# METHODE DU GRADIENT À PAS OPTIMAL

Leçons 219, 229, 233

Références Ciarlet p. 188-191

Théorème

Une fonctionnelle est dite elliptique si elle est $\mathcal{C}^1$ et s'il existe une constante $\alpha$ telle que

$$(\nabla J(v) - \nabla J(u), v-u) \geqslant \alpha \|v-u\|^2$$

On définit la méthode de gradient optimal par la donnée de $u_0$ et la construction

$$J(u_k - \rho(u_k)\nabla J(u_k)) = \inf_{\rho \in \mathbb{R}} J(u_k - \rho \nabla J(u_k))$$

$$u_{k+1} = u_k - \rho(u_k)\nabla J(u_k).$$

Pour une fonctionnelle elliptique définie sur $\mathbb{R}^n$, cette méthode converge.

Résumé

I - $J$ est strictement convexe et coercive ;

$$J(v) - J(u) \geqslant (\nabla J(u), v-u) + \frac{\alpha}{2}\|v-u\|^2.$$

II - $\|u_k - u\| \leq \frac{1}{\alpha}\|\nabla J(u_k)\| \xrightarrow{k \to +\infty} 0$

III - Cas d'une fonctionnelle quadratique elliptique

$$J(v) = \frac{1}{2}(Av, v) - (b, v)$$

I] • J est $\mathcal{C}^1$, on peut donc écrire Taylor reste intégral :

$$J(v) - J(u) = \int_0^1 (\nabla J(u+t(v-u)), v-u)\, dt$$

$$= (\nabla J(u), v-u) + \int_0^1 (\nabla J(u+t(v-u)) - \nabla J(u), v-u)\, dt$$

$$\geq (\nabla J(u), v-u) + \int_0^1 \alpha t \|v-u\|^2 dt$$

$$\geq (\nabla J(u), v-u) + \frac{\alpha}{2}\|v-u\|^2$$

Donc J est strictement convexe et coercive car :

$$J(v) \geq J(0) + (\nabla J(0), v) + \frac{\alpha}{2}\|v\|^2$$

II] • Comme J est strictement convexe, on a un minimum en un point $u$ tel que $\nabla J(u) = 0$.

Et on peut écrire :

$$\alpha \|u_k - u\|^2 \leq (\nabla J(u_k) - \nabla J(u), u_k - u)$$

$$= (\nabla J(u_k), u_k - u)$$

$$\leq \|\nabla J(u_k)\| \cdot \|u_k - u\|.$$

(*) On a donc $\|u_k - u\| \leq \frac{1}{\alpha}\|\nabla J(u_k)\|$. (fondamental !)

→ il suffit de montrer que le terme de droite tend vers 0 pour avoir la convergence.

• Montrons que les directions de descentes successives sont orthogonales.

Considérons la fonction

$$\varphi_k : \rho \longmapsto J(u_k - \rho \nabla J(u_k))$$

On a

$$\varphi_k(\rho + h) = J(u_k - \rho\nabla J(u_k)) - (\nabla J(u_k - \rho \nabla J(u_k)), h \nabla J(u_k))$$

Donc $\varphi_k'(\rho) = -(\nabla J(u_{k+1}), \nabla J(u_k))$

Or $\rho(u_k)$ est défini comme le minimum de $\varphi_k^*$, qui est convexe et coercive (car J l'est !), on a donc

$$(\nabla J(u_{k+1}), \nabla J(u_k)) = -\varphi_k'(\rho) = 0.$$

Et même, par définition de $u_{k+1} = u_k - \rho(u_k)\nabla J(u_k)$,

$$(\nabla J(u_{k+1}), u_{k+1} - u_k) = 0.$$

Donc l'inégalité de I donne :

$$J(u_{k+1}) - J(u_k) \geqslant \frac{\alpha}{2} \|u_{k+1} - u_k\|^2.$$

- À partir de là, on va supposer $\nabla J(u_k) \neq 0$ pour $k \geqslant 0$, car si ce gradient est nul c'est qu'on a trouvé exactement $u$ en un nombre fini d'étapes.

→ on peut donc dire que la suite $(J(u_k))_k$ est
  - * décroissante
  - * minorée

Donc $J(u_{k+1}) - J(u_k) \xrightarrow{k \to +\infty} 0$

d'où $\|u_{k+1} - u_k\| \xrightarrow{k \to +\infty} 0$

- On se rapproche de $\|\nabla J(u_k)\|$, regardons :

$$\|\nabla J(u_k)\|^2 = (\nabla J(u_k), \nabla J(u_k) - \nabla J(u_{k+1}))$$
$$\leqslant \|\nabla J(u_k)\| \, \|\nabla J(u_k) - \nabla J(u_{k+1})\|$$

donc $\|\nabla J(u_k)\| \leqslant \|\nabla J(u_k) - \nabla J(u_{k+1})\|$

- Mais remarquons que comme $(J(u_k))_k$ est décroissante, ~~elle est bornée~~ la suite $(u_k)_k$ est bornée car J est coercive.

Dès lors la dérivée de J n'a besoin d'être considérée que sur un compact, sur lequel elle est uniformément continue car continue !

Or $\|u_k - u_{k+1}\| \xrightarrow{k\to+\infty} 0$ donc $\|\nabla J(u_k) - \nabla J(u_{k+1})\| \xrightarrow{k\to+\infty} 0$

Mais c'est un majorant de $\|\nabla J(u_k)\|$, et on a donc bien la convergence.

III | Une petite remarque pratique d'application aux systèmes linéaires :

soit $J(v) = \frac{1}{2}(Av, v) - (b, v)$

on a $\nabla J(v) = Av - b$

donc trouver le minimum de $J$ résout notre système $Ax = B$.

~~Mais on a~~ alors

$$0 = (\nabla J(u_k), \nabla J(u_{k+1})) = (Au_k - b, A(u_k - \rho(u_k)(Au_k - b)) - b)$$
$$= (Au_k - b, Au_k - b) + (Au_k - b, -\rho(u_k)A(Au_k - b))$$
$$= (Au_k - b, w_k - \rho(u_k)Aw_k)$$
$$= (w_k, w_k - \rho(u_k)Aw_k)$$

Mais on a alors, en posant $w_k = \nabla J(u_k) = Au_k - b$ :

$$0 = (\nabla J(u_{k+1}), \nabla J(u_k))$$
$$= (w_k - \rho(u_k)Aw_k, w_k)$$
$$= (w_k, w_k) - \rho(u_k)(Aw_k, w_k)$$

On peut donc donner un algorithme utilisable en pratique :

$$\begin{cases} w_k = Au_k - b \\ \rho(u_k) = \dfrac{\|w_k\|^2}{(Aw_k, w_k)} \\ u_{k+1} = u_k - \rho(u_k)\, w_k \end{cases}$$