Méthode du gradient à pas conjugué

Benjamin Hellouin

Dumas, A.VIII "Minimisation d'une fonctionnelle"

Soit $A \in S_n^{++}$ et $b \in \mathbb{R}^n$. On cherche à résoudre le système linéaire $Ax = b$ itérativement. On pose $x_0 \in \mathbb{R}^n$ arbitrairement et $r_0 = b - Ax_0$. L'*espace de Krylov de dimension k* associé à r_0 est défini comme $K_k = Vect(r_0, \dots A^k(r_0))$.

Théorème 1. *Il existe une unique suite x_k vérifiant les conditions équivalentes suivantes :*
- *x_{k+1} minimise la fonctionnelle $J(x) = \frac{1}{2} < Ax, x > - < b, x >$ dans l'espace affine $x_0 + K_k$;*
- *$r_{k+1} = b - Ax_{k+1}$ est orthogonal à K_k.*

De plus, cette suite stationne en au plus n itérations vers une limite u telle que $Au = b$.

Notons que la condition (i), si elle était vraie dans l'espace entier, suffirait à garantir que x_k est une solution du système.

Preuve. On montre d'abord l'équivalence des deux conditions. Montrons que J est strictement convexe : on a

$$J(\theta x + (1-\theta)y) = \theta J(x) + (1-\theta)J(y) - \theta(1-\theta) < A(x-y), (x-y) >$$

et J est coercive. Par conséquent, $g : y \mapsto J(x_0 + y)$ admet un unique minimum dans $K_k(r_0)$ qu'on peut noter $x_{k+1} - x_0$. Ainsi,

$$\forall y \in K_k, J(x_{k+1} + y) > J(x_{k+1}) \quad \text{soit} \quad \frac{1}{2} < Ay, y > + < Ax_{k+1} - b, y > > 0$$

(avec un passage au symétrique). En remplaçant y par ty, on obtient un polynôme de degré 2, toujours positif, qui ne s'annule qu'en 0 : il vaut donc t^2, i.e. $< Ax_{k+1} - b, y >= 0$. Ceci étant vrai pour tout $y \in K_k$, il s'agit exactement de la propriété (ii). Inversement, si on suppose la propriété (ii) vraie pour x_{k+1}, alors ce polynôme admet un minimum en 0 pour toute direction y, et il s'agit donc d'un minimum sur $x_0 + K_{k_0}$.

Pour démontrer la convergence de la suite (x_k), on constate que la suite K_k est croissante pour l'inclusion. Plus précisément, elle est strictement croissante puis stationnaire à partir d'un rang k_0 (puisque elle est incluse dans $\mathbb{R}^n$). On distingue les cas :
- $k_0 = n - 1$: alors K_{k_0} est l'espace entier et x_{k+1} est un minimum absolu, d'où $Ax_{k+1} = b$.
- $k_0 < n - 1$: alors $A^{k_0+1}(r_0)$ appartient à K_{k_0}, soit

$$A_{k_0+1}(r_0) = \sum_{i=0}^{k_0} \alpha_i A^i(r_0)$$

avec $\alpha_0 \neq 0$ soit, en divisant par α_0 :

$$A\left(\frac{1}{\alpha_0}A^{k_0}(r_0) - \sum_{i=1}^{k_0}\frac{\alpha_i}{\alpha_0}A^{i-1}(r_0) + x_0\right) = b.$$

Autrement dit, la solution de $Ax = b$ appartient à l'espace $x_0 + K_{k_0}$, et il s'agit bien de x_{k_0+1} par la propriété (i).

Corollaire 1. *Par construction, les suites r_k et $d_k = x_{k+1} - x_k$ sont orthogonales (conjuguées !) resp. pour les produits scalaires $< .,. >$ et $< A.,. >$.*

Preuve. Le premier point vient directement du fait que $r_k \in K_k$ pour tout k et de la propriété (ii). D'autre part, on a aussi $d_k \in K_k$ et, pour $l < k$,

$$< Ad_k, d_l >=< Ax_{k+1} - Ax_k, d_l >=< r_k - r_{k+1}, d_l >= 0$$

puisque r_k et r_{k+1} sont orthogonaux à K_l.

La famille d_k correspond au pas de la suite x_k, et est conjugué par rapport à A, ce qui justifie le nom de la méthode.