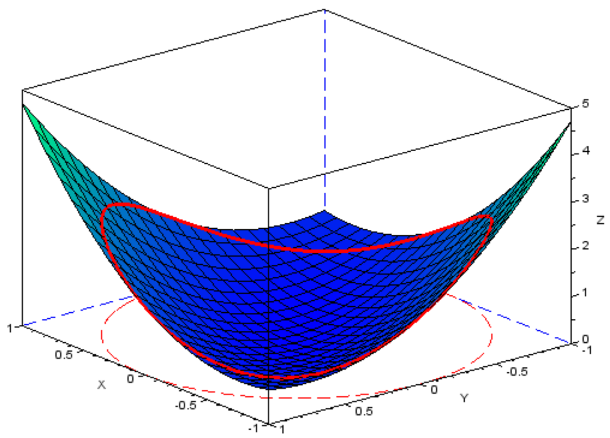
- *on détermine le plus grand ouvert Ω inclus dans A (il s'agit en général de remplacer inégalités larges par inégalités strictes) ;*
- *si f est $\mathscr{C}^1$ sur Ω, on détermine ses points critiques sur Ω ;*
- *on étudie « à la main » les extrema de f sur la frontière $F = A \setminus \Omega$ de A ;*
- *on compare la valeur de f aux points critiques de Ω et aux extrema sur la frontière de A pour conclure.*

Exercice. On considère la fonction $f : (x, y) \mapsto x^2 + y^2$ définie sur la boule fermée $\mathscr{B}_f(0, 1)$.

1. Justifier l'existence d'un maximum et d'un minimum global pour f sur $\mathscr{B}_f(0, 1)$.
2. Déterminer en quels points ces extrema sont atteint.



Exercice. Déterminer les extrema globaux de $f : (x, y) \mapsto \sqrt{1 + x^2 + y^2} + (x + y)^2 - 1$ 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 :

$$\mathscr{S} : \left\{ \begin{array}{r} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \quad \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n = b_p \end{array} \right. .$$

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 $\mathscr{S}$ se récrit :

$$\mathscr{S} : \left\{ \begin{array}{c} g_1(x) = b_1 \\ g_2(x) = b_2 \\ \vdots \\ g_p(x) = b_p \end{array} \right. .$$

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

$$\mathscr{S}_0 : \left\{ \begin{array}{c} g_1(x) = 0 \\ g_2(x) = 0 \\ \vdots \\ g_p(x) = 0 \end{array} \right. .$$

Propriété 6 (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 un élément de $\mathscr{C}$, alors :

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

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

$$\mathscr{C} = \underbrace{x_0}_{\text{sol. part.}} + \underbrace{\mathscr{H}}_{\text{ens. des sol. de l'éq. hom.}}.$$

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 $x = x_0 + h$.

□

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

Pour tout $1 \leq i \leq p$, g_i est polynomiale donc de classe $\mathscr{C}^1$, et on a pour tout $x \in \mathbb{R}^n$:

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

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

$$\mathscr{S}_0 : \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = 0 \\ \quad\vdots \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n = 0 \end{cases} \Leftrightarrow \begin{cases} \langle \nabla g_1, x \rangle = 0 \\ \langle \nabla g_2, x \rangle = 0 \\ \quad\vdots \\ \langle \nabla g_p, x \rangle = 0 \end{cases} \Leftrightarrow \begin{cases} x \perp \nabla g_1 \\ x \perp \nabla g_2 \\ \quad\vdots \\ x \perp \nabla g_p \end{cases}$$

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

Propriété 7

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

Remarque. Dans l'exemple précédent, on a $\nabla g_1 = (1,1,1)$, $\nabla g_2 = (0,1,-1)$, de sorte que $\mathscr{H} = \text{Vect}((1,1,1),(0,1,-1))^\perp = \text{Vect}(-2,1,1)$.

2.2 Extrema sous contrainte d'égalités linéaires

Théorème 8 (*des extrema liés sous contrainte d'égalités linéaires*)

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 :

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

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, g est aussi de classe $\mathscr{C}^1$ sur $]-\eta,\eta[$ et $g'(t) = \langle \nabla f(a+th), h \rangle$ pour tout $t \in]-\eta,\eta[$.
D'autre part pour tout $t \in]-\eta,\eta[$, $a+th$ appartient à $\mathscr{C} \cap \mathscr{B}_o(a,r)$, d'où :

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

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

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

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

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

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

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. $g'(0) = \langle \nabla f(a), h \rangle$ correspond à la dérivée de f en a dans la direction du vecteur h. Ainsi, 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. Détermination des points critiques sous contrainte d'égalités linéaires.

*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 :

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

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

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

Méthode. Nature locale d'un point critique sous contrainte d'égalités linéaires.

*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 $f : (x_1, \dots, x_n) \mapsto \sum_{k=1}^{n} x_k^2$ sous la contrainte $\mathscr{C} : x_1 + \cdots + x_n = n$.

Méthode. Nature globale d'un point critique sous contrainte d'égalités linéaires.

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 $f(a_1 + h_1, \dots, a_n + h_n) - f(a_1, \dots, a_n)$ pour $(h_1, \dots, h_n) \in \mathscr{H}$.*
- *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 $f : (x_1, \dots, x_n) \mapsto \sum_{k=1}^{n} x_k^2$ sous la contrainte $\mathscr{C} : x_1 + \cdots + x_n = n$.