Chapitre 2: Premiers algorithmes

Ivan Lejeune[*]

26 février 2024

1 Chapitre 1 : Optimisation convexe

2 Chapitre 2 : Premiers algorithmes

On considère $U \subset \mathbb{R}^n$ ouvert et $f: U \to \mathbb{R}$.
On suppose dans la suite que l'on sait montrer l'existence du minimum
On cherche un processus de construction d'une suite de points $(x_k)_{k\in\mathbb{N}}$ telle que (x_k) converge vers le minimum de f.

Definition 2.0.1. Un algorithme **converge** si la suite de points converge vers le minimum recherché.

On suppose également qu'on connait des algorithmes qui convergent pour $n = 1$ (cas des fonctions $f: I \subset \mathbb{R} \to \mathbb{R}$).

2.1 Méthode de relaxation

La principe de la méthode de relaxation consiste à se ramener à la minimisation de fonctions à une variable, de la manière suivante :

(i) On fixe une suite de directions de descente $(d_k)_{k\in\mathbb{N}}$,

(ii) On fixe un point de départ u_1.

(iii) Connaissant u_k, on calcule $\rho_{k+1} \in \mathbb{R}$ qui réalise le minimum de la fonction

$$\phi_k: \rho \mapsto f(u_k + \rho \cdot d_{k+1})$$

(iv) On pose $u_{k+1} = u_k + \rho_{k+1} d_{k+1}$.

Théorème 2.1.1. Si f est $\mathscr{C}^1$ et elliptique, alors la méthode de relaxation converge.

Preuve IMPORTANTE A COMPRENDRE.

Nous avons vu au chapitre précédent que si f est elliptique, alors

$$\forall (x, y) \in \mathbb{R}^n, \underbrace{f(y) - f(x) - f'(x)(y - x) \geq \frac{\alpha}{2}\|y - x\|^2}_{(E)}$$

et que

$$\underline{x} \text{ est un minimum de } f \iff f'(\underline{x}) = 0.$$

Considérons la fonction

$$\begin{aligned} \phi_k: \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(u_k + t d_{k+1}) \end{aligned}$$

[*]Cours inspiré de M. Marche

qui atteint son minimum en ρ_{k+1}, donc

$$\phi_k'(\rho_{k+1}) = f'(x_{k+1}) \cdot d_{k+1} = 0$$

et puisque $x_{k+1} - x_k$ est parallèle à d_{k+1}, on a

$$f'(x_{k+1}) \cdot (x_{k+1} - x_k) = 0$$

En outre, puisque x_{k+1} minimise f sur la droite passant par x_k et dirigée par d_{k+1}, on a

$$f(x_{k+1}) \leq f(x_k)$$

Ainsi, la suite des $f(x_k)$ est décroissante, et minorée, donc convergente.
Considérons maintenant l'inégalité (E) avec $x = x_{k+1}$ et $y = x_k$.

$$f(x_k) - f(x_{k+1}) - f'(x_{k+1})(x_k - x_{k+1}) \geq \frac{\alpha}{2}||x_k - x_{k+1}||^2$$

Comme $f'(x_{k+1}) \cdot (x_k - x_{k+1}) = 0$, on a

$$f(x_k) - f(x_{k+1}) \geq \frac{\alpha}{2}||x_k - x_{k+1}||^2$$

Puisque la suite $(f(x_k))$ est convergente, on a

$$\lim_{k\to\infty} ||x_k - x_{k+1}|| = 0$$

On a également, pour tout $p \in \{1, \ldots, n\}$,

$$\lim_{k\to\infty} x_{k+p} - x_k = 0$$

Par ailleurs, on a par ellipticité de f :

$$\forall (x, y) \in \mathbb{R}^n, (f'(y) - f'(x))(y - x) \geq \alpha||y - x||^2$$

Soit $\underline{x}$ le minimum de f. Puisque $f'(\underline{x}) = 0$, on a, en appliquant l'inégalité précédente à $y = x_k$ et $x = \underline{x}$:

$$\alpha||x_k - \underline{x}||^2 \leq (f'(x_k) \cdot (x_k - \underline{x}))$$

d'où

$$\alpha||x_k - \underline{x}||^2 \leq ||f'(x_k)|| \cdot ||x_k - \underline{x}||$$

et donc

$$||x_k - \underline{x}|| \leq \frac{||f'(x_k)||}{\alpha}$$

Il reste donc à montrer que $f'(x_k)$ tend vers 0.
Il suffit de montrer que :

$$\forall p \in \{1, \ldots, n\}, \lim_{k\to\infty} f'(x_k) \cdot e_p = 0$$

Soit donc $p \in \{1, \ldots, n\}$, on a pour $j \in \mathbb{N}$:

$$f'(x_{p+j_n}) \cdot d_{p+j_n} = 0$$

or $d_{p+j_n} = e_p$ donc

$$f'(x_{p+j_n}) \cdot e_p = 0$$

On effectue la division euclidienne de k par n :

$$k = \alpha n + i \text{ avec } i \in \{1, \ldots, n\}$$

On a alors

$$\begin{aligned} f'(x_k)\cdot e_p &= f'(x_{\alpha n+i})\cdot e_p \\ &= (f'(x_{\alpha n+1}) - f'(x_{\alpha n+p}))\cdot e_p \end{aligned}$$

puisque $f'(x_{\alpha n+p})\cdot e_p = 0$.
Ainsi,

$$|f'(x_k)\cdot e_p| \leq \|f'(x_{\alpha n+i}) - f'(x_{\alpha n+p})\|\cdot\|e_p\|$$

Or, lorsque k tend vers l'infini, α tend vers l'infini, et i est borné par définition.
Ainsi,

$$\lim_{\alpha\to\infty} x_{\alpha n+i} - x_{\alpha n} = 0$$

et

$$\lim_{\alpha\to\infty} x_{\alpha n} - x_{\alpha n+p} = 0$$

Ainsi, pour $\alpha \to \infty$, on a

$$\|x_{\alpha n+i} - x_{\alpha n+p}\| \to 0$$

Enfin, par uniforme continuité de f' sur une boule $\overline{B}(0,M)$, on a

$$\|f'(x_{\alpha n+i}) - f'(x_{\alpha n+p})\| \to 0$$

et donc finalement

$$\lim_{k\to\infty} f'(x_k)\cdot e_p = 0$$

> **Remarque 2.1.2 sur le choix de M.** On veut qu'à partir d'un certain rang, tous les termes de la suite soient dans la boule $\overline{B}(0,M)$.
> Or, la suite $(f(x_k))$ est bornée et f est coercive, ce qui implique que la suite (x_k) est nécessairement bornée.

2.2 Méthode de gradient à pas optimal

On considère $f\colon\mathbb{R}^n \to \mathbb{R}$.
L'idée de la méthode est de choisir des directions de descente (... vers le minimum) privilégiées.
Connaissant x_k, on considère la fonction

$$\begin{aligned} \phi_k\colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f(x_k - t\underbrace{\nabla f(x_k)}_{d_{k+1}}) \end{aligned}$$

et t_{k+1} qui minimise ϕ_k sur $\mathbb{R}$.
On pose alors $x_{k+1} = x_k - t_{k+1}\nabla f(x_k)$.
Pourquoi un tel choix de descente selon la direction $-\nabla f(x_k)$?
Parce que c'est localement le meilleur choix possible :

$$f(x+t\xi) = f(x) + t\nabla f(x)\cdot\xi + o(t)$$

Ainsi, on observe qu'en négligeant le reste, on 'descend' le plus possible en rendant le terme $t\nabla f(x)\cdot\xi$ le plus négatif possible. Il faut donc choisir ξ parallèle à $\nabla f(x)$ et de direction opposée.

Théorème 2.2.1. Si f est $\mathscr{C}^1$ et elliptique, alors la méthode de gradient à pas optimal converge.

Preuve. La preuve est similaire à celle de la méthode de relaxation, en utilisant la propriété de coercivité de f :
La suite des $f(x_k)$ est décroissante et minorée, donc convergente.
Ainsi, la suite $(f(x_k))$ est convergente, et la suite (x_k) est bornée.

Par ailleurs, on a $\nabla f(x_{k+1}) \perp \nabla f(x_k)$ car

$$\begin{aligned}\phi_k'(t_{k+1}) = 0 &= f'(x_{k+1}\nabla f(x_k)) \\ &= (\nabla f(x_{k+1}), \nabla f(x_k))\end{aligned}$$

On a également $(\nabla f(x_{k+1}), x_{k+1} - x_k) = 0$
En outre, on a

$$f(x_k) - f(x_{k+1}) \geq \frac{\alpha}{2}||x_k - x_{k+1}||^2$$

d'où $||x_k - x_{k+1}|| \to 0$.
On peut de manière similaire à la preuve précédente, montrer que $(\nabla f(x_k))$ tend vers 0.
Pour conclure, pour $x = \underline{x}$ et $y = x_k$, on a

$$(\nabla f(\underline{x}) - \nabla f(x_k)), (x_k - \underline{x}) \geq \alpha||x_k - \underline{x}||^2$$

d'où

$$||\underline{x} - x_k|| \leq \frac{||\nabla f(\underline{x}) - \nabla f(x_k)||}{\alpha}$$

Comme $\nabla f(\underline{x}) = 0$, on a

$$||\underline{x} - x_k|| \leq \frac{||\nabla f(x_k)||}{\alpha} \to 0$$

et donc $x_k \to \underline{x}$.

2.3 Méthode de gradient à pas fixe ou variable

Pour ne pas avoir à résoudre le problème d'optimisation 1d à chaque itération, on peut fixer à priori une suite de réels $(\rho_k)_{k\in\mathbb{N}}$ et poser

$$x_{k+1} = x_k - \rho_{k+1}\nabla f(x_k)$$

Il est même possible de choisir $\rho_k = \rho$ constant.

Definition 2.3.1. Soit $f: \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$. On dit que ∇f est **Lipschitzienne** de constante $M > 0$ si

$$\forall (x, y) \in \mathbb{R}^n, ||\nabla f(x) - \nabla f(y)|| \leq M||x - y||$$

Théorème 2.3.2. On suppose que f est $\mathscr{C}^1$ elliptique (de constante α) et que ∇f est Lipschitzienne de constante M.
Si il existe deux constantes $a, b \in \mathbb{R}$ telles que

$$\forall K \in \mathbb{N}, 0 < a \leq \rho_k \leq b < \frac{2\alpha}{M^2}$$

alors la méthode de gradient à pas fixe converge et il existe $\beta < 1$ tel que tel que

$$\forall K \in \mathbb{N}, ||x_k - \underline{x}|| \leq \beta^k||x_0 - \underline{x}||$$