Université Paris I, Panthéon-Sorbonne

Régime Aménagé - UFR 02
Mathématique 2
Cours d'Optimisation Libre Dans $\mathbb{R}^n$

Alexis Fauth (Université Paris I)

1. Rappels Et Compléments.

1.1. Premières Définitions.

Définition 1.1.1. On dira que f est de classe C^2 sur un ensemble I si elle admet des dérivées partielles secondes au voisinage de $X \in I$ qui sont continues en X. On notera alors $f \in C^2(I)$.

Soit E un ouvert de $\mathbb{R}^n$ et, soit f une fonction de classe C^2 sur E.

Définition 1.1.2. (extrema local) Soit $f : E \in \mathbb{R}^n \mapsto \mathbb{R}$. On dit que f présente en $X^* \in E$ un extremum local, ssi il existe un voisinage $\mathcal{V}(X^*) = B(X^*, \rho)$, $\rho \in \mathbb{R}$ de X^* tel que,

$$\begin{array}{ll} \forall X \in B(X^*, \rho) \cap E, f(X^*) \leq f(X) & X^* \text{ sera alors appelé minimum local de } f \text{ sur } E \\ \forall X \in B(X^*, \rho) \cap E, f(X^*) \geq f(X) & X^* \text{ sera alors appelé maximum local de } f \text{ sur } E \end{array}$$

On dira qu'il s'agit d'un extrema local strict si les inégalités précedentes sont strict pour tout $X^* \neq X$.

Définition 1.1.3. (extrema global) On dit que $X^* \in E$ est un extremum global de $f : E \in \mathbb{R}^n \mapsto \mathbb{R}$ si,

$$\begin{array}{ll} \forall X \in E, f(X^*) \leq f(X) & X^* \text{ sera alors appelée minimum global de } f \text{ sur } E \\ \forall X \in E, f(X^*) \geq f(X) & X^* \text{ sera alors appelée maximum global de } f \text{ sur } E \end{array}$$

On dira qu'il s'agit d'un extrema local strict si les inégalités précedentes sont strict pour tout $X^* \neq X$

Définition 1.1.4. (Convexe, concave) Une fonction f sur $I \in \mathbb{R}^n$ est convexe (resp concave) si pour tout $X, Y \in I$, $\lambda \in [0, 1]$,

$$\begin{aligned} f(\lambda X + (1 - \lambda Y)) &\leq \lambda f(X) + (1 - \lambda) f(Y) \\ \text{resp } f(\lambda X + (1 - \lambda Y)) &\geq \lambda f(X) + (1 - \lambda) f(Y) \end{aligned}$$

Définition 1.1.5. (Strict convexité, concavité) Une fonction f sur $I \in \mathbb{R}^n$ est strictement convexe (resp concave) si pour tout $X \neq Y \in I$, $\lambda \in]0, 1[$,

$$\begin{aligned} f(\lambda X + (1 - \lambda Y)) &< \lambda f(X) + (1 - \lambda) f(Y) \\ \text{resp } f(\lambda X + (1 - \lambda Y)) &> \lambda f(X) + (1 - \lambda) f(Y) \end{aligned}$$

Proposition 1.1.1. Une fonction f de classe C^2 sur $I \in \mathbb{R}$ est dite convexe (resp concave) si pour tout $X \in I$, $f''(X) \geq 0$ (resp $f''(X) \leq 0$).

1.2. Optimisation Dans $\mathbb{R}$.

On rappel ici la méthodologie pour trouver les extremas d'une fonction $f : \mathbb{R} \to \mathbb{R}$,

1. La fonction est-elle de classe C^1 ?
2. Détermination des extremas en résolvant $f'(x) = 0$.

3. On cherche la nature de(s) extrema(s) $\rightarrow$ Tableau de variation, la fonction est-elle concave, convexe ?

Exemple de Cobb Douglas :

$$f(x) = x^{\alpha}\left(\frac{R-p_1x}{p_2}\right)^{1-\alpha} \qquad R>0,\ p_1>0,\ p_2>0,\ x>0, \alpha\in]-1,1[$$

f est de classe C^1 pour $x>0$ et $\dfrac{R-p_1x}{p_2}>0$,

$$f'(x) = \alpha x^{\alpha-1}\left(\frac{R-p_1x}{p_2}\right)^{1-\alpha} - \frac{p_1}{p_2}(1-\alpha)x^{\alpha}\left(\frac{R-p_1x}{p_2}\right)^{-\alpha}$$

On résout $f'(x)=0$,

$$\begin{aligned}
& \alpha x^{\alpha-1}\left(\frac{R-p_1x}{p_2}\right)^{1-\alpha} - \frac{p_1}{p_2}(1-\alpha)x^{\alpha}\left(\frac{R-p_1x}{p_2}\right)^{-\alpha} &= 0\\
\Leftrightarrow\quad & \alpha x^{-1}\left(\frac{R-p_1x}{p_2}\right)^{1} - \frac{p_1}{p_2}(1-\alpha) &= 0\\
\Leftrightarrow\quad & \alpha\left(\frac{R-p_1x}{p_2}\right) &= x\frac{p_1}{p_2}(1-\alpha)\\
\Leftrightarrow\quad & \alpha\frac{R}{p_2} &= x\left(\alpha(\frac{p_1}{p_2})+\frac{p_1}{p_2}(1-\alpha)\right)\\
\Leftrightarrow\quad & \alpha\frac{R}{p_2} &= \frac{p_1}{p_2}x\\
\Leftrightarrow\quad & \frac{\alpha R}{p_1} &= x^*
\end{aligned}$$

Je vous laisse determiner si cette fonction est convexe ou concave.

2. Optimisation Dans $\mathbb{R}^n$.

2.1. Condition Necessaire du Premier Ordre.

Théorème 2.1.1. (CN1) Soit f de classe $C^1(E)$. Si f admet un extremum local au point $X^* = (x_1, \cdots, x_n)$ de l'ouvert E, alors X^* vérifie

$$\bigtriangledown f(X^*) = 0_{\mathbb{R}^n}$$

où $\bigtriangledown f(X^*)^t = (\frac{\partial f(X^*)}{\partial x_1}, \cdots, \frac{\partial f(X^*)}{\partial x_n})$. X^* sera appelé point candidat ou point critique.

Preuve. (sur $\mathbb{R}^2$) par hypothèse, f admet un extremum local en $X^*=(x^*,y^*)$ sur la boule ouverte centrée en X^* et de rayon α, on $\forall X\in B(X^*,\alpha), f(X)\le f(X^*)$ (resp $f(X^*)\le f(X)$) si l'optimum est un maximum (resp minimum) local.

Soit $H(h_1,h_2)$ un élément quelconque de $\mathbb{R}^2$. Considérons la fonction g de la variable réelle t définie comme suit : $\forall t\in\mathbb{R}, g(t) = f(X^*+tH)$. Elle vérifie donc,

- $g(0) = f(X^*)$

- Elle admet un extremum local en $t=0$. En effet, pour tout t tel que $X^*+tH\in B(X^*,\alpha)$, autrement dit pour tout t tel que $|t|<\frac{\alpha}{||H||}$, on a : $g(t) = f(X^*+tH)\le g(0) = f(X^*)$.

- Comme f est de classe C^1 sur E, a fortiori sur $B(X^*,\alpha)$, alors g est dérivable au voisinage de 0 et on a, pour tout t de l'ouvert $\left]-\frac{\alpha}{||H||};\frac{\alpha}{||H||}\right[$, $g'(t) = \frac{\partial f(X^*+tH)}{\partial x}h_1 + \frac{\partial f(X^*+tH)}{\partial y}h_2$. En particulier $g'(0) = \frac{\partial f(X^*)}{\partial x}h_1+\frac{\partial f(X^*)}{\partial y}h_2$.

Puisque g admet un maximum local en 0 et est de classe C^1 au voisinage de 0, on peut appliquer la condition du premier ordre $g'(0)=0$. D'où, $\frac{\partial f(X^*)}{\partial x}h_1+\frac{\partial f(X^*)}{\partial y}h_2 = 0$. Comme l'égalité est vraie pour tout $H=(h_1,h_2)$, on en déduit, $\frac{\partial f(X^*)}{\partial x} = \frac{\partial f(X^*)}{\partial y} = 0$. ■

Remarque 2.1.1. La CN1 signifie géométriquement que le plan tangent à la surface d'équation $y = f(x_1, x_2)$ au point X^* de coordonées $(x_1^*, x_2^*, f(x_1^*, x_2^*))$ est horizontal. On peut donc avoir trois cas possible, Fig[1], Fig[2] et Fig[3]. On a ainsi besoin d'une condition du second ordre pour caractériser la nature des points critiques.

2.2. Condition Nécessaire Du Second Ordre.

Les conditions du second ordre permettent de distinguer les cas d'un maximum local et d'un minimum local

Théorème 2.2.1. (Schwartz) Soit f une fonction de classe C^2 sur un ensemble ouvert $E \in \mathbb{R}^n$. Le théorème de Schwartz nous assure que

$$\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_j \partial x_i} \qquad \forall i, j \in \{1, \cdots, n\}$$

Contre exemple :

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Définition 2.2.1. (Hessien) Soit f une fonction de classe C^2 sur un ensemble ouvert $E \in \mathbb{R}^n$. On définit la matrice hessienne, $H(f, X)$, de f au point $X = (x_1, \cdots, x_n) \in \mathbb{R}^n$ par,

$$H(f,X) = \begin{pmatrix} \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_1} & \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_2} & \cdots & \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(x_1,\cdots,x_n)}{\partial x_n^2} \end{pmatrix}$$

D'après le théorème de Schwartz, la matrice hessienne est symétrique.

Définition 2.2.2. Une fonction $f : \mathbb{R}^n \mapsto \mathbb{R}$ de classe C^2 sur $\mathbb{R}^n$ est convexe (resp concave) si sa matrice Hessiene est semi-définie postive (resp négative) pour tout $x \in \mathbb{R}^n$.

Définition 2.2.3. Une fonction $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ de classe C^2 sur $\mathbb{R}^n$ est strictement convexe (resp concave) si sa matrice hessiene est définie postive (resp négative) pour tout $x \in \mathbb{R}^n$.

Théorème 2.2.2. (CN2 pour un max local) Soit f de classe C^2 sur E et soit X^* de l'ouvert E. Si f admet un maximum local en X^*, alors on a, $\forall H = (h_1, \cdots, h_n) \in \mathbb{R}^n$,

$$\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1^2} h_1^2 + 2\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1 \partial x_2} h_1 h_2 + 2\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1 \partial x_3} h_1 h_3 + \cdots + \frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_n^2} h_n^2 \leq 0$$

Soit sous forme matricielle, $\forall Z \in \mathbb{R}^n$,

$$Z^t H(f, X^*) Z \leq 0$$

Soit encore, la matrice hessienne au point X^* est semi-définie négative.

Preuve. (Dans $\mathbb{R}^2$) La démonstration et la suite directe de celle donnée pour la condition nécessaire du premier ordre. La fonction d'une variable réelle g est de classe C^2 dans un voisinage de 0 puisque f est de classe C^2 au voisinage de $B(X^*, \alpha)$ de X^*. On a, $g''(t) = \frac{\partial^2 f(X^*+tH)}{\partial x^2} h_1^2 + 2\frac{\partial^2 f(X^*+tH)}{\partial x \partial y} h_1 h_2 + \frac{\partial^2 f(X^*+tH)}{\partial y^2} h_2^2$. Comme g atteint un max local en 0, la condition nécessaire du second ordre pour une fonction d'une variable réelle est vérifiée en ce point. On a donc : $g''(0) \leq 0$, soit $\forall H = (h_1, h_2) \in \mathbb{R}^2$, $\frac{\partial^2 f(X^*)}{\partial x^2} h_1^2 + 2\frac{\partial^2 f(X^*)}{\partial x \partial y} h_1 h_2 + \frac{\partial^2 f(X^*)}{\partial y^2} h_2^2 \leq 0$.
■

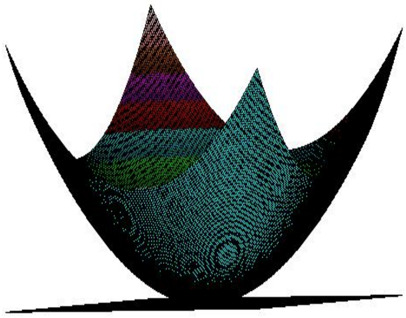

Fig. 1 – Fonction convexe d'équation $f(x, y) = x^2 + y^2$ et son plan tangent d'équation $z = x + y$. Le point critique est ici un minimun global strict.

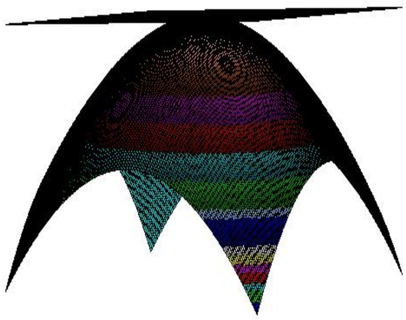

Fig. 2 – Fonction concave d'équation $g(x, y) = -x^2 - y^2$ et son plan tangent d'équation $z = x + y$. Le point critique est ici un maximum global strict.

Fig. 3 – Fonction d'équation $h(x, y) = -x^2 + y^2$ et son plan tangent d'équation $z = x + y$. Le point critique est ici un point selle (ni minimum, ni maximum.

Théorème 2.2.3. (CN2 pour un min local) Sous les même hypothèses que celles du Th 2.2.2., on obtient alors, $\forall H = (h_1, \cdots, h_n) \in \mathbb{R}^n$,

$$\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1^2} h_1^2 + 2\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1 \partial x_2} h_1 h_2 + 2\frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_1 \partial x_3} h_1 h_3 + \cdots + \frac{\partial^2 f(x_1, \cdots, x_n)}{\partial x_n^2} h_n^2 \geq 0$$

Soit sous forme matricielle, $\forall Z \in \mathbb{R}^n$,

$$Z^t H(f, X^*) Z \geq 0$$

soit encore, la matrice hessienne au point X^* est semi-définie positive.

Théorème 2.2.4. (CN2) Les théorèmes 2.2.2. et 2.2.3. s'etendent pour des extremas globaux si les matrices hessienne sont définies de la même manière pour tout $X \in \mathbb{R}^n$ et non, seulement au point X^*. Et de manière évidente, les extremas globaux ou locaux seront strict si les hessiens sont définis positifs pour un minimum et définis négatifs pour un maximum.

2.3. Condition Suffisante du Second Ordre.

Théorème 2.3.1. (CS2 pour un max local) Si l'une des conditions suivantes est vérifiée,

- $$\bigtriangledown(f) = 0,\ \forall Z \in \mathbb{R}^n,\ Z^t H(f, X^*) Z \leq 0$$
- X^* est un point candidat et, la matrice hessienne au point X^* est semi def négative

Alors f admet un max local au point X^*

Théorème 2.3.2. (CS2 pour un max local strict) Si l'une des conditions suivantes est vérifiée,

- X^* est un point candidat et, $\forall Z \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\}$, $Z^t H(f, X^*) Z < 0$
- X^* est un point candidat et, la matrice hessienne au point X^* est def négative

Alors f admet un max local strict au point X^*

Preuve. (Dans $\mathbb{R}^2$) La fonction f étant de classe C^2 sur E, elle admet un devellopement de Taylor à l'ordre 2 en X^*. Pour tout $H = (h_1, h_2) \in \mathbb{R}^2$, on a,

$$f(X^* + H) = f(X^*) + f'_x(X^*) h_1 + f'_y(X^*) h_2 + \frac{1}{2}\left[f''_{x^2}(X^*) h_1^2 + 2f''_{xy}(X^*) h_1 h_2 + f''_{y^2}(X^*) h_2^2\right] + ||H||^2 \epsilon(H)$$

avec $\lim_{H \to 0} \epsilon(H) = 0$.
Comme X^* est un point candidat, on a,

$$f(X^* + H) - f(X^*) = \frac{1}{2}\left[f''_{x^2}(X^*) h_1^2 + 2f''_{xy}(X^*) h_1 h_2 + f''_{y^2}(X^*) h_2^2\right] + ||H||^2 \epsilon(H)$$

avec $\lim_{H \to 0} \epsilon(H) = 0$.
Considérons tout d'abord la forme quadratique sur $\mathbb{R}^2$: $q(H) = f''_{x^2}(X^*) h_1^2 + 2f''_{xy}(X^*) h_1 h_2 + f''_{y^2}(X^*) h_2^2$. Par hypothèse elle est définie négative. Elle vérifie donc $q(H) \leq -\alpha ||H||^2$ avec $\alpha > 0$ (le nombre $-\alpha$ est la borne supérieure de la fonction continue $q(H)$ sur l'ensemble $\{H \in \mathbb{R}^n / ||H|| = 1\}$).
D'autre part, comme $\lim_{H \to 0} \epsilon(H) = 0$, il existe $\eta > 0$ tel que $||H|| < \eta \implies |\epsilon(H)| < \frac{\alpha}{2}$. Puisque $\alpha > 0$, on en déduit : $||H|| < \eta \implies \epsilon(H) < \frac{\alpha}{2}$.

En définitive, pour $||H|| < \eta$, on a,

$$\frac{1}{2}\left[f''_{x^2}(X^*) h_1^2 + 2f''_{xy}(X^*) h_1 h_2 + f''_{y^2}(X^*) h_2^2\right] + ||H||^2 \epsilon(H) \leq -\frac{\alpha}{2} ||H||^2 + ||H||^2 \epsilon(H) = ||H||^2 \left(\epsilon(H) - \frac{\alpha}{2}\right) < 0$$

Par conséquent, $\forall H$ tel que $||H|| < \eta$, $f(X^* + H) - f(X^*) < 0$. X^* est donc bien max local strict. ■

Théorème 2.3.3. (CS2 pour un max global) Si l'une des conditions suivantes est vérifiée,

- X^* est un point candidat et, $\forall Z \in \mathbb{R}^n$, $Z^t H(f, X) Z \leq 0$
- X^* est un point candidat et, la matrice hessienne pour tout $X \in E$ est semi def négative

Alors f admet un max global au point X^*

Théorème 2.3.4. (CS2 pour un max global strict) Si l'une des conditions suivantes est vérifiée,

- X^* est un point candidat et, $\forall Z \in \mathbb{R}^n - \{0_{\mathbb{R}}^n\}$, $Z^t H(f, X) Z < 0$
- X^* est un point candidat et, la matrice hessienne pour tout $X \in E$ est définie négative

Alors f admet un max global strict au point X^*

Théorème 2.3.5. (CS2 pour un point selle) Considérons un point X^* vérifiant la condition du premier ordre d'optimalité et non celles du second ordre, i.e. que la matrice H n'est pas définie, aussi bien pour tout $X \in E$ que pour X^*, on ne peut donc pas le caracteriser comme maximum ou minimum, local, global, strict, non-strict. On dira alors que X^* est un point selle.

2.4. La Simplification Dans $\mathbb{R}^2$.

Proposition 2.4.1. Soit A une matrice de $\mathcal{M}_{n \times n}(\mathbb{R})$ diagonalisable de valeurs propres λ_i avec $i = \{1, \cdots, n\}$. Alors,

$$Det(A) = \prod_{i=1}^{n} \lambda_i$$

$$tr(A) = \sum_{i=1}^{n} \lambda_i$$

Proposition 2.4.2. Dans le cas d'une fonction définie sur $\mathbb{R}^2$ cette proposition simplifie les calculs. En effet, il n'est pas necessaire de determiner de manière explicite les valeurs propres du hessien, par exemple, si $Det(H) > 0$ alors necessairement les vaps sont de même signes, ont est donc en présence d'un max ou d'un min, et si la trace est positive, les vap sont positives, le hessien est donc défini positif, il s'agit d'un min strict (global si Det(H)>0 pour tout $X \in E$).

3. Applications.

Exercice 1. Soit la fonction f définie par,

$$f(x, y) = 2x^2 + 3y^2 + 4xy + 2x - y + 6$$

Déterminer le(s) extrema(s) de f et le(s) caracteriser.

Correction. La fonction est de classe C^∞ sur l'ouvert $\mathbb{R}^2$ en tant que fonction polynômiale, on peut donc determiner son gradient et son hessien,

$$\bigtriangledown f(x, y) = (4x + 4y + 2, 6y + 4x - 1)$$

Résolvons $\bigtriangledown f(x, y) = (0, 0)$ pour déterminer le(s) point(s) critique(s), $X^* = (x^*, y^*)$,

$$\begin{cases} 4x + 4y + 2 & = 0 \\ 4x + 6y - 1 & = 0 \end{cases}$$

Ainsi, $X^* = (-2, 3/2)$. Déterminons maintenant la nature de ce point critique. Pour cela on écrit le hessien,

$$H(f,X) = \begin{pmatrix} 4 & 4 \\ 4 & 6 \end{pmatrix}$$

La matrice hessienne étant de taille 2×2 pour avoir le signe des valeurs propres il suffit de regarder son déterminant. Det(H)=8 donc les vaps sont de même signes et differentes de 0 et, tr(H)=10 donc les vaps sont positives strictement. On est donc en présence d'un minimum strict, et comme ces résultats ont été trouvé pour tout $X \in \mathbb{R}^2$ il s'agit d'un minimum global strict.

Exercice 2. Soit f la fonction définie sur l'ouvert $]0,1[\times]0,1[$ de $\mathbb{R}^2$ par,

$$f(x,y) = (1-x)^n + (1-y)^n + (x+y)^n \qquad n \geq 2$$

.

1. Déterminer les dérivées partielles d'ordre 1 de f.
2. En déduire que le seul point en lesquelles les dérivées partielles d'ordre 1 de f s'annulent simultanément est le point $\left(\frac{1}{3}, \frac{1}{3}\right)$.
3. Démontrer que f présente un minimum local en ce point.

Correction.

1. Il est clair que la fonction f est de classe C^2 sur l'ouvert $]0,1[\times]0,1[$ en tant que fonction polynômiale, on peut donc déterminer son gradient et sa matrice hessienne.
$\frac{\partial f}{\partial x}(x,y) = -n(1-x)^{n-1} + n(x+y)^{n-1} \quad \frac{\partial f}{\partial y}(x,y) = -n(1-y)^{n-1} + n(x+y)^{n-1}$
Ainsi,
$$\triangledown f(X) = \left(-n(1-x)^{n-1} + n(x+y)^{n-1}, -n(1-y)^{n-1} + n(x+y)^{n-1}\right)$$

2. On résout directement le système, en tenant compte que $a^{n-1} = b^{n-1} \Leftrightarrow a = b$ lorsque a et b sont positifs (ce qui est le cas pour x et y), on a
$$\left\{\begin{array}{l} \frac{\partial f}{\partial x}(x,y) = 0 \\ \frac{\partial f}{\partial y}(x,y) = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} -n(1-x)^{n-1} + n(x+y)^{n-1} = 0 \\ -n(1-y)^{n-1} + n(x+y)^{n-1} = 0 \end{array}\right.$$
$$\Leftrightarrow \left\{\begin{array}{l} -(1-x)^{n-1} + (x+y)^{n-1} = 0 \\ -(1-y)^{n-1} + (x+y)^{n-1} = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} (1-x)^{n-1} = (x+y)^{n-1} \\ (1-y)^{n-1} = (x+y)^{n-1} \end{array}\right.$$
$$\Leftrightarrow \left\{\begin{array}{l} 1-x = x+y \\ 1-y = x+y \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} 3x = 1 \\ 3y = 1 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} x = \frac{1}{3} \\ y = \frac{1}{3} \end{array}\right.$$

3. Le point $\left(\frac{1}{3}, \frac{1}{3}\right)$ étant un point critique de f (il annule les deux dérivées partielles d'ordre 1), et $]0,1[\times]0,1[$ étant un ouvert de $\mathbb{R}^2$, on doit déterminer les dérivées partielles du second ordre au point $\left(\frac{1}{3}, \frac{1}{3}\right)$
$$\left\{\begin{array}{l} \frac{\partial^2 f}{\partial x^2}(x,y) = n(n-1)(1-x)^{n-2} + n(n-1)(x+y)^{n-2} \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y) = n(n-1)(x+y)^{n-2} \\ \frac{\partial^2 f}{\partial y^2}(x,y) = n(n-1)(1-y)^{n-2} + n(n-1)(x+y)^{n-2} \end{array}\right.$$

$$\Rightarrow \begin{cases} \dfrac{\partial^2 f}{\partial x^2}\left(\dfrac{1}{3},\dfrac{1}{3}\right) = 2n(n-1)\left(\dfrac{2}{3}\right)^{n-2} \\ \dfrac{\partial^2 f}{\partial x \partial y}\left(\dfrac{1}{3},\dfrac{1}{3}\right) = n(n-1)\left(\dfrac{2}{3}\right)^{n-2} \\ \dfrac{\partial^2 f}{\partial y^2}\left(\dfrac{1}{3},\dfrac{1}{3}\right) = 2n(n-1)\left(\dfrac{2}{3}\right)^{n-2} \end{cases}$$

Par conséquent,

$$H\left(f\left(\frac{1}{3},\frac{1}{3}\right)\right) = \begin{pmatrix} 2n(n-1)\left(\frac{2}{3}\right)^{n-2} & n(n-1)\left(\frac{2}{3}\right)^{n-2} \\ n(n-1)\left(\frac{2}{3}\right)^{n-2} & 2n(n-1)\left(\frac{2}{3}\right)^{n-2} \end{pmatrix}$$

et donc,

$$\begin{aligned} Det(H\left(f\left(\frac{1}{3},\frac{1}{3}\right)\right)) &= \left[2n(n-1)\left(\frac{2}{3}\right)^{n-2}\right]^2 - \left[n(n-1)\left(\frac{2}{3}\right)^{n-2}\right]^2 \\ &= 4n^2(n-1)^2\left(\frac{2}{3}\right)^{2n-4} - n^2(n-1)^2\left(\frac{2}{3}\right)^{2n-4} \\ &= 3n^2(n-1)^2\left(\frac{2}{3}\right)^{2n-4} > 0 \end{aligned}$$

ce qui implique que le point $\left(\dfrac{1}{3},\dfrac{1}{3}\right)$ est un extrémum local de f sur l'ouvert $]0,1[\times]0,1[$

Ensuite, $Tr\left(H\left(f\left(\frac{1}{3},\frac{1}{3}\right)\right)\right) = 4n(n-1)\left(\dfrac{2}{3}\right)^{n-2} > 0$ donc le point $\left(\dfrac{1}{3},\dfrac{1}{3}\right)$ est un minimum local de f sur l'ouvert $]0,1[\times]0,1[$.

Exercice 3. Soit $g(x) = \ln\ln x$.

1. Donner le domaine de définition de g.
2. Montrer que g est concave sur son domaine de définition.
3. En déduire l'inégalité
$$\forall a > b > 1, \quad \ln\frac{a+b}{2} > \sqrt{\ln a \ln b}.$$

Correction.

1. $D_g =]1,+\infty[$.
2. La fonction g est une composé de fonction ln elle est donc dérivable sur l'ouvert D_g et, $g'(x) = \frac{1}{x\ln x}$. $g''(x) = \frac{-(\ln(x)+1)}{(x\ln x)^2} < 0$ car $\ln(x)+1 > 1 > 0$. Donc g est strictement concave sur $]1,+\infty[$.
3. Par concavité, on a $g\left(\frac{a+b}{2}\right) > \frac{g(a)+g(b)}{2}$, autrement dit $\ln\left(\ln\frac{a+b}{2}\right) > \frac{\ln(\ln a)+\ln(\ln b)}{2} = \ln((\ln a\ln b)^{1/2})$. On prend l'exponentielle et on obtient $\ln\left(\frac{a+b}{2}\right) > (\ln a\ln b)^{1/2}$.

4. Optimisation Sous Contraintes, Ouverture.

Un programme d'optimisation sous contrainte est de la forme, pour E un sous ensemble de $\mathbb{R}^n$ dévrivant $(x_1, x_2, \cdots, x_n)$,

$$\begin{cases} \min_{X\in E} & f(X) & \\ sc & g_1(X) & \leq 0 \\ & g_2(X) & \leq 0 \\ & g_3(X) & = 0 \\ & \cdots & \end{cases} \quad \text{ou} \quad \begin{cases} \max_{X\in E} & f(X) & \\ sc & g_1(X) & \leq 0 \\ & g_2(X) & \leq 0 \\ & g_3(X) & = 0 \\ & \cdots & \end{cases}$$

Pour des cas simple, on pourra utiliser la méthode de subsitution, on reprend l'exemple de Cobb Douglas,

$$\begin{cases} \max_{(x,y)\in E\times F} x^\alpha y^{1-\alpha} \\ sc\; xp_1 + yp_2 = R \end{cases}$$

ce programme d'optimisation équivaut à résoudre le même programme que celui proposé dans la première partie avec une seule variable.

$$\left\{ \begin{array}{c} \max_{(x,y)\in\mathbb{R}_+^2} x^\alpha y^{1-\alpha} \\ sc\, xp_1 + yp_2 = R \end{array} \right. \quad \Leftrightarrow \left\{ \begin{array}{c} \max_{(x,y)\in\mathbb{R}_+^2} x^\alpha \left(\frac{R-p_1x}{p_2}\right)^{1-\alpha} \\ sc\, y = \frac{R-p_1x}{p_2} \end{array} \right.$$

On trouve donc, $x^* = \alpha\frac{R}{p_1}$ et en reprenant la contrainte d'équation $g(x,y) = xp_1 + yp_2 = R$ on trouve, $y^* = (1-\alpha)\frac{R}{p_2}$

De manière évidente, on ne peut pas toujours utiliser cette méthode, par exemple si l'on a un nombre élevé de contraintes, s'il ne s'agit pas de contraintes d'égalité mais d'inégalités... Heureusement, il existe des outils plus puissants, les méthodes de Programmation Linéaire comme, par exemple, la méthode du Simplexe si la fonction à optimiser est linéaire (cette méthode à déjà été entrevue par ceux qui ont fait un bac ES). Pour d'autres programmes plus complexes, les méthodes du Lagrangien ou de Karush, Kuhn and Tucker (KKT) semblent appropriées. Ces deux dernières méthodes sont présentés sur la brochure de cours distribué en début de semestre. Enfin, notons que les méthodes proposé dans ce mini cours d'optimisation libre sur $\mathbb{R}^n$ ainsi que les méthodes du Lagrangien et de KKT imposent de pouvoir dériver les fonctions, on peut donc se poser une question, comment optimiser une fonction s'il elle n'est pas différentiable ?