Algorithme du gradient à pas optimal

Leçons concernées

- **162 :** Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.
- **219 :** Extremums : existence, caractérisation, recherche. Exemples et applications.
- **226 :** Suites vectorielles et réelles définies par une relation de récurrence un+1=f(un). Exemples. Applications à la résolution approchée d'équations.
- **233 :** Analyse numérique matricielle. Résolution approchée de systèmes linéaires, recherche d'éléments propres, exemples.
- **229 :** Fonctions monotones. Fonctions convexes. Exemples et applications.
- **253 :** Utilisation de la notion de convexité en analyse.

Remarque importante : Favorisez la partie "lemme de Kantorovitch" si vous voulez inclure ce développement dans les leçons 229 et 253.

Réference

- *Bernis - Analyse pour l'agrégation de mathématiques*

Le but de ce développement est d'étudier l'*algorithme du gradient à pas optimal*, appliqué à la fonctionnelle quadratique $\phi(x) = \frac{1}{2}||x||_A^2 - < x, b >$, où $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique définie positive.

Théorème. *L'application ϕ atteint son minimum en un unique vecteur $\overline{x}$.*
Soit de plus, pour $a \in \mathbb{R}^n$ donné, la suite (x_k) définie par récurrence par les relations :

$$\begin{cases} x_0 & = & a & \\ x_{k+1} & = & x_k & si\ \nabla\phi(x_k) = 0 \\ x_{k+1} & = & x_k - \alpha_k \nabla\phi(x_k) & sinon,\ avec\ \alpha_k = argmin_{t>0}[\phi(x_k - t\nabla\phi(x_k))] \end{cases}$$

Alors (x_k) converge vers $\overline{x}$, et on a de plus l'estimation de la vitesse de convergence :

$$\forall k \in \mathbb{N}, ||x_{k+1} - \overline{x}|| \leq \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}\right)^{k+1} ||x_0 - \overline{x}||$$

où $||.||$ désigne la norme euclidienne, λ_{max} (resp.λ_{min}) la plus grande (resp. petite) valeur propre de A

La preuve va s'appuyer sur le lemme suivant, que nous allons démontrer :

Lemme. *(de Kantorovitch)*
Pour tout x non nul de $\mathbb{R}^n$, on a : $\dfrac{||x||^4}{||x||^2_{A^{-1}}||x||^2_A} \geq 4\dfrac{\lambda_{max}\lambda_{min}}{(\lambda_{max}+\lambda_{min})^2}$

Preuve du lemme : Par le théorème spectral, A admet une base orthonormée de vecteurs propres $(e_1, ..., e_n)$ associés aux valeurs propres strictement positives $(\lambda_1, ..., \lambda_n)$. Ecrivons $x = \sum_{i=1}^{n} x_i e_i$. Alors :

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

en utilisant l'inégalité $ab \leq \frac{a^2+b^2}{2}$ valable pour tout réels a et b.

Or, $t \in [\lambda_{min}; \lambda_{max}] \longrightarrow \frac{t}{\lambda_{max}} + \frac{\lambda_{min}}{\lambda_i}$ est une fonction convexe, en tant que somme de fonctions convexes. Donc, $\forall i, g(\lambda_i) \leq g(\lambda_{min}) = g(\lambda_{max}) = 1 + \frac{\lambda_{min}}{\lambda_{max}}$.

On a alors au final :

$$||x||_A||x||_{A^{-1}} \leq \frac{1}{2}\frac{\lambda_{max}+\lambda_{min}}{\sqrt{\lambda_{max}\lambda_{min}}}||x||^2$$

En élevant au carré, on retrouve alors le résultat.

Preuve du théorème : Tout d'abord, ϕ est continue coercive, donc admet un minimum. Ce minimum est en particulier un point critique. Or, $\forall x \in \mathbb{R}^n, \nabla\phi(x) = Ax - b$. Donc les points critiques sont solutions de $Ax = b$, et il n'y en a qu'une seule, que nous allons noter $\overline{x}$. En particulier, il n'y a qu'un seul minimum, et il ne peut être que $\overline{x}$, d'où l'intérêt de l'algorithme du gradient.

Soit $k \in \mathbb{N}$ tel que $\nabla\phi(x_k) \neq 0$.

Commençons par exprimer x_{k+1} en cherchant α_k. Soit f définie par : $\forall t \in \mathbb{R}, f(t) = \phi(x_k - t\nabla\phi(x_k))$. Alors un calcul direct donne $\forall t \in \mathbb{R}, f(t) = \phi(x_k) - t < Ax_k - b, \nabla\phi(x_k) > +\frac{t^2}{2}||\nabla\phi(x_k)||^2_A$ soit $f(t) = \phi(x_k) - t||\nabla\phi(x_k)||^2 + \frac{t^2}{2}||\nabla\phi(x_k)||^2_A$. On en déduit que le minimum est atteint en $\alpha_k = \dfrac{||\nabla\phi(x_k)||^2}{||\nabla\phi(x_k)||^2_A}$. En particulier, $f'(\alpha_k) = 0$ soit $- < \nabla\phi(x_k - \alpha_k\nabla\phi(x_k)), \nabla\phi(x_k) >= 0$ soit $< \nabla\phi(x_{k+1}), \nabla\phi(x_k) >= 0$, ce qui sera utile par la suite.

Notons pour alléger $g_k = \nabla\phi(x_k)$. Estimons l'erreur en norme A :

$$||x_{k+1} - \overline{x}||^2_A =< A(x_{k+1} - \overline{x}), x_{k+1} - x_k > + < A(x_{k+1} - \overline{x}), x_k - \overline{x} >$$

Or, $x_{k+1} - x_k = -\alpha_k g_k$ et $A(x_{k+1} - \overline{x}) = Ax_{k+1} - b = g_{k+1}$. Donc, g_k et g_{k+1} étant orthogonaux, le premier terme est nul.

Donc, en se rappelant que A est symétrique :

$$\begin{aligned}||x_{k+1}-\overline{x}||_A^2 &=< A(x_{k+1}-x_k), x_k-\overline{x} > + < A(x_k-\overline{x}), x_k-\overline{x} > \\ &=< x_{k+1}-x_k, A(x_k-\overline{x}) > + ||x_k-\overline{x}||_A \\ &= -\alpha_k < g_k, g_k > + ||x_k-\overline{x}||_A^2 \\ &= -\frac{||g_k||^4}{||g_k||_A^2} + ||x_k-\overline{x}||_A^2\end{aligned}$$

Or, $||x_k-\overline{x}||_A^2 =< A(x_k-\overline{x}), x_k-\overline{x} >=< g_k, A^{-1}(Ax_k-b) >= ||g_k||_{A^{-1}}^2$. Au final :

$$||x_{k+1}-\overline{x}||_A^2 = \left(1-\frac{||g_k||^4}{||g_k||_A^2||g_k||_{A^{-1}}^2}\right)||x_k-\overline{x}||_A^2$$

On reconnait un terme qui fait penser au lemme de Kantorovitch. On obtient ainsi :

$$\begin{aligned}||x_{k+1}-\overline{x}||_A^2 &\leq \left(1-4\frac{\lambda_{max}\lambda_{min}}{(\lambda_{max}+\lambda_{min})^2}\right)||x_k-\overline{x}||_A^2 \\ ||x_{k+1}-\overline{x}||_A &\leq \left(\frac{\lambda_{max}-\lambda_{min}}{\lambda_{max}+\lambda_{min}}\right)||x_k-\overline{x}||_A \\ ||x_{k+1}-\overline{x}||_A &\leq \left(\frac{\lambda_{max}-\lambda_{min}}{\lambda_{max}+\lambda_{min}}\right)^{k+1}||x_0-\overline{x}||_A\end{aligned}$$

Ceci suffit à prouver la convergence de l'algorithme du gradient. On va cependant affiner notre inégalité pour travailler, à la place, avec la norme euclidienne. On remarque tout d'abord, en décomposant dans la base de vecteur propre de A, que :

$$\sqrt{\lambda_{min}}||.|| \leq ||.||_A \leq \sqrt{\lambda_{max}}||.||$$

Ceci donne alors :

$$||x_{k+1}-\overline{x}|| \leq \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}\left(\frac{\lambda_{max}-\lambda_{min}}{\lambda_{max}+\lambda_{min}}\right)^{k+1}||x_0-\overline{x}||$$

Remarque : On peut observer l'influence du conditionnement de A sur la vitesse de convergence. Sachant que A est symétrique définie positive, on a alors $cond(A) = \frac{\lambda_{max}}{\lambda_{min}}$. L'inégalité devient ainsi :

$$||x_{k+1}-\overline{x}|| \leq \sqrt{cond(A)}\left(\frac{cond(A)-1}{cond(A)+1}\right)^{k+1}||x_0-\overline{x}||$$

On observe alors que plus le conditionnement de A est proche de 1, plus la convergence est rapide, et inversement.

Donnons un complément utile, car très proche de l'algorithme du gradient à pas optimal. Il existe un autre algorithme, assez analogue, appelé *algorithme du gradient conjugué*. Soit $r_k = b - Ax_k = -g_k$. Nous avons vu précédemment que pour tout k, $r_k \perp r_{k+1}$. On a donc une relation d'orthogonalité intéressante, mais que pour les résidus d'ordre k et $k+1$ à chaque étape k.

Soit $K_k = Vect(r_0, ..., r_k)$. On peut, pour améliorer l'algorithme, demander à ce que x_{k+1} vérifie :

$$x_{k+1} \in (x_0 + K_k) \text{ et } r_{k+1} = b - Ax_{k+1} \perp K_k$$

Ceci définit l'algorithme du gradient conjugué. On constate alors que nous avons une relation d'orthogonalité plus forte que pour l'algorithme du gradient à pas optimal : cette fois-ci, r_{k+1} est orthogonal à la famille $(r_0, ..., r_k)$! En particulier, la famille des résidus est une famille libre de $\mathbb{R}^n$. Ceci garantie alors que cet algorithme converge en au plus n itérations, rang à partir duquel la suite devient constante égale à $\overline{x}$.

On peut démontrer que K_k est l'*espace de Krylov* associé à r_0, c'est-à-dire $K_k = Vect(r_0, Ar_0, ..., A^k r_0)$. On peut alors dans la pratique prendre une base orthogonale (p_k) de cet espace de Krylov, pour le produit scalaire $< .|. >_A$. On parle alors de vecteurs *conjugués par rapport à A*, d'où le nom de l'algorithme.