Titre : **Algorithme du gradient à pas optimal**
Recasages : 162,219,223,226,229,233,253
Thème : Calcul différentiel, analyse convexe, espaces vectoriels normés en dimension finie
Références : Ciarlet (189,190)

Rappelons que l'algorithme du gradient à pas optimal est défini comme suit :

$$\begin{cases} u^0 \in \mathbb{R}^n \text{ fixé} \\ \text{Pour } k \geqslant 0 \text{ faire} \\ \quad \text{Trouver } \rho_k \text{ tel que } J(u^k - \rho_k \nabla J(u^k)) = \inf_{\rho \in \mathbb{R}} J(u^k - \rho \nabla J(u^k)) \\ \quad u^{k+1} := u^k - \rho_k \nabla J(u^k) \\ \text{fin} \end{cases}$$

Théorème 1. *Soit $J : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle différentiable et α-convexe. Si ∇J est une fonction lipschitzienne sur $\mathbb{R}^n$, alors l'algorithme du gradient à pas optimal converge vers $u \in \mathbb{R}^n$ réalisant $J(u) = \inf_{x \in \mathbb{R}^n} J(x)$.*

Commençons par noter que l'α-convexité de J assure l'existence d'un unique minimum global u, caractérisé par l'équation d'Euler $\nabla J(u) = 0$, le problème de minimisation est donc bien posé. De plus, on peut supposer que pour tout $k \geqslant 0$, $\nabla J(u^k) \neq 0$ (sinon $u^k = u$ et l'algorithme a convergé en un nombre fini d'itérations).
Montrons que le problème de minimisation intermédiaire est bien posé : Soit $k \geqslant 0$, on pose

$$\begin{array}{rccl} \alpha_k : & \mathbb{R} & \longrightarrow & \mathbb{R}^n \\ & \rho & \longmapsto & u^k - \rho \nabla J(u^k) \end{array}$$

Et $\varphi_k := J \circ \alpha_k : \mathbb{R} \to \mathbb{R}$, cette fonction est différentiable, et on a par dérivation des composées

$$\varphi_k'(\rho) = dJ_{\alpha_k(\rho)} \circ d\alpha_{k\ \rho} = \left\langle \nabla J(u^k - \rho \nabla J(u^k)), \alpha_k'(\rho) \right\rangle = -\left\langle \nabla J(u^k - \rho \nabla J(u^k)), \nabla J(u^k) \right\rangle$$

Soient $\rho, \sigma \in \mathbb{R}$, on a

$$\begin{aligned} (\varphi_k'(\rho) - \varphi_k'(\sigma))(\rho - \sigma) &= \left\langle \nabla J(u^k - \sigma \nabla J(u^k)) - \nabla J(u^k - \rho \nabla J(u^k)), \nabla J(u^k) \right\rangle (\rho - \sigma) \\ &= \left\langle \nabla J(\alpha_k(\sigma)) - \nabla J(\alpha_k(\rho)), (\rho - \sigma) \nabla J(u^k) \right\rangle \\ &= \left\langle \nabla J(\alpha_k(\sigma)) - \nabla J(\alpha_k(\rho)), (u^k - \sigma \nabla J(u^k)) - (u^k - \rho \nabla J(u^k)) \right\rangle \\ &= \langle \nabla J(\alpha_k(\sigma)) - \nabla J(\alpha_k(\rho)), \alpha_k(\sigma) - \alpha_k(\rho) \rangle \\ &\geqslant \alpha \left\| \alpha_k(\sigma) - \alpha_k(\rho) \right\|^2 \\ &= \alpha |\rho - \sigma|^2 \left\| \nabla J(u^k) \right\|^2 \end{aligned}$$

Donc φ_k est $\alpha \left\| \nabla J(u^k) \right\|^2$-convexe et la problème de minimisation intermédiaire admet une unique solution ρ_k, caractérisée par l'équation d'Euler :

$$\left\langle \nabla J(u^k - \rho_k \nabla J(u^k)), \nabla J(u^k) \right\rangle = \left\langle \nabla J(u^{k+1}), \nabla J(u^k)) \right\rangle = 0$$

En particulier, deux directions de descente successives $\nabla J(u^{k+1})$ et $\nabla J(u^k)$ sont orthogonales, on a de plus

$$\left\langle \nabla J(u^{k+1}), u^k - u^{k+1} \right\rangle = \left\langle \nabla J(u^{k+1}), u^k - (u^k - \rho_k \nabla J(u^k) \right\rangle = 0$$

Nous pouvons à présent appliquer ces différents résultats pour montrer la convergence de notre algorithme : Par α-convexité de J, on a

$$J(u^k) \geqslant J(u^{k+1}) + \left\langle \nabla J(u^{k+1}), u^k - u^{k+1} \right\rangle + \frac{\alpha}{2} \left\| u^k - u^{k+1} \right\|^2 = J(u^{k+1}) + \frac{\alpha}{2} \left\| u^k - u^{k+1} \right\|^2$$

La suite $(J(u^k))_{k\in\mathbb{N}}$ est donc décroissante, et minorée par $\inf_{x\in\mathbb{R}^n} J(x)$ par définition : il s'agit d'une suite convergente. Donc la suite $(J(u^k) - J(u^{k+1}))_{k\in\mathbb{N}}$ converge vers 0, il en va donc de même de $(u^k - u^{k+1})_{k\in\mathbb{N}}$ toujours par l'inégalité ci-dessus.
Nous voulons enfin montrer que la suite $(u^k)_{k\in\mathbb{N}}$ converge vers u, on a, par α-convexité

$$\begin{aligned}
\alpha \left\| u^k - u \right\|^2 &\leqslant \left\langle \nabla J(u^k) - \nabla J(u), u^k - u \right\rangle \\
&= \left\langle \nabla J(u^k), u^k - u \right\rangle \\
&\leqslant \left\| \nabla J(u^k) \right\| \left\| u^k - u \right\|
\end{aligned}$$

Donc $\alpha \left\| u^k - u \right\| \leqslant \left\| \nabla J(u^k) \right\|$, il suffit donc de montrer que $(\nabla J(u^k))_{k\in\mathbb{N}}$ converge vers 0. Or, on a

$$\begin{aligned}
\left\| \nabla J(u^k) \right\|^2 &= \left\langle \nabla J(u^k), \nabla J(u^k) \right\rangle \\
&= \left\langle \nabla J(u^k) - \nabla J(u^{k+1}), \nabla J(u^k) \right\rangle \\
&\leqslant \left\| J(u^k) - J(u^{k+1}) \right\| \left\| \nabla J(u^k) \right\| \\
&\leqslant L \left\| u^k - u^{k+1} \right\| \left\| \nabla J(u^k) \right\|
\end{aligned}$$

Où L est le rapport de Lipschitz de ∇J sur $\mathbb{R}^n$, on en déduit $\left\| \nabla J(u^k) \right\| \leqslant L \left\| u^k - u^{k+1} \right\|$, ce qui termine la preuve.