

Méthode de gradient à pas optimal

[BERNIS, p 159]

ÉNONCÉ :

Théorème : L'application

$$\begin{aligned} \Phi : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{2} \|x\|_A^2 - txb \end{aligned}$$

admet un minimum en $\bar{x}$, unique solution de l'équation $Ax = b$, $A \in \mathcal{S}_n^{++}(\mathbb{R})$.

De plus, pour $a \in \mathbb{R}^n$ distinct de $\bar{x}$ et $(x_k)_{k \in \mathbb{N}}$ la suite de $\mathbb{R}^n$ définie récursivement par :

$$\begin{cases} x_0 = a \\ \alpha_k = \frac{\|\nabla \Phi(x_k)\|^2}{\|\nabla \Phi(x_k)\|_A^2}, \text{ si } x_k \neq \bar{x}, 0 \text{ sinon} \\ x_{k+1} = x_k - \alpha_k \nabla \Phi(x_k) \end{cases}$$

converge vers $\bar{x}$ à une vitesse géométrique.

DÉVELOPPEMENT :

LEMME : (Inégalité de KANTOROVITCH :) Pour $x \in \mathbb{R}^n \setminus \{0\}$, on a :

$$\frac{\|x\|^4}{\|x\|_A^2 \|x\|_{A^{-1}}^2} \geq 4 \frac{\lambda_{\max} \lambda_{\min}}{(\lambda_{\max} + \lambda_{\min})^2}$$

Démonstration. Soit $x \in \mathbb{R}^n \setminus \{0\}$ fixé.
En vertu du théorème spectral, la matrice A admet une base or-

thonormée $(e_1, \dots, e_n)$ de vecteurs propres. Pour $i = 1, \dots, n$, on notera λ_i la valeur propre de e_i pour la matrice A . De plus, si λ est valeur propre de A alors λ^{-1} est valeur propre de A^{-1} . Ainsi, il vient :

$$\begin{aligned} \|x\|_A \|x\|_{A^{-1}} &= \left(\sum_{i=1}^n \lambda_i x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \lambda_i^{-1} x_i^2 \right)^{\frac{1}{2}} \\ &= \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_{\max}} x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \frac{\lambda_{\min}}{\lambda_i} x_i^2 \right)^{\frac{1}{2}} \\ &\leq \frac{1}{2} \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_{\max}} + \frac{\lambda_{\min}}{\lambda_i} \right) x_i^2 \right) \end{aligned}$$

Or l'application : $g : \begin{matrix} [\lambda_{\min}, \lambda_{\max}] \\ t \end{matrix} \longrightarrow \mathbb{R} \longmapsto \frac{t}{\lambda_{\max}} + \frac{\lambda_{\min}}{t}$ est convexe. De plus, $g(\lambda_{\max}) = g(\lambda_{\min}) = 1 + \frac{\lambda_{\min}}{\lambda_{\max}}$. D'où, pour tout $t \in [\lambda_{\min}, \lambda_{\max}]$, on a :

$$g(t) \leq 1 + \frac{\lambda_{\min}}{\lambda_{\max}}$$

Ainsi, on obtient l'inégalité :

$$\|x\|_A \|x\|_{A^{-1}} \leq \frac{1}{2} \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{\frac{1}{2}} \left(1 + \frac{\lambda_{\min}}{\lambda_{\max}} \right) \|x\|^2$$

On obtient ainsi l'inégalité souhaitée en élevant au carré. □

Démonstration. (théorème) : Notons, pour $k \in \mathbb{N}$, $g_k = \nabla \Phi(x_k)$.

Soit $p \in \mathbb{N}$. Alors on a :

$$\begin{aligned}
\|x_{p+1} - \bar{x}\|_A^2 &= \langle A(x_{p+1} - \bar{x}), x_{p+1} - x_p \rangle + \langle A(x_{p+1} - \bar{x}), x_p - \bar{x} \rangle \\
&= -\alpha_p \underbrace{\langle g_{p+1}, g_p \rangle}_{=0} + \langle A(x_{p+1} - \bar{x}), x_p - \bar{x} \rangle \\
&= \langle A(x_{p+1} - x_p), x_p - \bar{x} \rangle + \langle A(x_p - \bar{x}), x_p - \bar{x} \rangle \\
&= \langle x_{p+1} - x_p, A(x_p - \bar{x}) \rangle + \|x_p - \bar{x}\|_A^2 \\
&= -\alpha_p \langle g_p, g_p \rangle + \|g_p\|_{A^{-1}}^2 \\
&= -\frac{\|g_p\|^4}{\|g_p\|_A^2} + \|g_p\|_{A^{-1}}^2 \\
&= \left(1 - \frac{\|g_p\|^4}{\|g_p\|_A^2 \|g_p\|_{A^{-1}}^2}\right) \|x_p - \bar{x}\|_A^2 \\
&\leq \left(1 - 4 \frac{\lambda_{\max} \lambda_{\min}}{(\lambda_{\max} + \lambda_{\min})^2}\right) \|x_p - \bar{x}\|_A^2 \quad (\text{lemme}) \\
&= \left(\frac{(\lambda_{\max} - \lambda_{\min})^2}{(\lambda_{\max} + \lambda_{\min})^2}\right) \|x_p - \bar{x}\|_A^2
\end{aligned}$$

Comme $\sqrt{\lambda_{\min}} \|\cdot\| \leq \|\cdot\|_A \leq \sqrt{\lambda_{\max}} \|\cdot\|$ et par ce qui précède, on obtient par une récurrence évidente l'inégalité :

$$\|x_k - \bar{x}\|_A \leq \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^k \|x_0 - \bar{x}\|_A$$

et finalement :

$$\|x_k - \bar{x}\| \leq \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}\right)^k \|x_0 - \bar{x}\|$$

□

Remarques :

- Φ admet un minimum puisqu'elle est continue et coercive. Ce minimum étant un point critique il vérifie donc $Ax = b$.
- L'étude de la fonction $f(t) = \Phi(x_k - t\nabla\Phi(x_k))$ conduit au minimum α_k et comme $f'(\alpha_k)$, on a $\langle g_{k+1}, g_k \rangle = 0$.