Université de M'sila
Faculté des Mathématiques de l'Informatique
Département des mathématiques
Licence 3ième année
Module: Optimisation sans contraintes (2020/2021)

Minimisation d'une fonction de plusieurs variables

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de plusieurs variables. La forme générale d'un problème d'optimisation sans contraintes est donné par

$$(P) : \begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases}$$

Points critiques ou candidats

Soit $x^* \in \mathbb{R}^n$, x^* est un *point critique* ou *candidat* ou *point stationnaire* si

$$\nabla f(x^*) = 0.$$

Condition d'optimalité du 1er ordre

Théorème (*Condition nécessaire du premier ordre*)
Soit x^* un point extremum local de f alors

$$\nabla f(x^*) = 0.$$

Cette équation est appelée équation d'Euler.

Preuve
Sans perte de généralité, soit $x^* \in \mathbb{R}^n$ un minimum local de f sur le voisinage V. Il existe $r > 0$ tel que

$$x^* + td \in V \text{ avec } t < r \text{ et } d \in \mathbb{R}^n$$

Nous avons donc,

$$\begin{aligned} f(x^* + td) &\geq f(x^*) \\ &\Rightarrow \frac{f(x^* + td) - f(x^*)}{t} \geq 0 \\ &\Rightarrow \lim_{t \to 0} \frac{f(x^* + td) - f(x^*)}{t} \geq 0 \end{aligned}$$

c'est à dire $\langle \nabla f(x^*), d \rangle \geq 0$, et comme d est libre dans $\mathbb{R}^n$, on peut prendre $-d$ on trouve

$$\langle \nabla f(x^*), d \rangle \leq 0$$

d'où

$$\langle \nabla f(x^*), d \rangle = 0 \Rightarrow \nabla f(x^*) = 0. \qquad \blacksquare$$

Condition d'optimalité du 2nd ordre

Théorème (*Condition nécessaire de deuxième ordre*)

1) *Si f présente un minimum local en x^* alors*

$$\nabla f(x^*) = 0 \text{ et } \nabla^2 f(x^*) \text{ } \textit{est une matrice semi définie positive.}$$

2) *Si f présente un maximum local en x^* alors*

$$\nabla f(x^*) = 0 \textit{ et } \nabla^2 f(x^*) \text{ } \textit{est une matrice semi définie négative.}$$

Preuve

1) Posons

$$\varphi(t) = f(x^* + td),$$

on suppose ici que $x^* + td \in V$ pour tout $d \in \mathbb{R}^n$ et $t < r$. En utilisant le développement de Taylor à l'ordre 2 de φ au voisinage de 0

$$\varphi(t) = \varphi(0) + \varphi^{'}(0)\,t + \frac{t^2}{2}\varphi^{''}(0) + t^2\epsilon(t).$$

avec $\lim_{t\to 0}\epsilon(t) = 0$. Alors

$$\begin{aligned} f(x^* + td) &= f(x^*) + \langle \nabla f(x^*), d\rangle\, t + \frac{t^2}{2} d^T \nabla^2 f(x^*)\, d + t^2\epsilon(t) \\ &= f(x^*) + \frac{t^2}{2} d^T \nabla^2 f(x^*)\, d + t^2\epsilon(t) \end{aligned}$$

ce qui implique

$$\frac{f(x^* + td) - f(x^*)}{t^2} = \frac{1}{2} d^T \nabla^2 f(x^*)\, d + t^2\epsilon(t) \geq 0$$

on fait $t \to 0$, on trouve

$$d^T \nabla^2 f(x^*)\, d \geq 0,$$

comme d est arbitraire, $\nabla^2 f(x^*)$ est semi définie positive.

2) Une démonstration similaire. ■

Condition suffisante

Rappel/

Si A est définie positive alors

$$\forall x \in \mathbb{R}^n : x^T A x \geq \lambda_{\min} \|x\|^2,$$

avec $\lambda_{\min} > 0$ est la plus petite valeurs propres de A.

Preuve.

Comme A est symétrique, il existe une base orthonormé de vecteurs propres de $\mathbb{R}^n$. Pour $x \in \mathbb{R}^n$ on a

$$x = \sum_{i=}^{n} \alpha_i v_i,$$

donc

$$
\begin{aligned}
x^T A x &= \left\langle A(\sum_{i=}^{n} \alpha_i v_i), \sum_{i=}^{n} \alpha_i v_i \right\rangle = \left\langle \sum_{i=1}^{n} \alpha_i A v_i, \sum_{i=}^{n} \alpha_i v_i \right\rangle \\
&= \left\langle \sum_{i=}^{n} \alpha_i \lambda_i v_i, \sum_{i=}^{n} \alpha_i v_i \right\rangle \\
&\geq \lambda_{\min} \left\langle \sum_{i=}^{n} \alpha_i v_i, \sum_{i=}^{n} \alpha_i v_i \right\rangle = \lambda_{\min} \|x\|^2 . \qquad \blacksquare
\end{aligned}
$$

Théorème (*Condition suffisante***)**

1) Si $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est une matrice définie positive, alors f admet un minimum local en x^*.

2) Si $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est une matrice définie négative, alors f admet un maximum local en x^*.

Preuve.

Soit $\varphi(t) = f(x^* + td)$. Le développement de Taylor à l'ordre 2 de φ au voisinage de 0 donne

$$\varphi(t) - \varphi(0) = \varphi'(0)\, t + \frac{t^2}{2} \varphi''(0) + t^2 \epsilon(t) .$$

avec $\lim_{t \to 0} \epsilon(t) = 0$. Alors

$$
\begin{aligned}
f(x^* + td) - f(x^*) &= \frac{t^2}{2} d^T \nabla^2 f(x^*)\, d + t^2 \epsilon(t) \\
&= \frac{t^2}{2} d^T \nabla^2 f(x^*)\, d + t^2 \epsilon(t) \\
&\geq t^2 \frac{\lambda_{\min}}{2} \|d\| + t^2 \epsilon(t) \\
&\geq \frac{t^2}{2} \left(\|d\| \lambda_{\min} + 2\epsilon(t) \right)
\end{aligned}
$$

Puisque $\lambda_{\min} > 0$, il s'ensuit que pour $t > 0$ suffisement petit

$$\|d\| \lambda_{\min} + 2\epsilon(t) > 0$$

ainsi

$$f(x^* + td) - f(x^*) > 0 \Rightarrow f(x^* + td) > f(x^*) .$$

2) Même argument. $\blacksquare$

Condition nécessaire et suffisante

Théorème (*Condition nécessaire et suffisante***)**

1) Soit $f : \mathbb{R}^n \to \mathbb{R}$ *une fonction convexe. Alors*

$$x^* \in \mathbb{R}^n \text{ est un minimum global pour } f \Leftrightarrow \nabla f(x^*) = 0.$$

2) Soit $f : \mathbb{R}^n \to \mathbb{R}$ *une fonction concave. Alors*

$$x^* \in \mathbb{R}^n \text{ est un maximum global pour } f \Leftrightarrow \nabla f(x^*) = 0.$$

Preuve.

1) Pour la première implication, elle est déjà fait précédemment. Pour la deuxième, on utilise la convexité de f, on a

$$f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle \text{ pour tout } x \in \mathbb{R}^n.$$

Donc,

$$f(x) \geq f(x^*) \text{ pour tout } x \in \mathbb{R}^n \Rightarrow x^* \in \arg\min_{\mathbb{R}^n} f(x) \ \ (x^* \text{ est un minimum global}).$$

2) Même procédure. ■

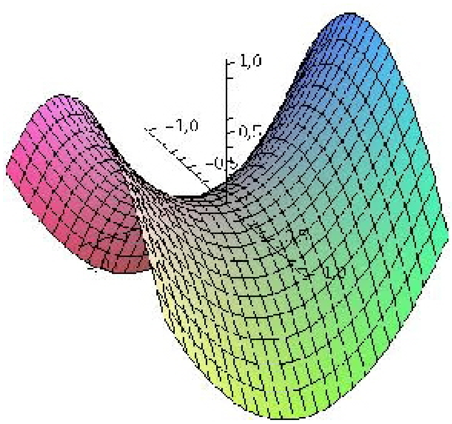
Point selle

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. La fonction f présente un *point selle* (*point col*) en $(\overline{x}, \overline{y}) \in \mathbb{R}^2$ s'il existe un voisinage V de $\overline{x}$ et W de $\overline{y}$ tel que

$$f(\overline{x}, y) \leq f(\overline{x}, \overline{y}) \leq f(x, \overline{y}) \text{ pour tout } x \in V \text{ et } y \in W$$

C'est à dire

$$f(\overline{x}, \overline{y}) = \max_{y \in W} f(\overline{x}, y) = \min_{x \in V} f(x, \overline{y})$$



Exemple. La fonction

$$f(x, y) = x^4 - y^4$$

présente un point selle en $(0, 0)$

$$\begin{aligned} f(0,0) &= \max_{y \in \mathbb{R}} f(\overline{x}, y) = \max_{y \in \mathbb{R}} \left(-y^4\right) \\ &= \min_{x \in \mathbb{R}} f(x, \overline{y}) = \min_{x \in \mathbb{R}} \left(x^4\right) \end{aligned}$$

Cas particulier

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. Soit $(\overline{x}, \overline{y}) \in \mathbb{R}^2$ tel que

$$\nabla f\left(\overline{x}, \overline{y}\right) = 0,$$

Alors:

1- Si $det\left(\nabla^2 f(\overline{x}, \overline{y})\right) > 0$, alors $(\overline{x}, \overline{y})$ est un point extremum :

-si $tr\left(\nabla^2 f\left(\overline{x}, \overline{y}\right)\right) > 0$ alors $(\overline{x}, \overline{y})$ est un minimum local,

-si $tr\left(\nabla^2 f\left(\overline{x}, \overline{y}\right)\right) < 0$ alors $(\overline{x}, \overline{y})$ est un maximum local,

2- Si $det(\nabla^2 f\left(\overline{x}, \overline{y}\right)) < 0$ alors $(\overline{x}, \overline{y})$ ni minimum ni maximum.

3- Si $det(\nabla^2 f\left(\overline{x}, \overline{y}\right)) = 0$, on ne peut rien conclure.