

Optimisation sous contraintes linéaires

Mais dites- moi... jusqu'où s'arrêteront-ils?

COLUCHE
La revue de presse

1 Position du problème

Soient $n \in \mathbb{N} \setminus \{0; 1\}$ et $p \in \mathbb{N}^*$ avec $p < n$. Considérons de plus $g_1, g_2, \dots, g_p, p$ formes linéaires sur $\mathbb{R}^n$. Autrement dit, pour chaque $i \in \llbracket 1; p \rrbracket$, il existe $\lambda_1, \lambda_2, \dots, \lambda_n$ tels que

$$\forall x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n, \quad g_i(x) = \sum_{k=1}^n \lambda_k x_k.$$

Notons $\mathcal{C}$, l'ensemble des solutions du système linéaire

$$(S) \quad \begin{cases} g_1(x) &= b_1 \\ \vdots & \vdots \\ g_p(x) &= b_p, \end{cases}$$

où l'inconnue est un élément x de $\mathbb{R}^n$, et $\mathcal{H}$ l'ensemble des solutions du système homogène associé :

$$(S)_0 \quad \begin{cases} g_1(x) &= 0 \\ \vdots & \vdots \\ g_p(x) &= 0. \end{cases}$$

En notant que $\mathcal{H} = \bigcap_{i=1}^p \text{Ker}(g_i)$, on constate que $\mathcal{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$.

Si $\mathcal{C}$ n'est pas vide, on parle d'espace affine. Dans ce cas, si $x_0 \in \mathcal{C}$ alors on peut écrire

$$\mathcal{C} = \{x_0 + h \in \mathbb{R}^n \mid h \in \mathcal{H}\}.$$

Preuve. À l'aide de la linéarité de g , on a les équivalences

$$x \in \mathcal{C} \iff \begin{cases} g_1(x) &= b_1 \\ \vdots & \vdots \\ g_p(x) &= b_p \end{cases} \iff \begin{cases} g_1(x) &= g_1(x_0) \\ \vdots & \vdots \\ g_p(x) &= g_1(x_0) \end{cases} \iff \begin{cases} g_1(x-x_0) &= 0 \\ \vdots & \vdots \\ g_p(x-x_0) &= 0 \end{cases} \iff x-x_0 \in \mathcal{H}.$$

DÉFINITION

extremum sous contrainte

Soit $f : \mathcal{C} \rightarrow \mathbb{R}$ où $\mathcal{C}$ est une partie de $\mathbb{R}^n$.

- On dit que f admet un **extremum local sous la contrainte** $\mathcal{C}$ en un point a si la restriction de f à $\mathcal{C}$ admet un extremum local en a .
- On dit que f admet un **extremum global sous la contrainte** $\mathcal{C}$ en un point a si la restriction de f à $\mathcal{C}$ admet un extremum global en a .

Remarque. Par exemple, f admet un maximum global sous la contrainte $\mathcal{C}$ en un point a si et seulement si :

$$\forall x \in \mathcal{C}, \quad f(x) \leq f(a).$$

Exemples.

- Une société de production d'épingles possède deux usines avec des fonctionnements différents. La première usine produit x tonne(s) d'épingles pour un coût $C_1(x)$ et la seconde produit la même quantité avec un coût $C_2(x)$ (en euros). Une analyse des coûts donne :

$$C_1(x) = 3 + 4x \quad \text{et} \quad C_2(x) = 4x^2 + 2.$$

Le coût total d'une production d'une quantité x_1 par la première usine et x_2 , pour la seconde est

$$C_T(x_1, x_2) = C_1(x_1) + C_2(x_2) = 5 + 4x_1 + 4x_2^2.$$

Pour une production d'une tonne, il faut rajouter la contrainte linéaire $g(x_1, x_2) = x_1 + x_2 = 1$.

Dans cet exemple, si on souhaite minimiser le coût d'une tonne d'épingles en répartissant la production entre les deux usines, on cherche le minimum global de la fonction C_T sous la contrainte $\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 1\}$.

- *Exemple de microéconomie.*

Considérons un consommateur dont les achats portent sur plusieurs produits, $X_1, X_2, \dots, X_n$. La satisfaction éprouvée par ce consommateur dépend des quantités de X_i dont il peut disposer. Autrement dit, l'utilité qu'il obtient, U , est fonction des quantités consommées des produits considérés; elle peut s'écrire :

$$U = f(x_1, x_2, \dots, x_n)$$

où x_i représente la quantité de biens X_i . Le consommateur rationnel de l'analyse économique désire se procurer des quantités x_i des biens X_i telles qu'il puisse obtenir la plus grande satisfaction possible. Le problème qu'il doit résoudre est celui de la maximisation de sa fonction d'utilité sous la contrainte de son revenu. Notons R est ce revenu et p_i le prix du bien X_i . Ce revenu et ces prix sont supposés s'établir sur des marchés sur lesquels l'individu considéré ne peut exercer aucune influence; autrement dit, R , et p_i sont des constantes données. Admettons également que notre consommateur dépense la totalité de son revenu, au cours de la période considérée, pour acheter. Dans ces conditions, la considère cherche à maximiser la fonction f sous la contrainte linéaire

$$R = \sum_{i=1}^n p_i \cdot x_i.$$

2

Recherche des extrema sous contraintes d'égalités linéaires

2.1

Par substitution

Reprenons l'exemple des épingles. Répartissons la production de la tonne entre les deux usines :

- Soit $x \in [0; 1]$, la production de l'usine 1.
- La production de la seconde usine est $y = 1 - x$.

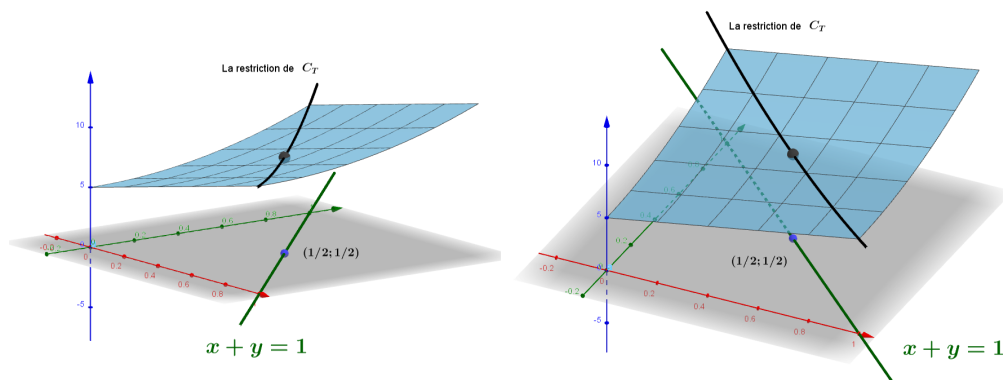
Le coût total est alors :

$$C(x) = C_1(x) + C_2(1 - x) = 3 + 4x + 4(1 - x)^2 + 2 = 9 - 4x + 4x^2.$$

La fonction C est polynomiale du second degré. Elle est donc dérivable sur $\mathbb{R}$ avec pour tout $x \in [0; 1]$, $C'(x) = -4 + 8x = 8(x - 1/2)$. La fonction C est décroissante sur l'intervalle $[0; 1/2]$ et croissante sur $[1/2; 1]$. On a donc un minimum en $x = 1/2$, le coût minimal est alors :

$$C_{\min} = C(1/2) = 8.$$

En conclusion, on minimise le coût de production en répartissant équitablement la production de la tonne entre les deux usines.



Exercice 1



1. Déterminer le minimum de $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 + 10$ sous la contrainte $3x + y = 10$.
2. Tester la pertinence du résultat à l'aide du code Python ci-dessous.

p. 14

OC1

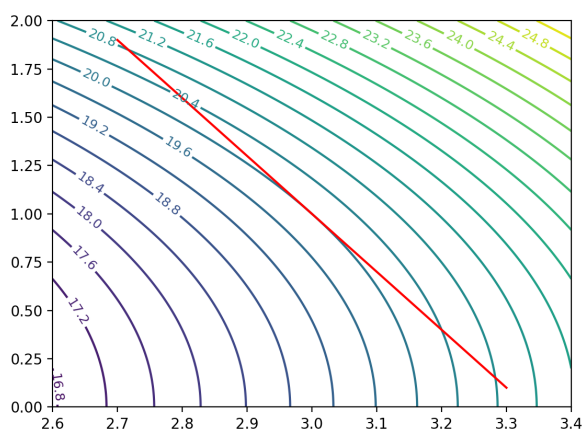
Editeur

```
import numpy as np
import matplotlib.pyplot as plt

x=np.linspace(2.6,3.4,200)
y=np.linspace(0,2,200)
X , Y = np.meshgrid(x,y)
Z = X**2+Y**2+10

graphe = plt.contour(X,Y,Z,25)
plt.clabel(graphe,inline=1,fontsize=9)

x=np.array([2.7,3.3])
plt.plot(x,-3*x+10,'r')
plt.show()
```



Exercice 2



Soit f la fonction définie par :

D'après EDHEC 2020

$$\forall (x, y, z) \in \mathbb{R}^3, \quad f(x, y, z) = xe^{x(y^2+z^2+1)}$$

p. 14

On souhaite étudier les extrema de f sous la contrainte linéaire : $\begin{cases} x = 1 \\ y + z = 0 \end{cases}$.

Montrer que, sous la contrainte, f présente un minimum global au point $(1, 0, 0)$. Quelle est sa valeur?

OC2

2.2 Par du calcul différentiel

Les énoncés

THÉORÈME

condition nécessaire d'ordre 1

Soient $f : \mathcal{O} \rightarrow \mathbb{R}$ où $\mathcal{O}$ est une partie ouverte de $\mathbb{R}^n$ et $a \in \mathcal{O}$.

Si $\left\{ \begin{array}{l} \rightarrow f \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathcal{O}. \\ \rightarrow f \text{ admet un extremum sous la contrainte } \mathcal{C} \text{ en un point } a. \end{array} \right.$

Alors pour tout $h \in \mathcal{H}$, la dérivée de f en a dans la direction h est nulle, c'est-à-dire

$$\forall h \in \mathcal{H}, \quad \langle \nabla f(a), h \rangle = 0.$$

Autrement dit, le gradient $\nabla f(a)$ appartient à $\mathcal{H}^\perp$.

Preuve. Soit $h \in \mathcal{H}$, $h \neq 0_{\mathbb{R}^n}$. Comme $\mathcal{O}$ est une partie ouverte, il existe $r \in \mathbb{R}_*^+$ tel que

$$\mathcal{B}(a, r) \subset \mathcal{O}.$$

On peut alors définir la fonction

$$g_{a,h} : t \in]-r; r[\mapsto f(a + th).$$

Précisons que $g_{a,h}$ est définie sur un intervalle ouvert de $\mathbb{R}$ et pour tout $t \in]-r; r[$, $a + th \in \mathcal{C}$. La fonction f est de classe $\mathcal{C}^1$, on sait (voir proposition sur les dérivées directionnelles) que $g_{a,h}$ l'est aussi sur $]-r; r[$ avec

$$\forall t \in]-r; r[, \quad g_{a,h}'(t) = \langle \nabla f(a + th), h \rangle.$$

avec en particulier

$$g_{a,h}'(0) = \langle \nabla f(a), h \rangle.$$

Or a est un extremum local sous contrainte pour f , par conséquent, 0 est un extremum de la fonction $g_{a,h}$. Nécessairement

$$g_{a,h}'(0) = 0$$

D'où le résultat. ■

Vocabulaire. Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, a un point de $\mathcal{O}$. Si ∇f_a appartient à $\mathcal{H}^\perp$, on dit que a est un **point critique de f sous contrainte $\mathcal{C}$** .

Remarque. Avec les notations du début de chapitre, pour tout $i \in \llbracket 1; p \rrbracket$, g_i est de classe $\mathcal{C}^1$ avec pour tout $x \in \mathcal{O}$

$$\nabla g_i(x) = (\lambda_1, \lambda_2, \dots, \lambda_n) = \text{cste}.$$

On notera donc simplement ∇g_i (on confond ici la constante avec la fonction). En particulier, on a

$$\forall x \in \mathbb{R}^n, \quad g_i(x) = \langle \nabla g_i, x \rangle.$$

LEMME

famille génératrice de l'orthogonal

Avec les notations introduites précédemment,

$$\mathcal{H}^\perp = \text{Vect}(\nabla g_1, \dots, \nabla g_p)$$

Preuve. Soit $x \in \mathbb{R}^n$. Raisonnons par équivalence.

$$\begin{aligned} x \in \mathcal{H} &\iff \forall i \in \llbracket 1; p \rrbracket, \quad g_i(x) = 0 \\ &\iff \forall i \in \llbracket 1; p \rrbracket, \quad \langle \nabla g_i, x \rangle = 0 \\ &\iff \forall y \in \text{Vect}(\nabla g_1, \dots, \nabla g_p), \quad \langle y, x \rangle = 0 \\ &\iff x \in \text{Vect}(\nabla g_1, \dots, \nabla g_p)^\perp. \end{aligned}$$

On a donc l'égalité

$$\mathcal{H} = \text{Vect}(\nabla g_1, \dots, \nabla g_n)^\perp.$$

Considérons les supplémentaires orthogonaux

$$\mathcal{H}^\perp = \left(\text{Vect}(\nabla g_1, \dots, \nabla g_p)^\perp \right)^\perp = \text{Vect}(\nabla g_1, \dots, \nabla g_p).$$

La dernière égalité résultant du fait qu'en dimension finie $(F^\perp)^\perp = F$ pour tout sous-espace vectoriel F .

THÉORÈME

condition d'ordre 1

Soient f une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ et $a \in \mathcal{O}$.

Si $\left\{ \begin{array}{l} \rightarrow f \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathcal{O}. \\ \rightarrow f \text{ admet un extremum sous la contrainte } \mathcal{C} \text{ en un point } a. \end{array} \right.$

Alors $\text{il existe } (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p \text{ tels que } \nabla f(a) = \sum_{i=1}^p \lambda_i \nabla g_i.$

Preuve. C'est une conséquence directe des deux énoncés précédents.

Exemples. Reprenons les fonctions définies aux exercices 1 et 2.

- On a une seule contrainte $g(x, y) = 3x + y = 10$ avec $\nabla g = (3, 1)$. De plus, on a vu que le minimum est atteint en $a = (3, 1)$ avec

$$\nabla f(a) = (2a_1, 2a_2) = 2\nabla g(3, 1).$$

- On a ici deux contraintes : $g_1(x, y, z) = x = 1$ et $g_2(x, y, z) = y + z = 0$. D'où

$$\nabla g_1(x, y, z) = (1, 0, 0) \quad \nabla g_2 = (0, 1, 1).$$

On a vu que le point critique est : $a = (1, 0, 0)$ et

$$\partial_1 \nabla f(x, y, z) = (1 + (y^2 + z^2 + 1)x) e^{x(y^2 + z^2 + 1)}$$

$$\partial_2 f(x, y, z) = 2x^2 y e^{x(y^2 + z^2 + 1)}$$

$$\partial_3 f(x, y, z) = 2x^2 z e^{x(y^2 + z^2 + 1)}.$$

On retrouve bien

$$\nabla f(a) = (2e, 0, 0) = 2e \cdot \nabla g_1 + 0 \cdot \nabla g_2.$$

Vocabulaire. Si la famille $(\nabla g_1, \dots, \nabla g_p)$ est libre, les réels $\lambda_1, \dots, \lambda_p$ sont uniques. Ils sont appelés multiplicateurs de Lagrange.

Exercice 3



◆ Déterminer les extrema possibles de la fonction f définie sur $\mathbb{R}^3$ par $f(x, y, z) = xyz$ sous la contrainte

$$\mathcal{C}: \begin{cases} x + y + z = 0 \\ x - y + z = 2. \end{cases} \quad \text{p. 14}$$

OC3

! Attention. La condition précédente n'est qu'une condition nécessaire et non suffisante. Elle ne fournit que des candidats pour les extrema et il convient ensuite de poursuivre l'analyse.

- Si l'on recherche un extremum *global*, on peut revenir à la définition et étudier le signe de

$$f(x) - f(a) \quad \text{pour } x \in \mathcal{C}$$

ou, à l'aide du changement de variable $h = x - a$, on regarde le signe de

$$f(a + h) - f(a) \quad \text{pour } h \in \mathcal{H}.$$

- Si on cherche un extremum *local*, on peut revenir au développement limité d'ordre 2. Notons q_a , la forme quadratique de f en a .

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h)$$

avec $\varepsilon(h) \xrightarrow{h \rightarrow 0} 0$ et $h \in \mathcal{H}$. Dans le cas d'un point critique sous la contrainte $\mathcal{C}$, on obtient la simplification :

$$\forall h \in \mathcal{H}, \quad f(a+h) - f(a) = \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

On regarde ensuite le signe de la forme quadratique (restreinte à $\mathcal{H}$).

- Si $q_a(h) \geq 0$ pour tout $h \in \mathcal{H}$, on a un minimum local.
 - Si $q_a(h) \leq 0$ pour tout $h \in \mathcal{H}$, on a un maximum local.
 - Sinon, on ne peut pas conclure.
- Dernier argument : se souvenir qu'il existe toujours un maximum et un minimum pour une fonction continue sur un fermé borné.

Exemples d'études globales

Exemples. Nous en donnons deux.

- Soient $f : (x, y, z) \in \mathbb{R}^3 \mapsto x^2 + y^2 + z^2 \in \mathbb{R}$ et $g : (x, y, z) \in \mathbb{R}^3 \mapsto x + y + z \in \mathbb{R}$.

Déterminons les extrema globaux de f sous la contrainte $\mathcal{C} : g(x, y, z) = 3$.

- *Étude du ou des points critiques sous contrainte.*

La fonction f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$ car polynomiale et

$$\forall (x, y, z) \in \mathbb{R}^2, \quad \nabla f(x, y, z) = (2x, 2y, 2z) = 2(x, y, z).$$

D'après les conditions d'ordre 1, $a = (x, y, z)$ donne un extremum sous la contrainte $g(x, y, z) = 3$ si

$$\exists \lambda \in \mathbb{R}, \quad \nabla f(a) = \lambda \nabla g = \lambda(1, 1, 1)$$

En particulier $x = y = z$. Puis $g(a) = 3$ donne $x = y = z = 1$. En particulier, on a $f(a) = 3$.

- *Étude locale au niveau du point critique.*

On vérifie que

$$\nabla f(a) = (2, 2, 2) \quad \text{et} \quad \nabla^2 f(a) = 2I_3.$$

En particulier, la forme quadratique devient

$$q_a(h) = {}^t H \nabla^2 f(a) H = 2 {}^t H H = 2 \|h\|^2 \geq 0.$$

Puis, pour tout $h \in \mathcal{H}$ (avec $\nabla f(a) \in \mathcal{H}^\perp$)

$$f(a+h) - f(a) = \underbrace{\langle \nabla f(a), h \rangle}_{=0} + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h)$$

On a pour h suffisamment petit $f(a+h) - f(a) \geq 0$. Le point a donne un minimum local sous la contrainte.

- *Étude globale*

Soit $X \in \mathbb{R}^3$ tel que $g(X) = 3$. On écrit

$$X = a + H = (1 + h_1; 1 + h_2; 1 + h_3).$$

La condition $g(X) = 3$ impose $h_1 + h_2 + h_3 = 0$ ou encore $h_3 = -h_1 - h_2$. Puis

$$\begin{aligned} f(X) - f(a) &= (1 + h_1)^2 + (1 + h_2)^2 + (1 + h_3)^2 - 3 \\ &= h_1^2 + 2h_1 + h_2^2 + 2h_2 + h_3^2 + 2h_3 \\ &= h_1^2 + 2h_1 + h_2^2 + 2h_2 + (h_1 + h_2)^2 - 2(h_1 + h_2) \\ f(X) - f(a) &= h_1^2 + h_2^2 + (h_1 + h_2)^2 \geq 0 \end{aligned}$$

Ce calcul étant valable pour tout $X \in \mathbb{R}^3$ vérifiant $g(X) = 3$ (soit, tout $X \in \mathcal{C}$), on a bien un minimum global sous la contrainte $\mathcal{C}$.

- On considère

$$f : x = (x_1, \dots, x_n) \in (\mathbb{R}_+)^n \mapsto \prod_{k=1}^n x_k \quad \text{et} \quad g : x \in \mathbb{R}^n \mapsto \sum_{k=1}^n x_k.$$

Étudions les extrema de f sous la contrainte $g(x) = c$ où $c \in \mathbb{R}_*^+$.

→ *Étude du ou des points critiques sous contrainte.*

En tant que fonction polynomiale, f est de classe $\mathcal{C}^1$ sur $(\mathbb{R}_*^+)^n$ avec pour tout $x \in (\mathbb{R}_*^+)^n$

$$\partial_i f(x) = \prod_{\substack{k=1 \\ k \neq i}}^n x_k = \frac{f(x)}{x_i}.$$

De plus, $\nabla g = (1, 1, \dots, 1)$.

Sur l'ouvert $(\mathbb{R}_*^+)^n$, si le point a donne un extremum de f sous la contrainte $g(x) = c$ alors il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(a) = \lambda \nabla g = (\lambda, \dots, \lambda).$$

Autrement dit, toutes les composantes de $\nabla f(a)$ sont identiques

$$\partial_1 f(a) = \partial_2 f(a) = \dots = \partial_n f(a) (= \lambda).$$

Soient $i, j \in \llbracket 1; n \rrbracket$

$$a_i \partial_i f(a) = \prod_{k=1}^n a_k = a_j \partial_j f(a)$$

Notons que $\lambda \neq 0$. D'où $a_i = a_j$. Et donc $a = (a_1, a_1, a_1, \dots, a_1)$. La condition $g(a) = c$ donne

$$a = \frac{c}{n} (1, 1, 1, \dots, 1)$$

et $f(a) = \left(\frac{c}{n}\right)^n$. En résumé, il existe un unique point critique sur $(\mathbb{R}_*^+)^n$ sous la contrainte $g(x) = c$.

- *Nature du point critique.*

Notons que la partie de $\mathbb{R}^n$ définie par

$$\mathcal{K}_c = (\mathbb{R}^+)^n \cap \{g(x) = c\}$$

est une partie fermée en tant qu'intersection de deux fermés. De plus, elle est bornée puisque $\mathcal{K}_c \subset [0; c]^n$. La fonction f étant continue sur ce fermé, bornée, elle y admet un maximum et un minimum. Notons de plus que si x appartient au "bord", c'est-à-dire si

$$x \in \mathcal{K}_c \setminus ((\mathbb{R}_*^+)^n \cap \{g(x) = c\})$$

alors l'une des composantes de x est nulle et $f(x) = 0$. Or, la fonction f est à valeurs dans $\mathbb{R}^+$, donc chaque point du bord permet d'avoir un minimum. Le maximum ne peut donc être atteint en un point du bord (f n'est pas la fonction nulle), il est donc atteint sur

$$\underbrace{(\mathbb{R}_*^+)^n}_{\text{ouvert}} \cap \{g(x) = c\}.$$

On sait alors que le maximum est atteint en un point critique sous contrainte. Or, il n'existe qu'un seul point critique sous contrainte (a). On a donc montré que $f(a)$ est le maximum de f sous la contrainte :

$$\forall x \in \mathcal{C}, \quad f(x) = \prod_{k=1}^n x_k \leq f(a) = \left(\frac{c}{n}\right)^n.$$

Formule que l'on réécrit sous la forme

$$\left(\prod_{k=1}^n x_k\right)^{1/n} \leq \frac{1}{n} \sum_{k=1}^n x_k.$$

C'est l'inégalité arithmético-géométrique.

Exercice 4



♦ Soit f définie sur $\mathbb{R}^3$ par

$$f(x, y, z) = x^2 - 2xy + yz + y - z.$$

Vérifier que $a = \left(\frac{4}{5}, \frac{3}{5}, \frac{1}{5}\right)$ est l'unique point critique sous la contrainte

$$\begin{cases} 2x - y = 1 \\ x + z = 1. \end{cases}$$

Est-ce un maximum ou minimum global? ni l'un ni l'autre?

p. 14

OC4

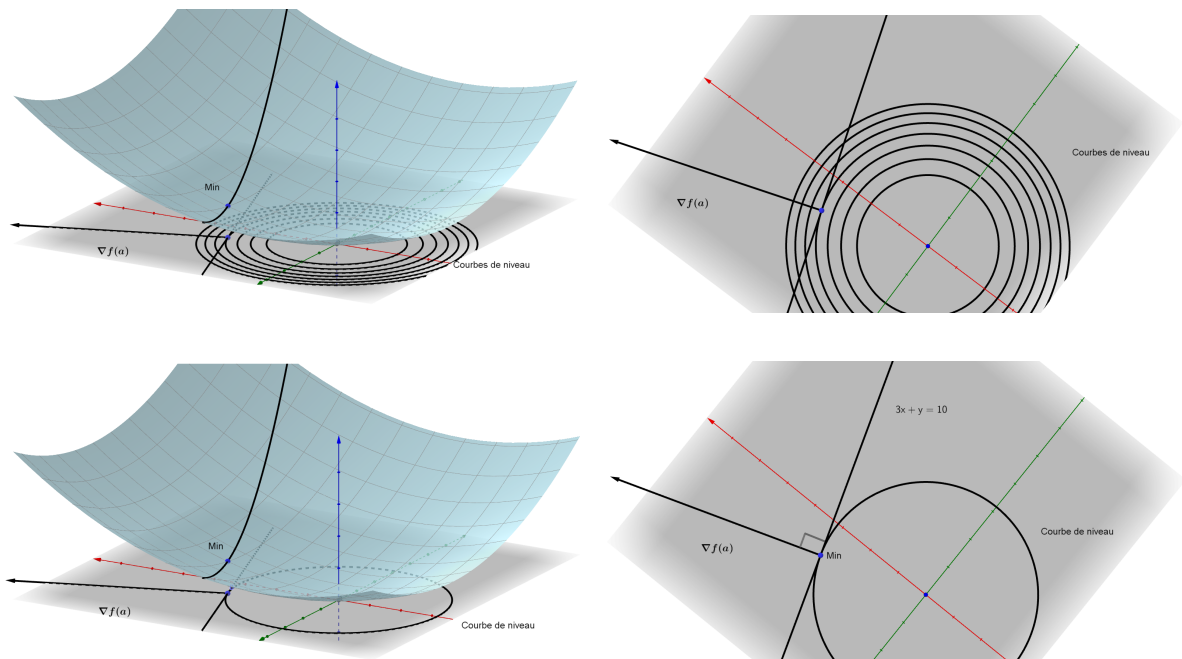
2.3 Interprétation géométrique avec une seule contrainte linéaire

Dans le cas d'une fonction de deux variables avec une seule contrainte linéaire, on cherche à minimiser la restriction de f à une droite d'équation $g(x, y) = \lambda x + \mu y = b$. ∇g est un vecteur normal à la droite $\mathcal{C}$. Or, on sait que :

- Si un extremum sous contrainte est atteint en a , $\nabla f(a)$ est colinéaire à ∇g . En particulier, $\nabla f(a)$ est un vecteur normal à la droite $\mathcal{C}$.
- Le vecteur $\nabla f(a)$ est aussi orthogonal à la ligne de niveau $\mathcal{L}_{f(a)}$.

On en déduit que dans le cas d'un extremum sous contrainte, la droite qui représente la contrainte est tangente à la ligne de niveau. Autrement dit, elle ne peut la franchir.

C'est ce qu'on constate avec l'exemple de la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 + 10$ sous la contrainte $3x + y = 10$.



Un second exemple est donné par la représentation Python de la question 2 de l'exercice 1, page 3.

3

Compléments sur les fonctions définies sur des fermés bornés

Commençons par résumer la démarche avant de l'appliquer à différents exemples.

Existence et calcul des extrema d'une fonction définie sur un fermé borné

Dans la suite, on considère une fonction f de classe $\mathcal{C}^1$ sur un ensemble fermé borné, noté $\mathcal{K}$.

→ I. Existence

On vérifie que l'ensemble de départ de f est bien une partie :

- fermé, par exemple, en montrant qu'il existe φ , une fonction de $\mathbb{R}^n$ et $\mathcal{C}^0$, telle que

$$\mathcal{K} = \{x \in \mathbb{R}^n \mid \varphi(x) \leq c\}.$$

- bornée, en montrant qu'il existe un réel M tel que pour tout $x \in \mathcal{K}$, $\|x\| \leq M$.

On sait alors que f admet un maximum global et un minimum global sur $\mathcal{K}$, mais on ne sait pas en quels points ils sont atteints.

→ II. Étude sur l'ouvert

Soit $\mathcal{O}$, le plus grand ouvert inclus dans $\mathcal{K}$. Autrement dit, $\mathcal{O}$ correspond à $\mathcal{K}$ privé de ses bords. On étudie alors les points critiques pour avoir des points candidats possibles où les extrema sont atteints.

→ III. Étude sur le bord

On étudie maintenant la fonction sur le bord $\mathcal{K} \setminus \mathcal{O}$. On retrouve le cas d'une recherche d'extremum sous contrainte.

→ IV. Conclusion

On compare les différentes valeurs de f aux points obtenus sur l'ouvert et sur le bord pour déterminer le maximum et le minimum global.

Exemple. Considérons la fonction f définie sur $\mathbb{R}^2$ par

$$f : (x, y) \mapsto 3x^2 + 2y^2 - 4x - 3y + 4xy + 3.$$

On souhaite étudier le minimum global de la restriction de f à l'ensemble

$$\mathcal{K} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y \leq 1\}.$$

→ I. Existence

L'ensemble $\mathcal{K}$ est délimité par les droites d'équation $x = 0$, $y = 0$ et $x + y = 1$. C'est un fermé. Par exemple, c'est l'intersection de trois fermés :

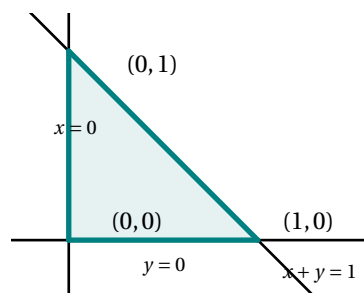
$$\mathbb{R}^+ \times \mathbb{R}, \quad \mathbb{R} \times \mathbb{R}^+ \quad \text{et} \quad \{(x, y) \in \mathbb{R}^2 \mid x + y \leq 1\}.$$

La partie $\mathcal{K}$ est aussi bornée. Par exemple

$$\mathcal{K} \subset \overline{\mathcal{B}}(0; 1).$$

C'est-à-dire $\forall x \in \mathcal{K}, \quad \|x\| \leq 1$.

La fonction f étant continue car polynomiale, il existe donc un maximum et un minimum sur $\mathcal{K}$.



→ II. Étude sur l'ouvert

Soit $\mathcal{O} = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < 1\}$. La fonction f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$ car polynomiale et

$$\partial_1 f(x, y) = 6x - 4 + 4y \quad \text{et} \quad \partial_2 f(x, y) = 4y - 3 + 4x.$$

Vérifier qu'on a un unique point critique donné par $a = (\frac{1}{2}; \frac{1}{4}) \in \mathcal{O}$ avec $f(a) = 13/8 = 1,625$.

→ III. Étude sur le bord

- Cas $y = 0$: $f(x, 0) = 3x^2 - 4x + 3$. On a un minimum valant $5/3 \approx 1,67$.

- Cas $x = 0$, $f(0, y) = 2y^2 - 3y + 3$. On a un minimum valant $15/8 = 1,875$.

- Cas $x + y = 1$.

Posons $g : (x, y) \in \mathbb{R} \mapsto x + y$. On étudie donc f sous la contrainte $g(x, y) = 1$. Étudions les points critiques sous cette contrainte. Il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

Dans ce cas, on a

$$\exists \lambda \in \mathbb{R}, \quad \begin{cases} 6x + 4y - 4 = \lambda \\ 4x + 4y - 3 = \lambda \end{cases} \quad \xLeftrightarrow_{L_1 - L_2} \quad \exists \lambda \in \mathbb{R}, \quad \begin{cases} 2x - 1 = 0 \\ 4x + 4y - 3 = \lambda \end{cases}$$

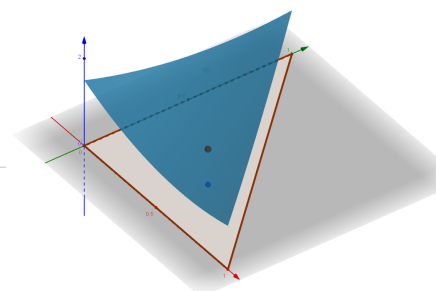
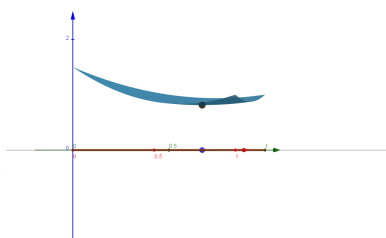
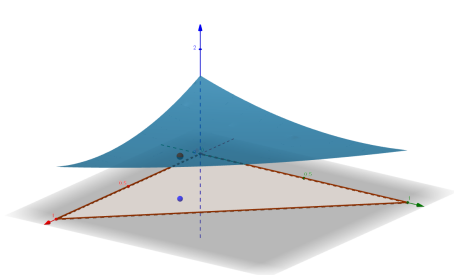
On obtient $x = 1/2$. La condition $x + y = 1$ donne en plus, $y = 1/2$.

$f(x, 1 - x)$ n'étant pas majoré, on a un minimum en $(1/2; 1/2)$ sous la contrainte $g(x, y) = 1$. Dans ce cas, on trouve $f(1/2; 1/2) = \frac{7}{4} = 1,75$.

→ IV. Conclusion

En comparant les valeurs des minima obtenus sur chacun des sous-ensembles de $\mathcal{K}$, le minimum global est atteint pour :

$$a = \left(\frac{1}{2}; \frac{1}{4}\right) \quad \text{et} \quad f(a) = 1,675.$$



Exercice 5



◆◆ Exemple

Soit $D = \{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq y \leq 1\}$ et f la fonction définie sur D par

$$f(x, y) = (y - x)^2 + 6xy.$$

p. 15

1. La fonction f admet-elle des extrema?
2. Déterminer les points critiques de f . Déterminer les extrema de f .

OC5

Exercice 6



◆◆ Exemple

Soient $D = [-1; 1]^2$ et f la fonction définie sur D par

$$f(x, y) = (y - x)^2 + 6xy.$$

p. 15

1. Justifier que la fonction f admet un maximum et un minimum.
2. 🔍 Vérifier que f a un unique point critique. Préciser sa nature.
3. 🔍 Donner le maximum et minimum de f .

OC6

Toute bonne chose a une fin, sauf le saucisson qui en a deux.

PROVERBE DANOIS



Exercices



Exercice 7. ♦ Soit f définie sur $\mathcal{O} = (\mathbb{R}_+^*)^3$ par $f(x, y, z) = x \ln x + y \ln y + z \ln z$.

OC7

- Justifier soigneusement que f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$.
- Soit $c \in \mathbb{R}_+^*$ et $\mathcal{C} = \{f(x, y, z) \in \mathcal{O} \mid x + y + z = 3c\}$. Montrer que f admet un unique point critique sous la contrainte $\mathcal{C}$.
- Justifier que ce point critique correspond à un minimum local sous contrainte pour f .

>> Solution p. ??

Exercice 8. ♦♦

D'après oraux ESCP # OC8

Soit $n \in \mathbb{N} \setminus \{0; 1\}$ et f la fonction définie sur $(\mathbb{R}_+^*)^n$ par $f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^4$.

Étudier les extrema locaux et globaux de f sous la contrainte $\sum_{i=1}^n x_i = n$.

>> Solution p. ??

Exercice 9. ♦♦ Soient $n \in \mathbb{N}^*$, $\alpha_1, \dots, \alpha_n$ des réels strictement positifs tels que $\sum_{i=1}^n \alpha_i = 1$. Soient f et g définies sur $(\mathbb{R}^+)^n$ par # OC9

$$f(x_1, \dots, x_n) = \prod_{i=1}^n x_i^{\alpha_i} \quad \text{et} \quad g(x_1, \dots, x_n) = \sum_{i=1}^n \alpha_i x_i.$$

Posons de plus $\mathcal{K} = \{(x_1, \dots, x_n) \in (\mathbb{R}^+)^n \mid g(x_1, \dots, x_n) = 1\}$.

- Déterminer les points critiques de f sur $(\mathbb{R}_+^*)^n$ sous la contrainte $g(x_1, \dots, x_n) = 1$.
- Montrer que f possède un maximum M sur $\mathcal{K}$ et que celui-ci est atteint sur $\mathcal{K} \cap (\mathbb{R}_+^*)^n$.
- En déduire que

$$\forall (x_1, \dots, x_n) \in (\mathbb{R}^+)^n, \quad \prod_{i=1}^n x_i^{\alpha_i} \leq \sum_{i=1}^n \alpha_i x_i.$$

>> Solution p. ??

Exercice 10. ♦♦

d'après EDHEC 2001 # OC10

On désigne par n et r deux entiers naturels vérifiant : $n \geq 2$ et $r \geq 3$. On considère une épreuve aléatoire pouvant aboutir à r résultats différents $R_1, R_2, \dots, R_r$ de probabilités respectives $x_1, x_2, \dots, x_r$. On admet que, pour tout i de $\llbracket 1, r \rrbracket$, $0 < x_i < 1$. On effectue n épreuves indépendantes du type de celle décrite ci-dessus.

Pour tout i de $\llbracket 1, r \rrbracket$, on note X_i la variable aléatoire qui vaut 1 si le résultat numéro i n'est pas obtenu à l'issue de ces n épreuves et qui vaut 0 sinon.

On désigne par X la variable aléatoire égale au nombre de résultats qui n'ont pas été obtenus à l'issue des n épreuves.

- Exprimer la variable X en fonction des variables $X_1, X_2, \dots, X_r$. Donner la loi de X_i pour tout i de $\{1, 2, \dots, r\}$.
- En déduire que l'espérance de X est

$$E(X) = \sum_{i=1}^r (1 - x_i)^n.$$

- En optimisant sous la contrainte $\sum_{i=1}^r x_i = 1$, justifier l'existence et donner la valeur du minimum local de $E(X)$.

>> Solution p. ??

Exercice 11. ♦♦♦ Optimisation et estimateur

OC11

Soit $n \in \mathbb{N}^*$. On désigne par $(X_1, \dots, X_n)$ un n -échantillon d'une loi de Poisson de paramètre λ inconnu.

Montrer que parmi toutes les combinaisons linéaires des $X_1, \dots, X_n$, il existe un unique estimateur T_n sans biais de λ et qui est de variance minimale. Reconnaitre cet estimateur.

>> Solution p. ??

Exercice 12. ♦♦♦ Entropie dans le cas discret

d'après ESCP 2021, num 2.15 # OC12

Soit un entier $n \in \mathbb{N}^*$. Pour toute variable aléatoire Z définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ qui prend n valeurs réelles distinctes $z_1, \dots, z_n$ (i.e. $Z(\Omega) = \{z_1, \dots, z_n\}$) et $p_k = \mathbf{P}(Z = z_k) \neq 0$ pour tout $k \in \llbracket 1, n \rrbracket$, on appelle entropie de Z , le nombre réel défini par :

$$H(Z) = - \sum_{k=1}^n p_k \ln(p_k).$$

Pour tout $x = (x_1, \dots, x_n) \in]0, 1[^n$, on note $h_n(x) = h_n(x_1, \dots, x_n) = - \sum_{k=1}^n x_k \ln(x_k)$.

1. Calculer l'entropie d'une variable aléatoire U qui suit une loi uniforme sur un ensemble $\{u_1, \dots, u_n\}$.
2. Justifier que la fonction h_n est de classe $\mathcal{C}^2$ sur l'ouvert $]0, 1[^n$ et calculer en tout point $x \in]0, 1[^n$, le gradient $\nabla h_n(x)$ et la matrice hessienne $\nabla^2(h_n)(x)$ de h_n .
3. Montrer que la condition nécessaire d'ordre 1 pour un extremum de la fonction h_n est vérifiée en un unique point critique x^* que l'on précisera.
4. Fixons $x \in]0, 1[^n$ tel que $x_1 + \dots + x_n = 1$ et notons $u = x - x^*$.
 - a) Vérifier que, pour tout $t \in [0, 1]$, on a : $x^* + tu \in]0, 1[^n$.
On note alors ψ la fonction définie sur $[0, 1]$ par : $\psi(t) = h_n(x^* + tu)$.
 - b) En utilisant la formule de Taylor avec reste intégral à l'ordre 1 pour ψ entre les points 0 et 1, montrer que h_n admet en x^* un maximum global sous la contrainte $x_1 + \dots + x_n = 1$. Ce maximum est-il atteint en d'autres points que x^* ?
5. Parmi les variables aléatoires prenant n valeurs (chacune avec une probabilité non nulle), quelles sont les lois de celles qui ont la plus grande entropie ?

>> Solution p. ??

Exercice 13. ♦ Soit f définie sur $\mathcal{O} = (\mathbb{R}_+^*)^3$ par $f(x, y, z) = x \ln x + y \ln y + z \ln z$. Pour tout $c \in \mathbb{R}$, on pose

OC13

$$\mathcal{C}_c = \{(x, y, z) \in \mathcal{O} \mid x + y + z = 3c\}.$$

1. Vérifier que f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$, puis montrer que f admet un unique point critique sous la contrainte $\mathcal{C}_c$.
2. Justifier que ce point critique correspond à un minimum local sous la contrainte $\mathcal{C}_c$.

>> Solution p. 15

Exercice 14. ♦♦ ♣ Soient $n \in \mathbb{N} \setminus \{0; 1\}$ et f la fonction définie sur $(\mathbb{R}_+^*)^n$ par

OC14

$$f(x_1, \dots, x_n) = \sum_{k=1}^n x_k^4.$$

Étudier les extrema locaux et globaux de f sous la contrainte $\sum_{i=1}^n x_i = n$.

>> Solution p. 15

Exercice 15. ♦♦ Variante de l'inégalité arithmético-géométrique

OC15

Soient $n \in \mathbb{N}^*$, $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+^*$ tels que $\sum_{i=1}^n \alpha_i = 1$. On définit en plus les fonctions f et g sur $(\mathbb{R}_+^*)^n$ par

$$f(x_1, \dots, x_n) = \prod_{i=1}^n x_i^{\alpha_i} \quad \text{et} \quad g(x_1, \dots, x_n) = \sum_{i=1}^n \alpha_i x_i.$$

Posons de plus $\mathcal{K} = \{(x_1, \dots, x_n) \in (\mathbb{R}_+^*)^n \mid g(x_1, \dots, x_n) = 1\}$.

1. Déterminer les points critiques de f sur $(\mathbb{R}_+^*)^n$ sous la contrainte $g(x_1, \dots, x_n) = 1$.
2. Montrer que f possède un maximum M sur $\mathcal{K}$ et que celui-ci est atteint sur $\mathcal{K} \cap (\mathbb{R}_+^*)^n$.
3. ♣ En déduire que

$$\forall (x_1, \dots, x_n) \in (\mathbb{R}_+^*)^n, \quad \prod_{i=1}^n x_i^{\alpha_i} \leq \sum_{i=1}^n \alpha_i x_i.$$

>> Solution p. 16

Optimisation et probabilité

Exercice 16. ♦♦ ♣ Optimisation et estimateur

OC16

Soit $\lambda \in \mathbb{R}_+^*$, un paramètre inconnu. Soit $(X_1, \dots, X_n)$ un échantillon d'une loi de Poisson de paramètre λ .

Montrer que parmi toutes les combinaisons linéaires des variables $X_1, \dots, X_n$, il existe un unique estimateur T_n sans biais de λ et qui est de variance minimale. Reconnaître cet estimateur.

>> Solution p. 17

Problème 17. ♦♦ Soient deux entiers $n \geq 2$ et $r \geq 3$. Une expérience aléatoire peut aboutir à r résultats différents $R_1, R_2, \dots, R_r$ de probabilités respectives $x_1, x_2, \dots, x_r \in]0; 1[$. On effectue ensuite n répétitions de l'expérience précédente et, pour tout $i \in \llbracket 1, r \rrbracket$, on note X_i la variable aléatoire qui vaut 1 si le résultat R_i n'est pas obtenu à l'issue de ces n épreuves et qui vaut 0 sinon. On désigne par X la variable aléatoire égale au nombre de résultats qui n'ont pas été obtenus à l'issue des n épreuves.

1. a) Exprimer la variable X en fonction des variables $X_1, X_2, \dots, X_r$.
b) Donner la loi de X_i pour tout $i \in \llbracket 1, r \rrbracket$.
c) En déduire l'espérance de X .

2. En optimisant sous la contrainte $\sum_{i=1}^r x_i = 1$, justifier l'existence et donner la valeur du minimum local de $E(X)$.
3. Dans la suite des questions, on se propose de retrouver le résultat précédent en raisonnant par substitution.
- a) Sachant que $\sum_{i=1}^r x_i = 1$, écrire $E(X)$ comme une fonction, notée $f : (]0; 1[)^{r-1} \rightarrow \mathbb{R}$, des $(r-1)$ variables $x_1, \dots, x_{r-1}$.
- b) Montrer que f est de classe $\mathcal{C}^2$ sur l'ouvert $(]0; 1[)^{r-1}$. Expliciter les dérivées partielles d'ordre 1 de f .
- c) Vérifier que le seul point critique de f est $a = \left(\frac{1}{r}, \frac{1}{r}, \dots, \frac{1}{r}\right)$.
- d) Soit J la matrice de $\mathcal{M}_{r-1}(\mathbb{R})$ dont tous les éléments sont égaux à 1. Vérifier que la matrice hessienne de f en a s'écrit sous la forme
- $$\nabla^2 f(a) = n(n-1) \left(\frac{r-1}{r} \right)^{n-2} (I_{r-1} + J).$$
- e) En remarquant que les valeurs propres de J sont 0 et $r-1$, déduire les valeurs propres de $\nabla^2 f(a)$.
- f) Justifier que f présente un minimum local au point a . Conclure sur la valeur de $E(X)$ correspondant à ce minimum local.

» Solution p. 16

Exercice 18. ♦♦♦ Entropie dans le cas discret

OC18

Soit h_1 la fonction définie sur $[0, 1]$ par :

$$h_1(x) = \begin{cases} -x \ln x & \text{si } x \in]0, 1] \\ 0 & \text{si } x = 0 \end{cases}.$$

1. Vérifier que h_1 est continue sur $[0; 1]$ et donner son graphe.

Soient $n \in \mathbb{N}^*$ et Y , une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ avec $Y(\Omega) = \{y_1, \dots, y_n\}$. On note pour tout $k \in \llbracket 1; n \rrbracket$, $p_k = \mathbf{P}(Y = y_k)$, puis on définit l'entropie de Y par

$$H(Y) = \sum_{k=1}^n h_1(p_k).$$

2. Calculer l'entropie d'une variable aléatoire qui suit une loi uniforme sur $\llbracket 1; n \rrbracket$.
3. Pour quelle valeur du paramètre p , l'entropie d'une loi de Bernoulli $\mathcal{B}(p)$ est maximale?
4. L'objectif de la suite est de déterminer parmi les variables aléatoires prenant n valeurs distinctes, celles qui maximise l'entropie. Pour cela, on définit la fonction h_n sur l'ouvert $]0; 1[^n$ par

$$\forall x = (x_1, \dots, x_n) \in]0; 1[^n, \quad h_n(x) = \sum_{k=1}^n h_1(x_k).$$

- a) Justifier que h_n est de classe $\mathcal{C}^2$ sur $]0, 1[^n$ et expliciter son gradient ∇h_n et sa matrice hessienne $\nabla^2 h_n$.
- b) Vérifier que la fonction h_n admet un unique point critique a sous la contrainte $\sum_{k=1}^n x_k = 1$.
- c) Soit $x \in]0, 1[^n$ tel que $\sum_{k=1}^n x_k = 1$ et notons $h = x - a$. Vérifier que, pour tout $t \in [0, 1]$, on a $a + th \in]0, 1[^n$.
On note alors ψ la fonction définie sur $[0, 1]$ par : $\psi(t) = h_n(a + th)$.
- d)  En utilisant la formule de Taylor avec reste intégral à l'ordre 1 pour ψ entre les points 0 et 1, montrer que h_n admet en a un maximum global sous la contrainte $x_1 + \dots + x_n = 1$. Ce maximum est-il atteint en d'autres points que a ?
5. Parmi les variables aléatoires prenant n valeurs (chacune avec une probabilité non nulle), quelles sont les lois de celles qui ont la plus grande entropie?

» Solution p. 17