CHAPITRE 19

Optimisation sous contraintes linéaires

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

COLUCHE
La revue de presse

Un dernier chapitre sur le calcul différentiel. On étudie ici la recherche d'extremum dans le cas particulier où la fonction est restreinte à un sous-espace affine (droite, plan ...).

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 $i \in [\![1;p]\!]$, il existe $\lambda_{i,1}, \lambda_{i,2}, \dots, \lambda_{i,n}$ tels que

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

Soient $b_1, \dots, b_p$ des réels. Notons $\mathscr{C}$, l'ensemble des solutions du système linéaire

$$\mathscr{S} \quad \left\{ \begin{array}{ccc} g_1(x) & = & b_1 \\ \vdots & \vdots & \vdots \\ g_p(x) & = & b_p, \end{array} \right.$$

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

$$\mathscr{S}_0 \quad \left\{ \begin{array}{ccc} g_1(x) & = & 0 \\ \vdots & \vdots & \vdots \\ g_p(x) & = & 0. \end{array} \right.$$

Comme $\mathscr{H} = \bigcap_{i=1}^{p} \operatorname{Ker}(g_i)$, on constate que $\mathscr{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$. Si $\mathscr{C}$ n'est pas vide et si $x_0 \in \mathscr{C}$ alors on peut écrire $\mathscr{C} = \{x_0 + h \in \mathbb{R}^n : h \in \mathscr{H}\}$. On dit alors que $\mathscr{C}$ est un espace affine.

Preuve. À l'aide de la linéarité de g, on a les équivalences : $x \in \mathscr{C} \iff$

$$\left\{ \begin{array}{ccc} g_1(x) & = & b_1 \\ \vdots & \vdots & \vdots \\ g_p(x) & = & b_p \end{array} \right. \iff \left\{ \begin{array}{ccc} g_1(x) & = & g_1(x_0) \\ \vdots & \vdots & \vdots \\ g_p(x) & = & g_1(x_0) \end{array} \right. \iff \left\{ \begin{array}{ccc} g_1(x-x_0) & = & 0 \\ \vdots & \vdots & \vdots \\ g_p(x-x_0) & = & 0 \end{array} \right. \iff x - x_0 \in \mathscr{H}.$$

■

DÉFINITION **extremum sous contrainte**

Soit $f : \mathcal{O} \to \mathbb{R}$ où $\mathcal{O}$ est une partie de $\mathbb{R}^n$.
- *On dit que f admet un **extremum global sous la contrainte** $\mathscr{C}$ en un point a si la restriction de f à $\mathscr{C}$ admet un extremum global en a.*
- *On dit que f admet un **extremum local sous la contrainte** $\mathscr{C}$ en un point a si la restriction de f à $\mathscr{C}$ admet un extremum local en a.*

Remarque. La fonction f admet un maximum global sous la contrainte $\mathscr{C}$ en a si et seulement si :

$$\forall x \in \mathscr{C}, \qquad f(x) \leqslant f(a).$$

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

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

Pour une production d'une tonne, il faut rajouter la contrainte linéaire $g(x_1, x_2) = x_1 + x_2 = 1$. 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 $\mathscr{C} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 1\}$.

- *Exemple de microéconomie*
Considérons un consommateur dont les achats portent sur plusieurs produits, $P_1, P_2, \ldots, P_n$. La satisfaction éprouvée par ce consommateur dépend des quantités de P_i. Autrement dit, l'utilité qu'il obtient, notée U, est fonction des quantités consommées des produits considérés. Elle peut s'écrire :

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

où x_i représente la quantité de biens P_i. L'analyse économique classique fait l'hypothèse que le consommateur rationnel désire se procurer des quantités x_i des biens X_i telles qu'il puisse obtenir la plus grande satisfaction possible. Il est donc question de la maximisation de la fonction d'utilité sous la contrainte de son revenu. Notons R ce revenu et p_i le prix du bien P_i. Ce revenu et ces prix sont supposés s'établir sur des marchés sur lesquels l'individu considéré ne peut exercer aucune influence. Ainsi R et p_i sont des constantes données. Admettons également que notre consommateur dépense la totalité de son revenu (quitte à rajouter une variable pour l'épargne). Dans ces conditions, le consommateur cherche à maximiser la fonction f sous la contrainte linéaire

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

2 Recherche d'extrema sous contraintes linéaires

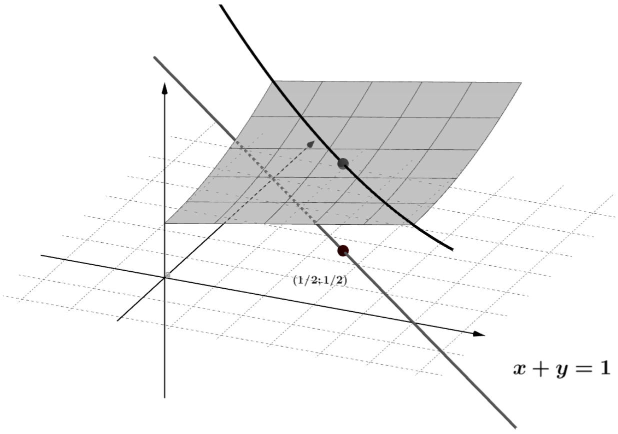
2.1 Par substitution

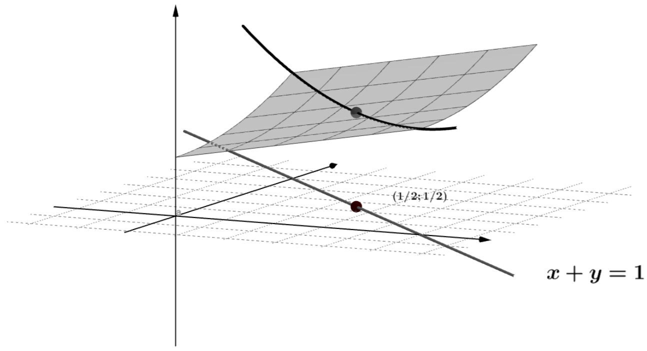
Reprenons l'exemple des épingles. Répartissons la production de la tonne entre les deux usines :

- 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 $x \in [0;1]$, $C'(x) = -4+8x = 8\left(x-1/2\right)$. 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.
Ci-dessous, la surface représentative de C_T et la courbe représentative de la restriction à $\mathscr{C}$.





Exercice 1

✧ 1. Déterminer le minimum de $f : (x,y) \in \mathbb{R}^2 \mapsto x^2+y^2+10$ sous la contrainte

$$3x+y=10.$$

p. 13

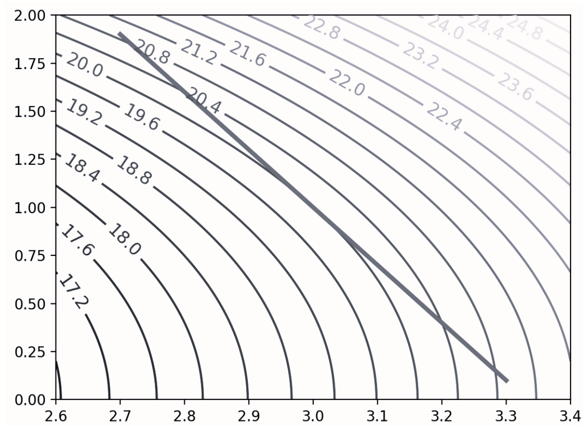
2. Tester la pertinence du résultat à l'aide du code Python ci-dessous.

Editeur

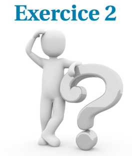
```
x=np.linspace(2.6,3.4,200)
y=np.linspace(0,2,200)
X , Y = np.meshgrid(x,y)
Z = X**2+Y**2+10

graphe = plt.contour(X,Y,Z,25)
plt.clabel(graphe,inline=1)
# lignes de niveau

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



Exercice 2



✧ Soit f la fonction définie sur $\mathbb{R}^3$ par $f(x,y,z) = xe^{x\left(y^2+z^2+1\right)}$.

1. En raisonnant par substitution, déterminer le minimum de de f sous la contrainte linéaire $: \left\{ \begin{array}{l} x=1 \\ y+z=0 \end{array} \right.$.

p. 13

2. Est-ce que f admet un maximum sous cette même contrainte?

2.2 Par du calcul différentiel

Les énoncés

Théorème — **condition nécessaire d'ordre 1**

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

Si
- *f est de classe $\mathscr{C}^1$ sur $\mathscr{O}$.*
- *f admet un extremum sous la contrainte $\mathscr{C}$ en un point a.*

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

$$\forall\, h \in \mathscr{H}, \qquad \left\langle \nabla f(a), h \right\rangle = 0.$$

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

Preuve. Soit $h \in \mathscr{H}$, $h \neq 0_{\mathbb{R}^n}$. Comme $\mathscr{O}$ est une partie ouverte, il existe $r \in \mathbb{R}_*^+$ tel que $\mathscr{B}(a,r) \subset \mathscr{O}$. 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 $t \in]-r; r[$, $a + th \in \mathscr{C}$. La fonction f est de classe $\mathscr{C}^1$, 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 $g_{a,h}$ définie sur un ouvert. Nécessairement

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

D'où le résultat. ∎

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

Remarque. Avec les notations du début de chapitre, pour tout $i \in [\![1; p]\!]$, g_i est de classe $\mathscr{C}^1$ avec :

$$\forall\, x \in \mathbb{R}^n, \qquad \nabla g_i(x) = (\lambda_1, \lambda_2, \ldots, \lambda_n) = \text{ Cste}.$$

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

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

Lemme — **famille génératrice de l'orthogonal**

Avec les notations introduites précédemment,

$$\mathscr{H}^{\perp} = \mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right)$$

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

$$\begin{aligned} x \in \mathscr{H} &\iff \forall\, i \in [\![1;p]\!], \quad g_i(x) = 0 \\ &\iff \forall\, i \in [\![1;p]\!], \quad \left\langle \nabla g_i, x \right\rangle = 0 \\ &\iff \forall\, y \in \mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right\rangle, \quad \langle y, x\rangle = 0 \\ &\iff x \in \mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right)^{\perp}. \end{aligned}$$

On a donc l'égalité
$$\mathscr{H} = \mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_n\right)^{\perp}.$$

Considérons les supplémentaires orthogonaux

$$\mathscr{H}^{\perp} = \left(\mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right)^{\perp}\right)^{\perp} = \mathrm{Vect}\left(\nabla g_1, \ldots, \nabla g_p\right).$$

La dernière égalité résulte du fait qu'en dimension finie $\left(\mathrm{F}^{\perp}\right)^{\perp} = \mathrm{F}$ pour tout sous-espace vectoriel F. ∎

THÉORÈME **condition d'ordre 1**

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

Si | — *f est de classe $\mathscr{C}^1$ sur $\mathscr{O}$.*
— *f admet un extremum sous la contrainte $\mathscr{C}$ en un point a.*

Alors *il existe $(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p$ 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
$$\begin{aligned}
\partial_1 \nabla f(x, y, z) &= \left(1 + \left(y^2 + z^2 + 1\right) x\right) e^{x\left(y^2+z^2+1\right)} \\
\partial_2 f(x, y, z) &= 2x^2 y e^{x\left(y^2+z^2+1\right)} \\
\partial_3 f(x, y, z) &= 2x^2 z e^{x\left(y^2+z^2+1\right)}.
\end{aligned}$$

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

Vocabulaire. Si la famille $\left(\nabla g_1, \ldots, \nabla g_p\right)$ est libre, les réels $\lambda_1, \ldots, \lambda_p$ 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
$$\mathscr{C} : \left\{ \begin{array}{rcl} x + y + z &=& 0 \\ x - y + z &=& 2. \end{array} \right.$$

p. 14

⚠ **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
$$f(x) - f(a) \quad \text{pour } x \in \mathscr{C}$$
ou, à l'aide du changement de variable $h = x - a$, on regarde le signe de
$$f(a + h) - f(a) \quad \text{pour } h \in \mathscr{H}.$$
- 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 $\varepsilon(h) \xrightarrow[h \to 0]{} 0$ et $h \in \mathscr{H}$. Dans le cas d'un point critique sous la contrainte $\mathscr{C}$, on obtient la simplification :
$$\forall\, h \in \mathscr{H}, \qquad f(a + h) - f(a) = \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h) \quad (\text{car } \nabla f(a) \in \mathscr{H}).$$
On regarde ensuite le signe de la forme quadratique (restreinte à $\mathscr{H}$).

— Si $q_a(h) \geqslant 0$ pour tout $h \in \mathcal{H}$, on a un minimum local.

— Si $q_a(h) \leqslant 0$ 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.

Deux exemples d'études globales

- 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

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

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

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

En particulier $x = y = z$. Comme $a \in \mathcal{C}$, $g(a) = 3$, puis $x = y = z = 1$. En particulier, on a $f(a) = 3$.

— *Étude locale au niveau du point critique*
On vérifie que

$$\nabla f(a) = (2, 2, 2) \quad \text{et} \quad \nabla^2 f(a) = 2\mathrm{I}_3.$$

En particulier, la forme quadratique devient

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

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 $f(a + h) - f(a) \geqslant 0$. Le point a donne un minimum local sous la contrainte.

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

$$\mathrm{X} = a + \mathrm{H} = (1 + h_1; 1 + h_2; 1 + h_3).$$

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

$$\begin{aligned}
f(\mathrm{X}) - f(a) &= (1 + h_1)^2 + (1 + h_2)^2 + (1 + h_3)^2 - 3 \\
&= {h_1}^2 + 2h_1 + {h_2}^2 + 2h_2 + {h_3}^2 + 2h_3 \\
&= {h_1}^2 + 2h_1 + {h_2}^2 + 2h_2 + (h_1 + h_2)^2 - 2(h_1 + h_2) \quad \text{car } h_3 = -h_1 - h_2 \\
f(\mathrm{X}) - f(a) &= {h_1}^2 + {h_2}^2 + (h_1 + h_2)^2 \geqslant 0.
\end{aligned}$$

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

- On considère

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

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 $\mathscr{C}^1$ sur l'ouvert $\left(\mathbb{R}_*^+\right)^n$ avec pour tout $x \in \left(\mathbb{R}_*^+\right)^n$

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

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

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

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

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

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

Soient $i, j \in [\![1; n]\!]$ $\qquad a_i \partial_i f(a) = \prod_{k=1}^{n} a_k = a_j \partial_j f(a)$

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

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

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

— *Nature du point critique*
Notons que la partie de $\mathbb{R}^n$ définie par

$$\mathscr{K}_c = \left(\mathbb{R}^+\right)^n \cap \left\{x \in \mathbb{R}^n \mid g(x) = c\right\}$$

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

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

alors l'une des coordonnées 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

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

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 (le point a). On a donc montré que $f(a)$ est le maximum de f sous la contrainte :

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

Formule que l'on réécrit sous la forme

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

C'est l'inégalité arithmético-géométrique.

Exercice 4

✦ Soit f définie sur $\mathbb{R}^3$ par $f(x, y, z) = x^2 - 2xy + yz + y - z$.

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

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

p. 14

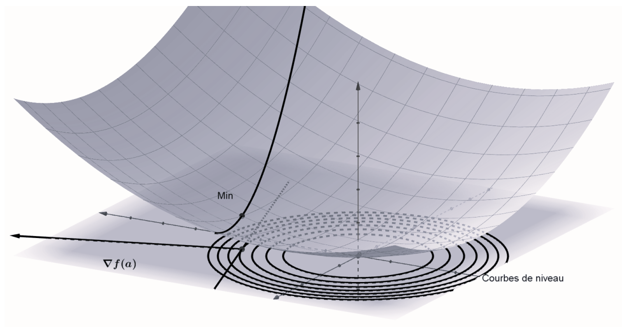
2. On a $f(a) = 1/5$. Est-ce un maximum ou minimum global ? ni l'un ni l'autre ?

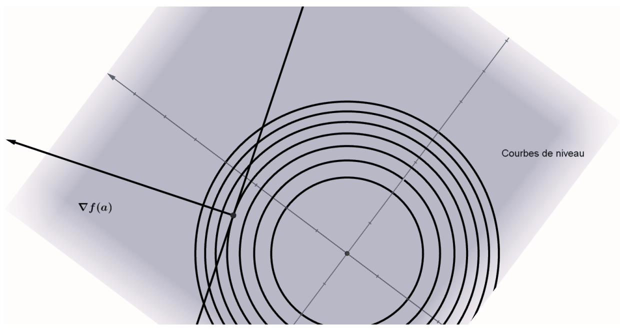
2.3 Interprétation géométrique avec une seule contrainte linéaire

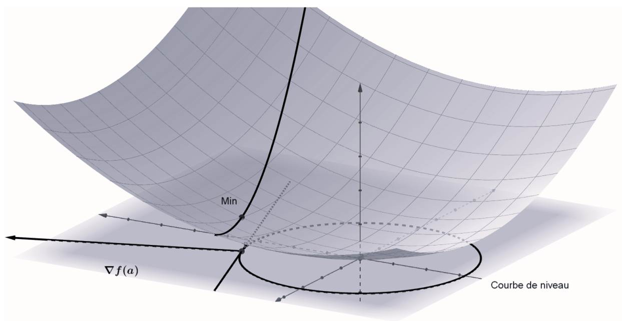
Dans le cas d'une fonction de deux variables avec une seule contrainte linéaire, on cherche à minimiser la restriction de f à une droite d'équation $g(x,y) = \lambda x + \mu y = b$. Le vecteur ∇g est un vecteur normal [1] à la droite $\mathscr{C}$. Or, on sait que :

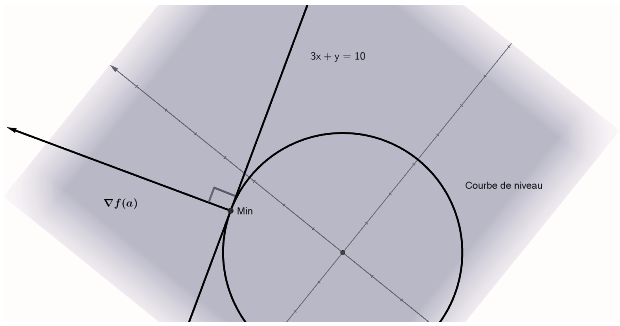
- Si un extremum sous contrainte est atteint en a, $\nabla f(a)$ est colinéaire à ∇g. En particulier, $\nabla f(a)$ est un vecteur normal à la droite $\mathscr{C}$.
- Le vecteur $\nabla f(a)$ est aussi orthogonal à la ligne de niveau $\mathscr{L}_{f(a)}$ (les détails page p.**??**).

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 $3x + y = 10$.









3 Cas des fonctions définies sur des fermés bornés

Commençons par résumer la démarche avant de l'appliquer à un exemple.

Méthode

Existence et calcul des extrema d'une fonction définie sur un fermé borné
Dans la suite, on considère une fonction f de classe $\mathscr{C}^1$ sur un ensemble fermé borné, noté $\mathscr{K}$.

1. Dit autrement, le vecteur est orthogonal aux vecteurs directeurs de la droite.

Méthode

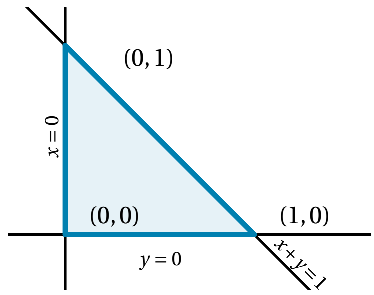
- *I. Existence*
  On vérifie que l'ensemble de départ de f est bien une partie :
  - fermé, par exemple, en montrant qu'il existe φ, une fonction continue de $\mathbb{R}^n$, telle que
  $$\mathcal{K} = \left\{x \in \mathbb{R}^n \,|\, \varphi(x) \leqslant c\right\}.$$
  - bornée, en montrant qu'il existe un réel M tel que pour tout $x \in \mathcal{K}$, $\|x\| \leqslant \mathrm{M}$.

  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 différents 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 $f(x, y) = 3x^2 + 2y^2 - 4x - 3y + 4xy + 3$.
On souhaite étudier le minimum global de la restriction de f à l'ensemble

$$\mathcal{K} = \left\{(x, y) \in \mathbb{R}^2 \,|\, x \geqslant 0,\ y \geqslant 0, x + y \leqslant 1\right\}.$$

- *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 :
  $$\mathbb{R}^+ \times \mathbb{R}, \quad \mathbb{R} \times \mathbb{R}^+ \quad \text{et} \quad \left\{(x, y) \in \mathbb{R}^2 \,|\, x + y \leqslant 1\right\}.$$
  La partie $\mathcal{K}$ est aussi bornée. Par exemple
  $$\mathcal{K} \subset \overline{\mathcal{B}}(0; 1).$$
  C'est-à-dire $\quad \forall (x, y) \in \mathcal{K}, \quad \|(x, y)\| \leqslant 1.$
  La fonction f étant continue car polynomiale, il existe donc un minimum sur $\mathcal{K}$.



- *II. Étude sur l'ouvert*
  Soit $\mathcal{O} = \{(x, y) \in \mathbb{R}^2 \,|\, x > 0,\ y > 0, x + y < 1\}$. La fonction f est de classe $\mathcal{C}^1$ sur $\mathcal{O}$ car polynomiale et
  $$\partial_1 f(x, y) = 6x - 4 + 4y \quad \text{et} \quad \partial_2 f(x, y) \quad = 4y - 3 + 4x.$$
  Vérifier que l'on a un unique point critique donné par $a = \left(\frac{1}{2}; \frac{1}{4}\right) \in \mathcal{O}$ avec $f(a) = 13/8 = 1,625$.

- *III. Étude sur le bord*

  - Cas $y = 0$: $f(x, 0) = 3x^2 - 4x + 3$. Une rapide étude de cette fonction polynomiale de degré 2 donne un minimum valant $5/3 \simeq 1,67$.
  - Cas $x = 0$: $f(0, y) = 2y^2 - 3y + 3$. On trouve 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

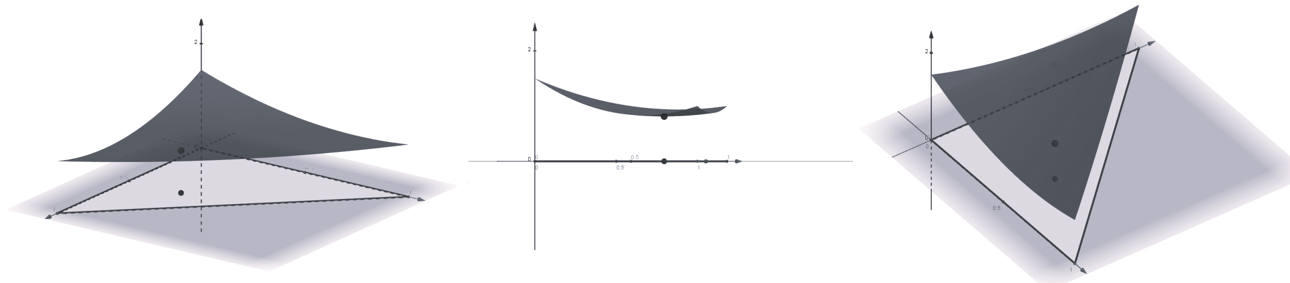
$$\exists \lambda \in \mathbb{R}, \quad \begin{cases} 6x+4y-4 &= \lambda \\ 4x+4y-3 &= \lambda \end{cases} \underset{L_1 \leftarrow L_1 - L_2}{\iff} \exists \lambda \in \mathbb{R}, \quad \begin{cases} 2x-1 &= 0 \\ 4x+4y-3 &= \lambda \end{cases}$$

On obtient $x = 1/2$. La condition $x + y = 1$ donne en plus, $y = 1/2$.
L'étude de la fonction $x \mapsto f(x, 1-x)$ donne 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 chaque sous-ensemble de $\mathscr{K}$, le minimum global est atteint pour :

$$a = \left(\frac{1}{2}; \frac{1}{4}\right) \quad \text{et} \quad f(a) = 1,625.$$

Illustrons le résultat par plusieurs vues de la surface de f :



Exercice 5

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

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

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.

p. 14

Exercices

Exercice 6. ✦ Soit f définie sur $\mathcal{O} = \left(\mathbb{R}_+^*\right)^3$ par $f(x,y,z) = x\ln x + y\ln y + z\ln z$. Pour tout $c \in \mathbb{R}$, on pose

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

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

≫ Solution p. 15

Exercice 7. ✦✦ Soient $n \in \mathbb{N} \setminus \{0;1\}$ et f la fonction définie sur $\left(\mathbb{R}_*^+\right)^n$ par

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

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

≫ Solution p. 15

Exercice 8. ✦✦ **Variante de l'inégalité arithmético-géométrique**
Soient $n \in \mathbb{N}^*$, $\alpha_1,\ldots,\alpha_n \in \mathbb{R}_*^+$ tels que $\sum_{i=1}^{n} \alpha_i = 1$. On définit en plus les fonctions f et g sur $\left(\mathbb{R}^+\right)^n$ par

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

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

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

$$\forall\, (x_1,\ldots,x_n) \in \left(\mathbb{R}^+\right)^n, \qquad \prod_{i=1}^{n} {x_i}^{\alpha_i} \leqslant \sum_{i=1}^{n} \alpha_i x_i.$$

≫ Solution p. 16

Optimisation et probabilité

Exercice 9. ✦✦ **Optimisation et estimateur**
Soit $\lambda \in \mathbb{R}_*^+$, un paramètre inconnu. Soit $(X_1,\ldots,X_n)$ un échantillon d'une loi de Poisson de paramètre λ.
Montrer que parmi toutes les combinaisons linéaires des variables $X_1,\ldots,X_n$, il existe un unique estimateur T_n sans biais de λ et qui est de variance minimale. Reconnaître cet estimateur.

≫ Solution p. 16

Problème 10. ✦✦ Soient deux entiers $n \geqslant 2$ et $r \geqslant 3$. Une expérience aléatoire peut aboutir à r résultats différents $R_1, R_2, \ldots, R_r$ de probabilités respectives $x_1, x_2, \ldots, x_r \in\,]0;1[$. On effectue ensuite n répétitions de l'expérience précédente et, pour tout $i \in [\![1,r]\!]$, 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 $X_1, X_2, \ldots, X_r$.
   b) Donner la loi de X_i pour tout $i \in [\![1;r]\!]$.
   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 $\mathbf{E}(X)$.
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 $\mathbf{E}(X)$ comme une fonction, notée $f : \left(]0;1[\right)^{r-1} \to \mathbb{R}$, des $(r-1)$ variables $x_1, \ldots, x_{r-1}$.
   b) Montrer que f est de classe $\mathscr{C}^2$ sur l'ouvert $\left(]0;1[\right)^{r-1}$. Expliciter les dérivées partielles d'ordre 1 de f.
   c) Vérifier que le seul point critique de f est $a = \left(\frac{1}{r}, \frac{1}{r}, \ldots, \frac{1}{r}\right)$.

d) Soit J la matrice de $\mathscr{M}_{r-1}(\mathbb{R})$ dont tous les éléments sont égaux à 1. Vérifier que la matrice hessienne de f en a s'écrit sous la forme

$$\nabla^2 f(a) = n(n-1)\left(\frac{r-1}{r}\right)^{n-2} (\mathrm{I}_{r-1} + \mathrm{J}).$$

e) En remarquant que les valeurs propres de J sont 0 et $r-1$, déduire les valeurs propre de $\nabla^2 f(a)$.

f) Justifier que f présente un minimum local au point a. Conclure sur la valeur de $\mathbf{E}(\mathrm{X})$ correspondant à ce minimum local.

≫ Solution p. 16

Exercice 11. ✦✦✦ **Entropie dans le cas discret**

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 $(\Omega, \mathscr{A}, \mathbf{P})$ avec $\mathrm{Y}(\Omega) = \{y_1, \ldots, y_n\}$. On note pour tout $k \in [\![1;n]\!]$, $p_k = \mathbf{P}(\mathrm{Y} = y_k)$, puis on définit l'entropie de Y par

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

2. Calculer l'entropie d'une variable aléatoire qui suit une loi uniforme sur $[\![1;n]\!]$.

3. Pour quelle valeur du paramètre p, l'entropie d'une loi de Bernoulli $\mathscr{B}(p)$ 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

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

a) Justifier que h_n est de classe $\mathscr{C}^2$ 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 $x_1 + \cdots + x_n = 1$. 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. 18

Indications et solutions

Indication de l'exercice 4
p. 7

2. Pour $X = a + h = a + (h_1, h_2, h_3) \in \mathscr{C}$, vérifier que

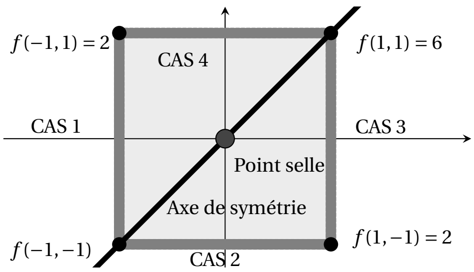
$$f(X) = \frac{1}{5} - 5{h_1}^2.$$

Conclure un maximum sous contrainte.

Indication de l'exercice 5
p. 10

2. Vérifier que $(0,0)$ est l'unique point critique et regarder le signe de $f(h,h)$ et $f(h,-h)$ pour étudier la nature de ce point critique.

3. Il faut étudier le maximum sur le bord de D. On peut distinguer 4 cas :



Indication de l'exercice 6
p. 11

2. Former le développement limité d'ordre 2 et vérifier que la matrice hessienne est $\nabla^2 f(a) = (1/c)I_3$. Relire page p.6 si besoin.

Indication de l'exercice 7
p. 11

Vérifier que $a = (1, 1, \dots, 1)$ est l'unique point critique. Utiliser deux fois l'inégalité de Cauchy-Schwarz pour justifier que le minimum est global.

Indication de l'exercice 8
p. 11

3. Poser $s = \sum_{i=1}^{n} \alpha_i x_i \neq 0$ et pour tout indice i, $\tilde{x}_i = x_i / s$ afin de se ramener au cas traité par les questions 1 et 2.

Indication de l'exercice 9
p. 11

Poser $T_n = \sum_{i=1}^{n} \alpha_i X_i$ et vérifier que le problème revient à trouver le minimum de la fonction définie par $f(\alpha_1, \dots, \alpha_n) = \sum_{i=1}^{n} {\alpha_i}^2$ sous la contrainte $\sum_{i=1}^{n} \alpha_i = 1$.

Indication de l'exercice 11
p. 12

4.d) Appliquer le théorème sur les dérivées directionnelles (p.**??** et p.**??**).

Exercice 1
p. 3

1. On a $y = 10 - 3x$ et il suffit alors de minimiser la fonction polynomiale

$$x \in \mathbb{R} \mapsto f(x, 10-3x) = x^2 + (10-3x)^2 + 10.$$

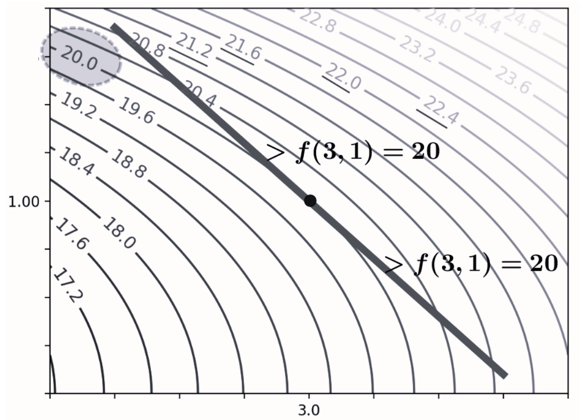
C'est un polynôme de degré 2 et de coefficient dominant positif. La fonction est convexe et le minimum existe, il est atteint en un point critique

$$0 = f'(x) = 2x + 2 \cdot 3(3x - 10) = 20(x-3).$$

D'où $x = 3$, puis $y = 1$ et le minimum sous la contrainte est uniquement atteint en $(3,1)$. Il vaut

$$f(3,1) = 20.$$

2.



De part et d'autre de la droite on constate que l'intersection avec la ligne de niveau ne se fait qu'avec des lignes de niveau dont la hauteur est supérieure à $20 = f(3,1)$. Dès lors, les point situés sur la droite ont une image supérieure à $f(3,1)$, ce qui confirme l'existence d'un minimum local atteint en $(3,1)$.

Exercice 2
p. 3

1. Pour tout $(x, y, z) \in \mathscr{C}$, $x = 1$, $z = -y$,

$$f(x, y, z) = 1 \cdot e^{1\left(y^2 + (-y)^2 + 1\right)} = e^{2y^2+1}.$$

Comme la fonction exponentielle est strictement croissante, les variations de $y \mapsto e^{2y^2+1}$ sont identiques à celle de $y \mapsto 2y^2 + 1$. On trouve un minimum global atteint en $y = 0$.

On en déduit que $z=0$. Couplé avec le fait que $x=1$, on a un minimum sous la contrainte atteint en $(1,0,0)$. Il vaut

$$f(1,0,0)=\mathrm{e}.$$

2. Comme

$$f(1,y,-y)\underset{y\to+\infty}{\longrightarrow}+\infty,$$

il n'y a pas de maximum sous la contrainte.

Exercice 3 p. 5

Avec les notations précédentes

$$\forall(x,y,z)\in\mathbb{R}^3,\quad \begin{aligned} g_1(x,y,z)&=x+y+z\\ g_2(x,y,z)&=x-y+z\end{aligned}$$

et $\qquad \nabla g_1(x,y,z)=(1,1,1),\quad \nabla g_2(x,y,z)=(1,-1,1).$

De plus $\qquad \nabla f(x,y,z)=(yz,xz,xy).$

Si (x,y,z) est un extremum sous la contrainte avec $\mathbb{R}^n$, un ouvert, alors, c'est aussi un point critique sous la contrainte et il existe $\lambda_1,\lambda_2\in\mathbb{R}$ tels que

$$\nabla f(x,y,z)=\lambda_1\nabla g_1+\lambda_2\nabla g_2$$

En rajoutant la contrainte, on obtient le système

$$\begin{cases} x+y+z=0\\ x-y+z=2\\ yz=\lambda_1+\lambda_2\\ xz=\lambda_1-\lambda_2\\ xy=\lambda_1+\lambda_2\end{cases}$$

Les lignes 3 et 5 donnent $yz=xy$. Notons que y ne peut être nul en raisonnant par l'absurde sur les lignes l et 2. D'où $z=x$ et les deux premières lignes deviennent

$$\begin{cases} 2x+y=0\\ 2x-y=2\end{cases}\iff\begin{cases} x=1/2\\ y=-1.\end{cases}$$

On obtient un unique point critique sous la contrainte

$$a=\left(\frac{1}{2};-1;\frac{1}{2}\right).$$

et un extremum possible est $f(a)=-1/4$.

Exercice 4 p. 7

1. Pour $(x,y,z)\in\mathbb{R}^3$, posons

$$g_1(x,y,z)=2x-y\quad\text{et}\quad g_2(x,y,z)=x+z.$$

Le point $a=(x,y,z)$ est un point critique sous contrainte si et seulement si il existe deux réels λ_1,λ_2 tels que

$$\nabla f(a)=\lambda_1\nabla g_1+\lambda_2\nabla g_2.$$

Or

$$\begin{aligned}\nabla f(a)&=(2x-2y,-2x+z+1,y-1)\\ \nabla g_1&=(2,-1,0)\\ \nabla g_2&=(1,0,1).\end{aligned}$$

D'où

$$\begin{cases} 2x-2y=2\lambda_1+\lambda_2\\ -2x+z+1=-\lambda_1\\ y-1=\lambda_2\end{cases}$$
$$\iff 2x-2y=-2(-2x+z+1)+(y-1)$$
$$\iff -2x-3y+2z=-3.$$

En rajoutant les contraintes, on obtient le système à trois inconnues pour trois équations

$$\begin{cases} 2x+3y-2z=3\\ 2x-y=1\\ x+z=1\end{cases}$$

La résolution de système linéaire (par un pivot de Gauss par exemple) donne bien $a=\left(\frac{4}{5},\frac{3}{5},\frac{1}{5}\right)$.

2. Soit $X=a+h\in\mathscr{C}$. Dans ce cas $h=(h_1,h_2,h_3)\in\mathscr{H}$. La définition de $\mathscr{H}$ donne

$$h_2=2h_1\quad\text{et}\quad h_3=-h_1.$$

Ainsi $X=\left(\frac{4}{5}+h_1,\frac{3}{5}+2h_1,\frac{1}{5}-h_1\right)$. On obtient en développant

$$f(X)=\frac{1}{5}-5h_1{}^2$$

On a donc pour tout $X\in\mathscr{C}$,

$$f(X)-f(a)=-5h_1{}^2\leqslant 0.$$

On a donc au point a un maximum global pour f sous la contrainte $\mathscr{C}$.

Exercice 5 p. 10

1. La fonction f est continue en tant que restriction d'une fonction polynomiale. De plus, D est une partie fermée et bornée de $\mathbb{R}^2$. On sait alors que f admet un minimum et maximum global sur D.

2. La fonction f est de classe $\mathscr{C}^1$ en tant que restriction d'une fonction polynomiale. De plus, pour tout $(x,y)\in]-1;1[^2$

$$\partial_1 f(x,y)=2(x-y)+6y=2x+4y.$$

Pas symétrie $\qquad \partial_2 f(x,y)=2y+4x.$

Des lors, (x,y) est point critique si et seulement si

$$\begin{cases} 2x+4y=0\\ 2y+4x=0\end{cases}\iff x=y=0.$$

Or, Pour $h\in]-1;1[$,

$$f(h,h)-f(0,0)=6h^2\geqslant 0,$$
$$f(h,-h)-f(0,0)=-4h^2\leqslant 0.$$

Il n'y a donc un changement de signe et ce point critique est un point selle ($f(0,0)$ n'est ni un maximum, ni un minimum).

3. D'après ce qui précède, on sait que le maximum et le minimum de f sur D sont atteints aux bords de D.

→ *Cas 1.*

Si $x=-1$, pour $y\in[-1;1]$

$$f(-1,y)=(y+1)^2-6y=y^2-4y+1=(y-2)^2-1.$$

Sur $[-1;1]$, la fonction polynomiale est décroissante donc

$$\forall y \in [-1;1],\quad f(-1,y) \leqslant f(-1,-1) = 6$$
$$-2 = f(-1,1) \leqslant f(-1,y)$$

→ *Cas 2.*
Si $y=-1$ et $x \in [-1;1]$, on a par symétrie de f

$$f(x,-1) = f(-1,x) \leqslant f(-1,-1) \in [-2;6].$$

→ *Cas 3.*
Si $x = 1$ pour $y \in [-1;1]$ et

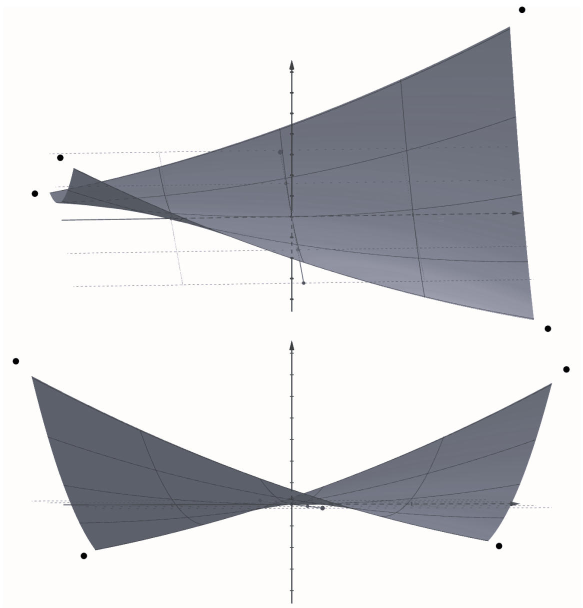
$$f(1,y) = (y-1)^2 + 6y = y^2 - 4y + 1 = \left(y^2-2\right)^2 - 3$$

$$-2 = f(1,-1) \leqslant f(1,y) \leqslant f(1,1) = 6$$

→ *Cas 4.*
Si $y = 1$ pour $x \in [-1;1]$, on a par symétrie

$$f(x,1) = f(1,x) \in [-2;6].$$

Finalement, f admet un maximum 6 atteint en $(1,1)$ et $(-1,-1)$, un minimum atteint en $(1,-1)$ et $(-1,1)$.



Exercice 6

p. 11

1. Les applications affines

$$(x,y,z) \in \mathscr{O} \mapsto x,\quad (x,y,z) \in \mathscr{O} \mapsto y$$
$$\text{et } (x,y,z) \in \mathscr{O} \mapsto z$$

sont de classe $\mathscr{C}^1$ sur $\mathscr{O}$ et à valeurs dans $\mathbb{R}_*^+$.
La fonction logarithme est $\mathscr{C}^1$ sur $\mathbb{R}_*^+$
Par composition et produits, les fonctions

$$(x,y,z) \in \mathscr{O} \mapsto x\ln(x),\quad (x,y,z) \in \mathscr{O} \mapsto y\ln(y)$$
$$\text{et } (x,y,z) \in \mathscr{O} \mapsto z\ln(z)$$

sont de classe $\mathscr{C}^1$ sur $\mathscr{O}$.
Par somme, f est de classe $\mathscr{C}^1$ sur $\mathscr{O}$.

- Posons $g : (x,y,z) \in \mathbb{R}^3 \mapsto x+y+z$. La fonction g est $\mathscr{C}^1$ avec

$$\nabla g(x,y,z) = (1,1,1).$$

De plus

$$\nabla f(x,y,z) = \big(\ln(x)+1, \ln(y)+1, \ln(z)+1\big).$$

Dès lors, $a \in \mathscr{O}$ est un point critique sous la contrainte s'il existe $\lambda \in \mathbb{R}$ tel que

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

Cela implique en particulier que $x = y = z$. La condition $a \in \mathscr{C}_c$ donne $x = y = z = c$.
On a bien un unique point critique donné par $a = (c,c,c)$.

2. Formons le développement limité à l'ordre 2. On montre comme dans la question 1 que f est de classe $\mathscr{C}^2$ sur $\mathscr{O}$ et

$$\nabla^2 f(a) = \frac{1}{c}\mathrm{I}_3.$$

Ensuite, pour tout $h \in \mathbb{R}^3$ tel que $a+h \in \mathscr{O}$,

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

où H est la matrice colonne des coordonnées de h dans la base canonique et

$$\varepsilon(h) \underset{h\to 0}{\longrightarrow} 0.$$

Dans le cas où $a+h \in \mathscr{C}_c$, on a $\nabla f(a) \perp h$ et

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2c}{}^t\mathrm{H}\mathrm{H} + \|h\|^2\varepsilon(h) \\ &= \left(\frac{1}{2c} + \varepsilon(h)\right)\|h\|^2. \end{aligned}$$

Dans un voisinage de $0_{\mathbb{R}^3}$, $\frac{1}{2c} + \varepsilon(h) \geqslant 0$ et

$$f(a+h) - f(a) \geqslant 0$$

Autrement dit, f admet un minimum local en a sous la contrainte $\mathscr{C}_c$.

Exercice 7

p. 11

→ *Point critique*
La fonction f est polynomiale, donc $\mathscr{C}^1$ avec pour tout $i \in [\![1;n]\!]$

$$\forall x = (x_1,\dots,x_n) \in \left(\mathbb{R}_*^+\right)^n,\quad \partial_i f(x) = 4{x_i}^3.$$

Posons $g(x) = \sum_{i=1}^{n} x_i$. La contrainte est linéaire avec

$$\nabla g(x) = (1,1,\dots,1).$$

Dans le cas d'un point critique sous la contrainte, il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla f(x) = \lambda \nabla g(x).$$

On en déduit que

$${x_1}^3 = {x_2}^3 = \ldots = {x_n}^3.$$

Par injectivité de la fonction cubique sur $\mathbb{R}$

$$x_1 = x_2 = \cdots = x_n.$$

La contrainte $x_1 + \ldots + x_n = n$ donne un unique point critique
$$a = (1, 1, \ldots, 1) \quad \text{avec} \quad f(a) = n.$$

→ *Étude globale*

À partir de l'inégalité de Cauchy-Schwarz sur $\mathbb{R}^n$ muni du produit scalaire canonique, on a
$$\forall\,(y_1, \cdots, y_n) \in \mathbb{R}^n, \qquad \left(\sum_{i=1}^n y_i\right)^2 \leqslant n \sum_{i=1}^n {y_i}^2.$$

Pour x vérifiant la contrainte et en appliquant deux fois l'inégalité précédente
$$n^4 = \left(\sum_{i=1}^n x_i\right)^4 \leqslant \left(n \sum_{i=1}^n {x_i}^2\right)^2 \leqslant n^2 \cdot n \sum_{i=1}^n {x_i}^4.$$

D'où $\qquad f(a) = n \leqslant f(x).$

On a un minimum global (et donc local).
Notons que l'unicité du point critique sous la contrainte permet d'affirmer qu'il n'y a pas de maximum local (ni global) sous la contrainte.

Exercice 8 p. 11

1. 2. Reprendre le second exemple page 6. On trouve M = 1 atteint uniquement en $(1, \ldots, 1)$.

3. Soient $(x_1, \ldots, x_n) \in \mathbb{R}^+$. Notons que si l'une des composantes est nulle alors l'inégalité est directement vérifiée. On suppose donc que toutes les composantes sont non nulles et on pose
$$s = \sum_{i=1}^n \alpha_i x_i \neq 0$$
et pour tout indice i, $\tilde{x}_i = x_i / s$ de sorte que
$$\sum_{i=1}^n \tilde{x}_i = \frac{1}{s} \sum_{i=1}^n \alpha_i x_i = 1.$$
Le point $(\tilde{x}_1, \ldots, \tilde{x}_n)$ vérifie donc la contrainte de la question 1. Il vient par la question 2
$$\prod_{i=1}^n (\tilde{x}_i)^{\alpha_i} \leqslant 1.$$
D'où $\qquad \displaystyle\prod_{i=1}^n {x_i}^{\alpha_i} \leqslant \prod_{i=1}^n s^{\alpha_i} = s^{\sum_{i=1}^n \alpha_i} = s.$

Ce qui conclut.

Exercice 9 p. 11

Soient $\alpha_1, \ldots, \alpha_n$ des réels tels que
$$\mathrm{T}_n = \sum_{i=1}^n \alpha_i \mathrm{X}_i.$$
La condition sans biais impose $\mathbf{E}(\mathrm{T}_n) = \lambda$ et par linéarité de l'espérance
$$\mathbf{E}(\mathrm{T}_n) = \sum_{i=1}^n \alpha_i \mathbf{E}(\mathrm{X}_i) = \lambda \sum_{i=1}^n \alpha_i.$$
Comme $\lambda \neq 0$, on a
$$\sum_{i=1}^n \alpha_i = 1.$$
De plus par indépendance et l'égalité en loi
$$\mathbf{V}(\mathrm{T}_n) = \sum_{i=1}^n {\alpha_i}^2 \mathbf{V}(\mathrm{X}_i) = \lambda \sum_{i=1}^n {\alpha_i}^2.$$
On cherche donc à minimiser
$$\sum_{i=1}^n {\alpha_i}^2 \quad \text{avec} \quad (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$$
sous la contrainte $\sum_{i=1}^n \alpha_i = 1$. En reprenant la démarche de l'exercice 7. On montre via l'inégalité de Cauchy-Schwarz qu'on a un minimum global au pont $(\alpha_i)_{i \in [\![1;n]\!]}$ avec
$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = \frac{1}{n}.$$
Dans ce cas $\qquad \displaystyle\mathrm{T}_n = \frac{1}{n} \sum_{i=1}^n \mathrm{X}_i.$

C'est l'estimateur de la moyenne empirique.

Exercice 10 p. 11

1.a) Vérifier que
$$\mathrm{X} = \sum_{i=1}^r \mathrm{X}_i.$$

1.b) Soit $i \in [\![1, r]\!]$. La variable X_i est à valeurs dans $\{0; 1\}$, elle suit donc une loi de Bernoulli. Le paramètre est donné par
$$\mathbf{P}(\mathrm{X}_i = 1) = \mathbf{P}\left(\overline{\mathrm{A}_1} \cap \overline{\mathrm{A}_2} \cap \ldots \cap \overline{\mathrm{A}_n}\right) = \prod_{i=1}^n \mathbf{P}\left(\overline{\mathrm{A}_i}\right)$$
où A_i désigne l'événement "le résultat R_i est obtenu à la i-ème épreuve". Ces événements sont indépendants et de même probabilité x_i, d'où
$$\mathbf{P}(\mathrm{X}_i = 1) = \prod_{i=1}^n (1 - x_i) = (1 - x_i)^n.$$
Ce qui conclut.

1.c) Par linéarité de l'espérance
$$\mathbf{E}(\mathrm{X}) = \sum_{i=1}^r \mathbf{E}(\mathrm{X}_i) = \sum_{i=1}^r (1 - x_i)^n.$$

2. Définissons f sur $[0; 1]^r$ par
$$f(x_1, \ldots, x_r) = \sum_{i=1}^r (1 - x_i)^n.$$
La fonction f est polynomiale, donc de classe $\mathscr{C}^1$ et
$$\nabla f(x_1, \ldots, x_r) = \left(-n(1 - x_1)^{n-1}, \ldots, -n(1 - x_r)^{n-1}\right)$$
La contrainte est $g(x_1, \ldots, x_r) = \sum_{i=1}^r x_i = 1$ avec
$$\nabla g = (1, 1, \ldots, 1).$$

Si $(x_1,\dots,x_r)$ est un point critique sous la contrainte, alors il existe $\lambda\in\mathbb{R}$ tel que

$$\nabla f(x_1,\dots,x_r)=\lambda\nabla g.$$

On en déduit que pour tout indice i

$$-n\left(1-x_i\right)^{n-1}=1.$$

Par injectivité de la fonction $t\mapsto -n(1-t)^{n-1}$ sur $[0;1]$, on montre que

$$x_1=x_2=\ldots=x_r.$$

Puis la condition $g(x_1,x_2,\dots,x_r)=1$ donne

$$x_1=x_2=\ldots=x_r=\frac{1}{r}.$$

On a donc un unique point critique sous la contrainte donné par

$$a=\frac{1}{r}(1,1,\dots,1)\in\mathbb{R}^r.$$

- Prouvons qu'on a un minimum local.

La fonction f est de classe $\mathscr{C}^2$ avec pour tous $k,\ell\in[\![1;r]\!]$

$$\partial^2_{k,\ell}f(x)=0 \quad \text{si } k\neq\ell$$
$$\partial^2_{k,k}f(x)=n(n-1)\left(1-x_k\right)^{n-2}.$$

En particulier, pour tout $k\in[\![1;r]\!]$, $\partial^2_{k,k}f(a)$ ne dépend pas de k. Notons α cette valeur commune. Le réel α est positif strictement et

$$\nabla^2 f(a)=\alpha\,\mathrm{I}_r.$$

La matrice hessienne a un spectre inclus dans $\mathbb{R}_*^+$. En revenant au développement limité à l'ordre 2, pour tout $h\in\mathscr{H}$

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

avec $$q_a(h)=\alpha\|h\|^2\geqslant 0.$$

On en déduit que $f(a)$ est un minimum local sous la contrainte.

3.a) Pour tout $i\in[\![1,r]\!]$, notons S_i l'événement la première épreuve donne le résultat R_i. $(\mathrm{S}_1,\mathrm{S}_2,\dots,\mathrm{S}_r)$ est un système complet d'événements donc

$$x_1+x_2+\cdots+x_r=\mathbf{P}(\mathrm{S}_1)+\mathbf{P}(\mathrm{S}_2)+\cdots+\mathbf{P}(\mathrm{S}_r)=1.$$

Il vient

$$\begin{aligned}\mathbf{E}(\mathrm{X})&=\sum_{i=1}^{r}\left(1-x_i\right)^n=\sum_{i=1}^{r-1}\left(1-x_i\right)^n+\left(1-x_r\right)^n\\&=\sum_{i=1}^{r-1}\left(1-x_i\right)^n+\left(\sum_{i=1}^{r-1}x_i\right)^n.\end{aligned}$$

On définit alors la fonction f par

$$\forall\,(t_1,t_2,\dots,t_{r-1})\in(]0,1[)^{r-1},$$
$$f(t_1,t_2,\dots,t_{r-1})=\sum_{i=1}^{r-1}\left(1-t_i\right)^n+\left(\sum_{i=1}^{r-1}t_i\right)^n$$

3.b) f est polynomiale donc de classe $\mathscr{C}^2$ sur $(]0;1[)^{r-1}$. De plus, pour tout indice $i\in[\![1;r-1]\!]$,

$$\forall\,(t_1,t_2,\dots,t_{r-1})\in(]0,1[)^{r-1},$$
$$\partial_i f(t_1,t_2,\dots,t_{r-1})=-n\left(1-t_i\right)^{n-1}+n\left(\sum_{k=1}^{r-1}t_k\right)^{n-1}.$$

3.c) Le point $t=(t_1,\dots,t_{r-1})\in(]0;1[)^{r-1}$ est point critique si et seulement si pour tout $i\in[\![1;r-1]\!]$,

$$\partial_i f(t)=-n\left(1-t_i\right)^{n-1}+n\left(\sum_{k=1}^{r-1}t_k\right)^{n-1}=0$$

en posant $s=\left(\sum_{k=1}^{r-1}t_k\right)$, il vient

$$\forall\,i\in[\![1;r-1]\!],\quad \left(1-t_i\right)^{n-1}=s^{n-1}.$$

Comme $t_i\in]0;1[$ et $x\in]0;1[\mapsto x^{n-1}$ est injective, on en déduit que tous les t_i sont identiques.

$$t_1=t_2=\cdots=t_{r-1}=\frac{s}{r-1}.$$

Ainsi, $\left(1-t_i\right)^{n-1}=s^{n-1}$, puis $1-t_i=s=1-\frac{s}{r-1}$. Finalement,

$$t_i=1/r.$$

On a bien un unique point critique donné par

$$a=\left(\frac{1}{r},\frac{1}{r},\dots,\frac{1}{r}\right).$$

3.d) Vérifier que :

→ Pour $i\neq j$

$$\partial^2_{i,j}f(n)=n(n-1)\left(\sum_{k=1}^{r-1}t_k\right)^{n-2}.$$

→ Pour $i=j$

$$\partial^2_{i,j}f(1)=n(n-1)\left(1-t_i\right)^{n-2}+n(n-1)\left(\sum_{k=1}^{r-1}t_k\right)^{n-2}.$$

En particulier, pour le point critique, il vient :

$$\partial^2_{i,j}f(a)=n(n-1)\left(\frac{r-1}{r}\right)^{n-2}\quad\text{et}$$
$$\partial^2_{i,i}f(a)=2n(n-1)\left(\frac{r-1}{r}\right)^{n-2}.$$

D'où le résultat sur $\nabla^2 f(a)$.

3.e) Posons $p=r-1$.

Rappelons la diagonalisation de la matrice Atilla $\mathrm{J}\in\mathscr{M}_{p,1}(\mathbb{R})$. *On a* $\mathrm{rg}(\mathrm{J})=1$. *Par le théorème du rang, 0 est valeur propre et*

$$\dim\mathrm{E}_0(\mathrm{J})=\dim\mathscr{M}_{p,1}(\mathbb{R})-\mathrm{rg}(\mathrm{J})=p-1.$$

Comme J *est diagonalisable (symétrique), la seconde valeur propre se déduit de la trace*

$$p\times 0+\lambda=\mathrm{Tr}(\mathrm{J})=p.$$

D'où $\lambda=p$. *Pour des questions de dimensions, il n'y a pas d'autre valeur propre.*

- Posons $A = \nabla^2 f(a)$.

$$\begin{aligned}
\lambda \in \mathrm{Sp}(J) &\iff \mathrm{rg}\left(J - \lambda I_p\right) < p \\
&\iff \mathrm{rg}\left((I_p + J) - (\lambda+1) I_p\right) < p \\
&\iff \mathrm{rg}\left(A - n(n-1)\left(\frac{r-1}{r}\right)^{n-2}(\lambda+1) I_p\right) < p \\
&\iff n(n-1)\left(\frac{r-1}{r}\right)^{n-2}(\lambda+1) \in \mathrm{Sp}(A).
\end{aligned}$$

On en déduit que le spectre de A est formé de deux valeurs propres

$$n(n-1)\left(\frac{r-1}{r}\right)^{n-2} \quad \text{et} \quad n(n-1)(r+1)\left(\frac{r-1}{r}\right)^{n-2}.$$

3.f) On a $\mathrm{Sp}\left(\nabla^2 f(a)\right) \subset \mathbb{R}_*^+$ et f est définie sur un ouvert, on peut affirmer que f admet en a un minimum local. Par le calcul, on montre que le minimum vaut

$$f(a) = r\left(\frac{r-1}{r}\right)^n.$$

Exercice 11 p. 12

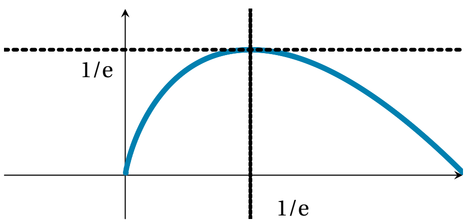
1. La fonction h_1 est dérivable sur $]0;1]$ par les théorèmes généraux avec pour tout $x \in]0;1]$

$$h_1'(x) = -\ln(x) - 1.$$

La dérivée change de signe en $1/\mathrm{e}$. De plus par les croissances comparées

$$\lim_{x\to 0^+} h_1(x) = \lim_{x\to 0^+} -x\ln(x) = 0 = h_1(0).$$

La fonction h_1 est continue en 0 (et donc sur $[0;1]$).



2. Dans cas

$$\forall k \in [\![1;n]\!], \quad p_k = \frac{1}{n}, \quad h_1(p_k) = \frac{1}{n}\ln(n)$$

et $\quad \mathrm{H}(Y) = \ln(n) \quad$ pour $\quad Y \hookrightarrow \mathscr{U}([\![1;n]\!])$.

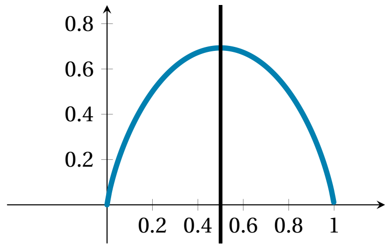
3. Si $Y \hookrightarrow \mathscr{B}(p)$, alors

$$p_1 = \mathbf{P}(Y=0) = 1-p, \quad p_2 = \mathbf{P}(Y=1) = p$$

et on trouve

$$\mathrm{H}(Y) = -(1-p)\ln(1-p) - p\ln(p).$$

Une étude de fonction donne un maximum pour $p = 1/2$.



4.a) Soit $k \in [\![1;n]\!]$, les applications coordonnées

$$x \in]0;1[^n \mapsto x_k$$

sont de classe $\mathscr{C}^2$ et à valeurs dans $]0;1[$. La fonction h_1 est de classe $\mathscr{C}^2$ sur $]0;1[$. Par composition, la fonction

$$x \in]0;1[^n \mapsto h_1\left(x_k\right)$$

est de classe $\mathscr{C}^2$. Par somme, h_n est de classe $\mathscr{C}^2$.
Soient $i, j \in [\![1;n]\!]$

$$\partial_i h_n(x) = h_1'\left(x_i\right) = -\left(\ln\left(x_i\right) + 1\right).$$

- Si $i \neq j$, $\quad \partial^2_{ij} h_n(x) = 0$;
- Si $i = j$, $\quad \partial^2_{ii} h_n(x) = \dfrac{1}{x_i}$.

Le gradient en x est :

$$\nabla h_n(x) = -\left(\ln\left(x_i\right) + 1, \ldots, \ln\left(x_n\right) + 1\right)$$

et la matrice hessienne est diagonale

$$\nabla^2 h_n(x) = \mathrm{diag}\left(-1/x_1, -1/x_2, \ldots, -1/x_n\right).$$

4.b) Posons $g : x \in \mathbb{R}^n \mapsto \sum_{k=1}^n x_k$. On a donc

$$\nabla g = (1, 1, \ldots, 1).$$

a est point critique sous la contrainte $g(x) = 1$ si et seulement si il existe $\lambda \in \mathbb{R}$ tel que

$$\nabla h_n(a) = \lambda \nabla g.$$

On en déduit que les coordonnées de $\nabla h_n(a)$ sont identiques et, par injectivité de la fonction logarithme que $x_1 = x_2 = \cdots = x_n$. Avec la contrainte $g(a) = 1$, on trouve

$$a = \frac{1}{n}(1, 1, \ldots, 1).$$

4.c) La i-ème coordonnée de $a + th$ est

$$(a+th)_i = \frac{1}{n} + t\left(x_i - \frac{1}{n}\right) = \frac{1}{n}(1-t) + t x_i.$$

Ce réel appartient au segment dont les extrémités sont $1/n$ et x_i. Comme $1/n$, $x_i \in]0;1[$, on peut affirmer que $(a+th)_i \in]0;1[$. C'est la définition de

$$a + th \in]0;1[^n.$$

4.d) Par le théorème sur les dérivées directionnelles, ψ est de classe $\mathscr{C}^2$ avec pour tout $t \in]0;1[$

$$\psi'(t) = \langle \nabla h_n(a+th), h \rangle.$$

En particulier, on a

$$\psi(0) = h_n(a) \quad \text{et} \quad \psi'(0) = \underbrace{\langle \nabla h_n(a), h \rangle}_{=0} = 0$$

car a est point critique sous la contrainte. De plus

$$\psi''(t) = q_{a+th}(h)$$

où q_a est la forme quadratique associée à la matrice hessienne $\nabla^2 h_n(a)$.
La formule de Taylor avec reste intégral s'écrit

$$\begin{aligned}\psi(1) &= \psi(0) + \psi'(0)(1-0) + \int_0^1 \psi''(t)(1-t)\,\mathrm{d}t \\ &= h_n(a) + \int_0^1 q_{a+th}(h)(1-t)\,\mathrm{d}t.\end{aligned}$$

Or on a vu à la question 4.a) que la matrice hessienne est diagonale avec des coefficients diagonaux strictement négatifs. Ainsi, la forme quadratique q_{a+th} est négative. Concluons

$$\begin{aligned}h_n(x) - h_n(a) &= \psi(1) - \psi(0) \\ &= \int_0^1 \underbrace{q_{a+th}(h)(1-t)}_{\leqslant 0}\,\mathrm{d}t \leqslant 0 \qquad (\star)\end{aligned}$$

Pour tout $x \in]0;1[^n$ avec $\sum\limits_{k=1}^{n} x_k = 1$

$$h_n(x) \leqslant h_n(a).$$

Le réel $h_n(a)$ est un maximum global sous la contrainte.

- Justifions que la maximum est atteint uniquement en a.

Argument 1
Il y a unicité du point point critique sous la contrainte donc il n'y a pas d'autre point où la maximum est atteint.

Argument 2
Précisons que pour $x \neq a$, l'inégalité $(\star)$ est stricte dès que $x \neq a$ car $h \neq 0$ et pour $t \in]0;1[$

$$q_{a+th}(h)(1-t) < 0$$

car le spectre est strictement inclus dans $\mathbb{R}_*^+$.

5. D'après ce qui précède, l'entropie est maximale lorsque la variable est uniforme puisque dans ce cas

$$p_1 = p_2 = \ldots = p_n = \frac{1}{n}.$$

Remarque. L'entropie (de Shannon) définie dans cet exercice est une mesure de l'information disponible. Plus l'entropie est petite, plus l'information est importante (prendre le cas extrême $p = 1$ dans la question 2). On comprend alors que dans le cas d'une loi uniforme où toutes les valeurs de la variable sont possibles avec la même probabilité, l'information disponible est la plus faible et l'entropie la plus grande.