Descente de gradient optimal

Chen Thomas
t.chen.thomas1[at]gmail.com

17 mai 2024

Attention

1. Ce document contient certainement des coquilles. N'hésitez pas à me le signaler. De même si vous avez une question.
2. Pour les recasages, ce sont les miens mais ce développement se case peut-être ailleurs et je n'y ai pas réfléchi.
3. Il se peut que ce développement dure plus de 15 minutes. J'ai essayé de le découper pour faire des recollements personnalisés.

Leçons

- 219 : Extremums : existence, caractérisation, recherche. Exemples et applications.
- 229 : Fonctions monotones. Fonctions convexes. Exemples et applications.
- 253 : Utilisation de la notion de convexité en analyse.

Références

[1] S. Francinou, H. Gianella, S. Nicolas. *Exercices de mathématiques. Oraux X-ENS, Analyse 4*. Cassini, 2012.

Théorème 1. *[1] Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^1$. On suppose qu'il existe $\alpha \in \mathbb{R}^{+*}$ tel que*

$$\forall x, y \in \mathbb{R}^n, f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2.$$

Alors f admet un unique minimum. De plus, en construisant la suite $x_{k+1} = x_k + t_k d_k$ avec $d_k = \nabla f(x_k)$ et $t_k \in \operatorname{argmin} f(x_k + td_k)$, cette suite converge, et ce vers l'unique minimum de f.

Démonstration. On va montrer que

1. f admet un unique minimum par coercivité.
2. $\varphi : t \mapsto f(x + t\nabla f(x))$ admet un unique minimum.
3. La suite $(x_k)_k$ converge et ce, vers a.

1. Soit $x \in \mathbb{R}^n$. Alors
$$f(x) \geq f(0) + \langle \nabla f(0), x \rangle + \frac{\alpha}{2}\|x\|^2 \underset{\|x\| \to +\infty}{\longrightarrow} +\infty.$$
Ainsi, f est coercive. En particulier, il existe $A > 0$ tel que $\forall \|x\| > A, f(x) \geq f(0)$. Sur le compact $B_f(0, A)$, f est continue donc admet un minimum noté a. Ainsi, $f(0) \geq a$ et $\forall x \in \mathbb{R}^n, f(x) \geq a$.

   Soit b un minimum de f. Alors
$$f(b) - f(a) = 0 \geq \langle \nabla f(a), b - a \rangle + \frac{\alpha}{2}\|b - a\|^2$$
mais a est point critique donc $0 \geq \|b - a\|^2$: $b = a$.

2. Supposons que $\nabla f(x) \neq 0$. φ est coercive par composition de limite. Par continuité, elle admet un minimum. Il manque l'unicité. Ce minimum est point critique, disons t_0. On a donc $\varphi'(t_0) = 0$. Or,

$$\forall t \in \mathbb{R}, \varphi'(t) = \mathrm{d}f_{x+t\nabla f(x)}(\nabla f(x)) = \langle \nabla f(x + t\nabla f), \nabla f(x) \rangle.$$

En $t = t_0$, on a donc

$$\langle \nabla f(x + t_0 \nabla f), \nabla f(x) \rangle.$$

On a alors

$$\varphi(t) = f(x + t\nabla f(x)) \geq f(x + t_0 \nabla f(x)) + \langle \nabla f(x + t_0 \nabla f), (t - t_0)\nabla f(x) \rangle + \frac{\alpha}{2}(t - t_0)^2 \|\nabla f(x)\|^2$$

donc $\varphi(t) > \varphi(t_0)$ pour tout $t \neq t_0$. On a bien unicité du minimum.

Si $\nabla f(x) = 0$, alors $f(a) - f(x) \geq 0 + \frac{\alpha}{2}\|a - x\|^2$ donc $f(x) \leq f(a)$ et $x = a$.

3. Une telle suite est bien définie. Si un des x_k vaut a, alors la suite stationne à a par le point précédent. On suppose donc que ce n'est jamais le cas. Par construction, la suite $(f(x_k))_k$ décroît. En effet,

$$f(x_{k+1}) = f(x_k + t_k d_k) \leq f(x_k + 0 d_k)$$

par construction de t_k. Elle est de plus minorée car f est minorée par 0. Ainsi, cette suite converge, disons vers ℓ. Ainsi, $f(x_k) - f(x_{k+1}) \underset{k \to +\infty}{\longrightarrow} 0$ donc

$$f(x_k) - f(x_{k+1}) = \underbrace{\langle \nabla f(x_{k+1}), t_k \nabla f(x_k) \rangle}_{=0} + \frac{\alpha}{2}\|x_k - x_{k+1}\|^2 \underset{k \to +\infty}{\longrightarrow} 0.$$

Par coercivité de f, il existe $A \in \mathbb{R}$ tel que $\forall \|x\| \geq A, f(x) \geq f(0)$. Ainsi, comme la suite $(f(x_k))_k$ tend en décroissant vers 0, on en déduit qu'à partir d'un certain rang K, $\forall k \geq K, f(x_k) < f(0)$ donc en contraposant, $\forall k \geq K, x_k \in B_f(0, A)$ qui est compact car fermée bornée en dimension finie. Soit φ extraction telle que $(x_{\varphi(k)})_k$ converge, disons vers x. Alors par continuité du gradient (f étant $\mathcal{C}^1$), on a

$$\nabla f(x_{\varphi(k)}) \underset{k \to +\infty}{\longrightarrow} \nabla f(x)$$

mais

$$\forall k \geq K, 0 = \langle \nabla f(x_{\varphi(k)+1}), \nabla f(x_{\varphi(k)}) \rangle \underset{k \to +\infty}{\longrightarrow} \|\nabla f(x)\|^2$$

par continuité du produit scalaire et que $x_{\varphi(k)+1} \underset{k \to +\infty}{\longrightarrow} x$. Ainsi, $\nabla f(x) = 0$ donc $x = a$.

Ainsi, $(x_k)_k$ admet une unique valeur d'adhérence. Cette suite vivant dans un compact, le fait qu'elle ait une unique valeur d'adhérence montre qu'elle converge, et ce, vers a.

□