

## 1. Algorithme du gradient

**Théorème 1.1.** — Soit  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$  une application  $\alpha$ -convexe et de classe  $\mathcal{C}^1$ . Alors  $\Phi$  admet un minimum unique  $x^* \in \mathbb{R}^n$ , approché par la suite  $(x_k)_{k \in \mathbb{N}}$  définie comme suit.

- Le point  $x_0 \in \mathbb{R}^n$  est arbitraire.
- Pour  $k \geq 0$ , si  $x_k = x^*$  alors  $x_{k+1} = x^*$ , sinon  $x_{k+1} = x_k + \lambda_k \nabla \Phi(x_k)$  où  $\lambda_k \in \mathbb{R}$  est l'unique réel minimisant la fonction  $\lambda \mapsto \Phi(x_k + \lambda \nabla \Phi(x_k))$ .

*Démonstration.* — • Montrons d'abord l'assertion d'existence et d'unicité du minimum en montrant, pour un certain  $\alpha_0 > 0$ , que pour tous  $x, y \in \mathbb{R}^n$ ,  $\Phi(y) \geq \Phi(x) + \langle \nabla \Phi(x) | y - x \rangle + \alpha_0 \|x - y\|^2$ . Soient  $x, y \in \mathbb{R}^n$ . D'après la formule de Newton-Leibniz,

$$\begin{aligned} \Phi(y) - \Phi(x) &= \int_0^1 \langle \nabla \Phi(x + s(y-x)) | y - x \rangle ds \\ &= \langle \nabla \Phi(x) | y - x \rangle + \int_0^1 \langle \nabla \Phi(x + s(y-x)) - \nabla \Phi(x) | s(y-x) \rangle s^{-1} ds \\ &\geq \langle \nabla \Phi(x) | y - x \rangle + \alpha \|x - y\|^2 \int_0^1 s ds. \end{aligned}$$

D'où l'inégalité en prenant  $\alpha_0 = \alpha/2$ . En particulier, la fonction  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$  est coercive et admet donc un minimum global  $x^* \in \mathbb{R}^n$  par un argument de compacité. Mais pour  $x \neq x^*$ , comme on a  $\nabla \Phi(x^*) = 0$  par la condition d'Euler, on obtient  $\Phi(x) \geq \Phi(x^*) + \alpha_0 \|x - x^*\|^2$  ce qui prouve l'unicité. Notons aussi qu'il n'y a pas d'autres minimums locaux car  $\Phi(x_0) \geq \Phi(x^*) \geq \Phi(x_0) + \alpha_0 \|x_0 - x^*\|^2$  en un tel minimum local  $x_0 \in \mathbb{R}^n$ , d'où  $\Phi(x_0) = \Phi(x^*)$  et  $x_0 = x^*$ .

• Montrons à présent la bonne définition de la suite mentionnée. Supposons que  $x_k \neq x^*$ , ou encore que  $\nabla \Phi(x_k) \neq 0$ . La fonction  $\theta_k : \mathbb{R} \rightarrow \mathbb{R}, \lambda \mapsto \Phi(x_k + \lambda \nabla \Phi(x_k))$  est de classe  $\mathcal{C}^1$  par composition et sa dérivée est donnée par  $\theta'_k(\lambda) = \langle \nabla \Phi(x_k + \lambda \nabla \Phi(x_k)) | \nabla \Phi(x_k) \rangle$ . Ainsi, pour  $\lambda, \mu \in \mathbb{R}$ ,

$$\begin{aligned} (\theta'_k(\lambda) - \theta'_k(\mu)) \cdot (\lambda - \mu) &= \langle \nabla \Phi(x_k + \lambda \nabla \Phi(x_k)) - \nabla \Phi(x_k + \mu \nabla \Phi(x_k)) | (\lambda - \mu) \nabla \Phi(x_k) \rangle \\ &\geq \alpha \|\nabla \Phi(x_k)\|^2 |\lambda - \mu|^2 \end{aligned}$$

Donc  $\theta_k$  est  $\alpha \|\nabla \Phi(x_k)\|^2$ -convexe et d'après ce qui précède, elle est minimisée de façon unique par un réel  $\lambda_k \in \mathbb{R}$  caractérisée par la propriété  $\theta'_k(\lambda_k) = 0$ .

• Montrons enfin la convergence de l'algorithme. Notons qu'en vertu de l'identité  $\theta'_k(\lambda_k) = 0$  et de l'expression de  $\theta'_k$ , on a la relation d'orthogonalité  $\nabla \Phi(x_k) \perp \nabla \Phi(x_{k+1})$ . Tout d'abord, la suite  $(x_k)_{k \geq 0}$  est contenue dans le compact  $K = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ , il suffit donc de montrer que  $(x_k)_{k \geq 0}$  admet  $x^*$  pour seule valeur d'adhérence. Mais la suite  $(\Phi(x_k))$  est décroissante et minorée par  $\Phi(x^*)$  donc converge. En particulier, on a,

$$\Phi(x_k) - \Phi(x_{k+1}) \xrightarrow[k \rightarrow \infty]{} 0.$$

D'après l'inégalité prouvée précédemment, on a aussi,

$$\begin{aligned} \Phi(x_k) - \Phi(x_{k+1}) &\geq \langle \nabla \Phi(x_{k+1}) | x_k - x_{k+1} \rangle + \alpha_0 \|x_k - x_{k+1}\|^2 \\ &= \lambda_k \langle \nabla \Phi(x_{k+1}) | \nabla \Phi(x_k) \rangle + \alpha_0 \|x_k - x_{k+1}\|^2 \\ &= \alpha_0 \|x_k - x_{k+1}\|^2. \end{aligned}$$

Donc la suite  $(x_k - x_{k+1})_{k \geq 0}$  converge aussi vers 0. La fonction  $\nabla \Phi$  étant uniformément continue sur le compact  $K$ , d'après le théorème de Heine, on a encore,

$$\nabla \Phi(x_k) - \nabla \Phi(x_{k+1}) \xrightarrow[k \rightarrow \infty]{} 0.$$

Puis, en vertu du théorème de Pythagore, on a,

$$\|\nabla \Phi(x_k)\|^2 \leq \|\nabla \Phi(x_k)\|^2 + \|\nabla \Phi(x_{k+1})\|^2 = \|\nabla \Phi(x_k) - \nabla \Phi(x_{k+1})\|^2.$$

Finalement,

$$\nabla \Phi(x_k) \xrightarrow[k \rightarrow \infty]{} 0.$$

On considère enfin  $(\xi_k)_{k \geq 0}$  une sous-suite de  $(x_k)_{k \geq 0}$  admettant un élément  $\xi \in K$  pour limite. Par continuité de  $x \mapsto \nabla \Phi(x)$ , la fonction  $\Phi$  étant supposée de classe  $\mathcal{C}^1$ , on a la convergence de  $\nabla \Phi(\xi_k)$  vers  $\nabla \Phi(\xi)$ . Mais nécessairement, on a  $\nabla \Phi(\xi) = 0$  donc  $\xi = x^*$  et la suite  $(x_k)$  admet donc  $x^*$  pour seule valeur d'adhérence, d'où la convergence voulue.  $\square$

---