Méthode de Kacmarz

Clarence Kineider

Leçons : 162, 208, 226, 233

Référence : $\emptyset$

On souhaite résoudre un système linéaire de la forme $Ax = b$ où $A \in GL_d(\mathbf{R})$ et $b \in \mathbf{R}^d$. Notons $\overline{x}$ l'unique solution du système.

On écrit la matrice A sous la forme $A = \begin{pmatrix} {}^t a_1 \\ \hline {}^t a_2 \\ \hline \vdots \\ \hline {}^t a_d \end{pmatrix}$, de sorte que les $a_i \in \mathbf{R}^d$ soient les vecteurs colonnes correspondants aux lignes de A. On pose pour $i \in \{1, \dots, d\}$, $u_i = \dfrac{a_i}{||a_i||} \in \mathbf{R}^d$ et $\beta_i = \dfrac{b_i}{||a_i||} \in \mathbf{R}$.
On définit les hyperplans affines $H_i = \{z \in \mathbf{R}^d \mid {}^t u_i z = \beta_i\} = \beta_i u_i + Vect(u_i)^\perp$ et les hyperplans vectoriels associés $E_i = Vect(u_i)^\perp$. Ces hyperplans affines sont l'écriture ligne par ligne de l'égalité $Az = b$. On a donc naturellement le résultat suivant :

Lemme : On a $\bigcap\limits_{i=1}^{d} H_i = \{\overline{x}\}$.

Démonstration : Soit $z \in \mathbf{R}$. On a :

$$\begin{aligned} z \in \bigcap_{i=1}^{d} H_i &\Leftrightarrow \forall i \in \{1, \dots, d\} \frac{{}^t a_i}{||a_i||} z = \frac{b_i}{||a_i||} \\ &\Leftrightarrow \forall i \in \{1, \dots, d\}\ {}^t a_i z = b_i \\ &\Leftrightarrow Az = b. \end{aligned}$$

$\square$

On fixe maintenant $x_0 \in \mathbf{R}^d \setminus \{0\}$ (pint de départ de notre algorithme itératif).
L'idée de la méthode de Kacmarz est de considérer les projections orthogonales successives de x_0 sur les hyperplans H_i. En notant M_i la projection orthogonale sur le sous-espace vectoriel E_i pour $i \in \{1, \dots, d\}$, on a :

$$
\begin{aligned}
x_1 &= M_1 x_0 + \beta_1 u_1 \\
x_2 &= M_2 x_1 + \beta_2 u_2 \\
&\vdots \\
x_d &= M_d x_{d-1} + \beta_d u_d \\
x_{d+1} &= M_1 x_d + \beta_1 u_1 \\
&\vdots
\end{aligned}
$$

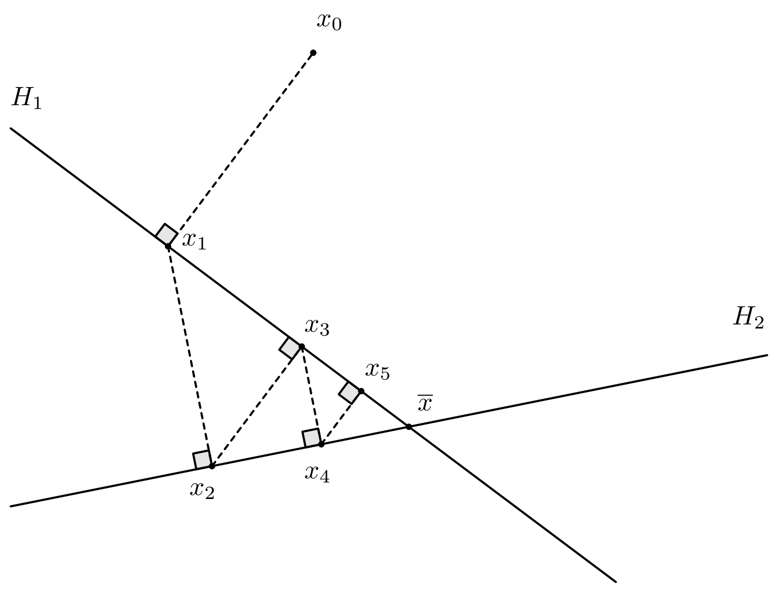

FIGURE 1 – Début de la méthode de Kacmarz dans $\mathbf{R}^2$

On obtient ainsi une suite $(x_n)_{n\in\mathbf{N}} \in (\mathbf{R}^d)^{\mathbf{N}}$.

Proposition : La suite $(x_n)_n$ converge vers la solution $\overline{x}$ du système $Ax = b$.

Démonstration : Posons, pour tout $n \in \mathbf{N}$, $\epsilon_n = x_n - \overline{x}$ (l'erreur au rang n), et montrons que $||\epsilon_n|| \underset{n\to\infty}{\longrightarrow} 0$.
Soit $n \in \mathbf{N}^*$, on a $\epsilon_n = x_n - \overline{x} = M_n x_{n-1} - (\overline{x} - \beta_n u_n)$ (pour les indices sous M, β et u, on considère le représentant de n [mod d] pris entre 1 et d).
Or, $\overline{x} \in H_n = E_n + \beta_n u_n$, donc $\overline{x} - \beta_n u_n \in E_n$, et donc $M_n(\overline{x} - \beta_n u_n) = \overline{x} - \beta_n u_n$. Donc $\epsilon_n = M_n(x_{n-1} + \beta_n u_n - \overline{x})$.
Comme $u_n \in E_n^{\perp}$, on a $M_n(\beta_n u_n) = 0$, donc $\epsilon_n = M_n(x_{n-1} - \overline{x}) = M_n \epsilon_{n-1}$.
De plus, les M_i sont des matrices de projections orthogonales, leur norme est donc égale à 1, et on a :

$$||\epsilon_n|| \leqslant ||M_n||||\epsilon_{n-1}|| = ||\epsilon_{n-1}||.$$

La suite $(||\epsilon_n||)_n$ est donc décroissante. Puisqu'elle est minorée par 0, elle converge.

Il reste à montrer que sa limite est 0. Pour cela, il suffit de trouver une sous-suite qui converge vers 0. On considère la sous-suite $(\epsilon_{nd})_n$, et on pose $M = M_d M_{d-1} \dots M_1$. On a ainsi $\epsilon_{nd} = M\epsilon_{(n-1)d}$.

Montrons que $||M|| < 1$. Soit $x \in \mathbf{R^d} \setminus \{0\}$.

— S'il existe $i \in \{1, \dots, d\}$ tel que $||M_i M_{i-1} \cdots M_1 x|| < ||x||$, alors

$$||Mx|| \leqslant ||M_d|| \cdot ||M_{d-1}|| \cdots ||M_{i+1}|| \cdot ||M_i M_{i-1} \cdots M_1 x|| = ||M_i M_{i-1} \cdots M_1 x|| < ||x||.$$

— Sinon, on a pour tout $i \in \{1, \dots, d\}$, $||M_i \cdots M_1 x|| = ||x||$. En particulier, $||M_1 x|| = ||x||$. Cela signifie que $x \in E_1$ et donc $M_1 x = x$. On en déduit par récurrence que $x \in \bigcap_{i=1}^{d} E_i = \{0\}$ car $\forall i \in \{1, \dots, d\}, E_i = Vect(u_i)^\perp$ et les u_i forment une base de $\mathbf{R}^d$. Donc $x = 0$, absurde.

Ainsi, $||Mx|| < ||x||$ pour tout $x \in \mathbf{R}^d \setminus \{0\}$. Or $||M|| = \sup_{x \in \mathbf{S}^{d-1}} \dfrac{||Mx||}{||x||}$, et comme la dimension est finie, $\mathbf{S}^{d-1}$ est compacte, le sup est atteint et donc $||M|| = \max_{x \in \mathbf{S}^{d-1}} \dfrac{||Mx||}{||x||} < 1$.

Ainsi, $\forall n \in \mathbf{N}, ||\epsilon_{nd}|| \leq ||M||^n ||\epsilon_0||$, avec $||M|| < 1$. La suite converge donc bien vers 0.

$\square$

Remarques :

- Quelle est la complexité de cette méthode ? À chaque itération, on doit calculer

$$x_{n+1} = M_{n+1} x_n + \beta_{n+1} u_{n+1}.$$

  Pas besoin de calculer un produit matrice-vecteur en $O(d^2)$ opérations puisque $M_i x = x - u_i \,{}^{\mathrm{t}}u_i x = x - \langle x, u_i \rangle u_i$. Le calcul de x_{n+1} à partir de x_n se fait donc en $O(d)$ opérations. On converge vers la solution à vitesse géométrique toute les d itérations, la complexité est donc en $O(d^2)$ opérations pour une convergence géométrique, comme les méthodes usuelles (Jacobi, Gauss-Seidel etc.)

- Cette méthode fonctionne également lorsque la matrice A n'est pas inversible. Alors l'intersection des hyperplans affines H_i n'est pas réduit à un unique point (mais est toujours l'ensemble des solutions du système) et la suite (x_n) convergera vers la projection orthogonale de x_0 sur cette intersection.

Merci à David XU *pour ce développement.*