Chapitre 3 Optimisation différentiable sans contraintes

Dans ce chapitre, nous allons caractériser le minimum (maximum) d'une fonction numérique $f: U \to \mathbb{R}$ où $U \subset \mathbb{R}^n$ ouvert en ajoutant l'hypothèse que f est dérivable.

Ensuite, nous allons présenter des méthodes numériques pour calculer le ou les minimums de f.

3.1 Conditions d'optimalité

Soit $f: U \to \mathbb{R}$ une fonction numérique définie dans un ouvert $U \subset \mathbb{R}^n$. On cherche à caractériser un point x minimisant, i.e.

$$f(x) = \min_{y \in U} f(y)$$

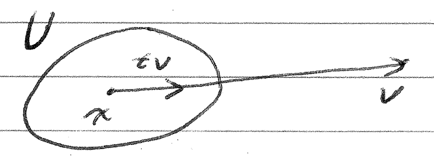
Théorème: condition nécessaire d'optimalité du 1er ordre

Soit $f: U \to \mathbb{R}$ de classe C^1. (local ou global)
Si $x \in U$ est un pt. minimisant, alors

$$\nabla f(x) = 0$$

Preuve:

On a que $f(x+tv) - f(x) \geq 0$ $\forall v \in \mathbb{R}^n$
$\forall t$ dans un intervalle proche de 0 $\Rightarrow$ $\forall t \in]0, \epsilon[$ $\epsilon > 0$
de sorte que $x + tv \in U$ $\forall t \in]0, \epsilon[$



$$\Rightarrow \quad \frac{f(x+tv) - f(x)}{t} \geq 0 \qquad \text{car } t>0$$

$$\lim_{\substack{t \to 0 \\ t>0}} \frac{f(x+tv) - f(x)}{t} \geq 0$$

$$\parallel$$

$$(\nabla f(x), v) \geq 0 \qquad \forall v \in \mathbb{R}^n$$

$$\Rightarrow \quad (\nabla f(x), -v) \geq 0 \qquad \forall v \in \mathbb{R}^n$$

$$\Rightarrow \quad (\nabla f(x), v) = 0 \qquad \forall v \in \mathbb{R}^n$$

$$\Rightarrow \quad \nabla f(x) = 0$$

Remarques:

1) le théorème ci-dessus est aussi valide pour les minimums locaux
$x \in U$ est un minimum local
s'il existe un voisinage V de x dans U
$$x \in V \subset U$$
t.q. $\quad f(y) \geq f(x) \quad \forall y \in V$

2) Un pt x tel que $\nabla f(x) = 0$
est dit un pt critique de f.

3) La condition $\nabla f(x) = 0$ n'est pas suffisante.
$f(x,y) = x^2 - y^2$
$f_x(0,0) = f_y(0,0) = 0 \quad \Rightarrow \quad \nabla f(0,0) = 0$
mais $(0,0)$ n'est pas un min. local de f.
C'est un pt-selle!

Par contre, si on ajoute le fait que f soit une fonction convexe, alors c'est une condition suffisante.

Théorème : condition nécessaire et suffisante d'optimalité dans le cas convexe.

Soit $f: U \to \mathbb{R}$ une fonction convexe de classe C^1 définie sur un ouvert U convexe de $\mathbb{R}^n$.
On a que :

$x \in U$ est un pt. minimisant de f

$$\Leftrightarrow \quad \nabla f(x) = 0$$

Preuve : $\Rightarrow$ par le théorème précédent

$\Leftarrow$ f est convexe $\Leftrightarrow$

$$f(y) \geq f(x) + \underbrace{(\nabla f(x), y-x)}_{=0} \quad \forall y \in U$$

$$\Rightarrow f(y) \geq f(x) \quad \forall y \in U$$

Exemple :

$$f(x) = \frac{1}{2}(Ax, x) - (b, x) \quad \text{avec } A > 0 \text{ et } b \in \mathbb{R}^n$$

On sait que f est convexe.
Or $\nabla f(x) = Ax - b$

Lo $\nabla f(x) = 0$ est équivalent à résoudre le système linéaire

$Ax = b$ de solution unique $x = A^{-1}b$

C'est le pt. minimisant de f.

Ajoutons maintenant l'hypothèse que f est de classe C^2.

Théorème: condition nécessaire d'optimalité du 2^e ordre

Soit $f: U \to \mathbb{R}$ de classe C^2.
Si $x \in U$ est un pt. minimisant (local ou global), alors

i) $\nabla f(x) = 0$

ii) $H_f(x) \geq 0$ où H_f est la matrice Hessienne de f.

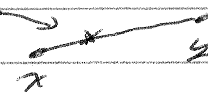
Preuve:

i) déjà fait. voir 1er ordre

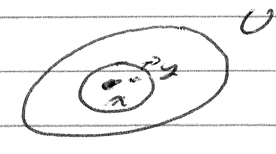
ii) On a, par Taylor,

$$f(y) = f(x) + (\nabla f(x), y-x) + \frac{1}{2}\,(H(y-x), y-x)$$

où $H = H_f(x + \xi(y-x))$



Ceci est valide $\forall y$ près de x



Mais $\nabla f(x) = 0$

$$\Rightarrow 0 \leq f(y) - f(x) = \frac{1}{2}\,(H(y-x), y-x)$$

Par continuité et $y \to x$

$$\Rightarrow (H_f(x)(y-x), y-x) \geq 0$$

On peut toujours choisir un y près de x t.q.

$$y - x = r\,v \quad \text{avec } v \in \mathbb{R}^n \text{ quelconque}$$

$$\Rightarrow H_f(x) \geq 0$$

Remarque :

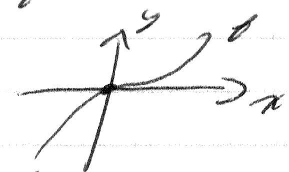
- la condition n'est pas suffisante en général.

Ex : $f(x) = x^3$ $\quad f : \mathbb{R} \to \mathbb{R}$

$f'(0) = 0$

$f''(0) = 0$

0 n'est pas un min. local



Voici une condition suffisante.

Théorème : condition suffisante d'optimalité du 2^e ordre.

Soit $f : U \to \mathbb{R}$ de classe C^2

a) Si $x \in U$ est un pt de U vérifiant

i) $\nabla f(x) = 0$,

ii) $H_f(x) > 0$,

alors x est un minimum local de f.

b) S'il existe un voisinage V d'un pt x dans U

$$x \in V \subset U$$

vérifiant

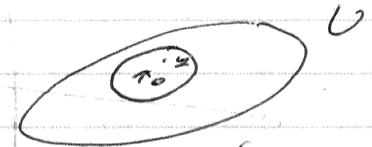
i) $\nabla f(x) = 0$,

ii) $H_f(y) \geq 0 \quad \forall y \in V$,

alors x est un minimum local dans U mais global dans V.

Preuve :

a) $H_f(x) > 0 \quad \Rightarrow \quad H_f(y) > 0 \quad \forall y$ près de x



$$\Rightarrow f(y) = f(x) + (\underbrace{\nabla f(x)}_{0}, y-x) + \frac{1}{2}(H(y-x), y-x)$$

mais $H(x + \xi(y-x)) > 0$

car $x + \xi(y-x)$ est près de x

$$\Rightarrow \quad f(y) = f(x) + \frac{1}{2}(H\, y-x, y-x) > f(x)$$

$\therefore$ x pt. minimisant (local)

b) Pour $y \in V$

$$f(y) = f(x) + (\underbrace{\nabla f(x)}_{0}, y-x) + \underbrace{\frac{1}{2}(H(y-x), y-x)}_{\geq 0}$$

$$f(y) \geq f(x)$$

Exemples :

1. $f(x,y) = x^2 + y^2 - 2xy \qquad f : \mathbb{R}^2 \to \mathbb{R}$

$$\nabla f = (2x - 2y, \; 2y - 2x) = 0 \Rightarrow x = y$$

$$H = H_f = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \geq 0 \qquad \forall (x,y) \in \mathbb{R}^2$$

car $\frac{1}{2}(Ha, a) = (a_1 - a_2)^2 \geq 0$

$\therefore$ Tous les pts (x,x) sont des minima (globaux) dans $\mathbb{R}^2$

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

$f_x = 3 - 3x^2 = 0$

$f_y = -4y + 4y^3 = 0$

On trouve $(\pm 1, \pm 1)$ et $(\pm 1, 0)$ pts critiques de f (6 pts)

$$H = \begin{pmatrix} -6x & 0 \\ 0 & 12y^2 - 4 \end{pmatrix}$$

$H > 0$ seulement aux pts $(-1, \pm 1)$
min. locaux