Méthode du gradient conjugué

Gaëtan Hello

Université d'Evry Val d'Essonne
UFR S&T - Laboratoire de Mécanique et d'Energétique d'Evry
gaetan.hello@ufrst.univ-evry.fr

Semestre d'automne 2013

Contexte Eléments-Finis

- La méthode des éléments-finis est une méthode numérique permettant la résolution approchée d'équations aux dérivées-partielles.
- Elle consiste à transformer un problème continu initial en un problème algébrique associé pouvant être résolu par des techniques automatisables/programmables.
- On se ramène ainsi à la résolution de :

$$K \cdot u = f \tag{1}$$

avec

  - $K \in \mathbb{R}^{n \times n}$ matrice de rigidité symétrique et définie positive,
  - $u \in \mathbb{R}^n$ vecteur des inconnues nodales,
  - $f \in \mathbb{R}^n$ vecteur des efforts nodaux.

Méthodes de résolution

Pour résoudre un système linéaire, il existe deux grandes familles de méthodes :

- les méthodes directes (nombre fini d'opérations) :
  - pivot de Gauss,
  - décomposition LU : $A = LU$ (L lower triangular, U upper triangular),
  - décomposition QR : $A = QR$ (Q orthogonale, R upper tri.),
  - décomposition de Cholesky : $A = LDL^T$ (A p.d., L lower tri., D diag.),
  - ...
- les méthodes itératives (nombre d'opérations lié à la convergence) :
  - Gauss-Seidel,
  - SOR (Successive Over Relaxation),
  - gradient conjugué (A p.d.),
  - GMRES (Generalized Minimal Residual Method)
  - ...

2 Méthode du gradient conjugué

Problème de minimisation

Comme la matrice de rigidité du système linéaire E.F. est symétrique définie positive, nous allons exploiter cette particularité pour définir une méthode de résolution adaptée. Etant donné le système linéaire :

$$A \cdot x = b \tag{2}$$

avec

- $A \in \mathbb{R}^{n \times n}$ matrice symétrique et définie positive,
- $x \in \mathbb{R}^n$ vecteur des inconnues,
- $f \in \mathbb{R}^n$ vecteur second membre.

Problème de minimisation

Si l'on définit la fonction :

$$\phi(x) = \frac{1}{2} x^T \cdot A \cdot x - x^T \cdot b \tag{3}$$

Alors, comme A est s.p.d., la solution x^* de $A \cdot x = b$ est aussi solution du problème de minimisation :

$$x^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}}\ \phi(x) \tag{4}$$

Méthode de la plus forte pente

En un point x_i donné, la fonction $\phi(x)$ décroit le plus vite dans la direction :

$$\begin{aligned} r(x_i) &= -\nabla\phi(x_i) \\ r(x_i) &= b - A \cdot x_i \end{aligned} \tag{5}$$

Connaissant la direction de plus forte décroissance, il s'agit maintenant de trouver la position selon cette direction où $\phi(x)$ est minimale. Si l'on note :

$$x_{i+1} = x_i + \alpha_i \cdot r(x_i) \tag{6}$$

Alors α_i vérifie (propriété du minimum local) :

$$\begin{aligned} \frac{\partial\phi}{\partial\alpha_i}\left(x_{i+1}\right) &= 0 \\ \frac{\partial\phi}{\partial\alpha_i}(x_i + \alpha_i \cdot r(x_i)) &= 0 \\ \alpha_i &= \frac{r_i^T \cdot r_i}{r_i^T \cdot A \cdot r_i} \end{aligned} \tag{7}$$

Méthode de la plus forte pente

Algorithme de la plus forte pente (Steepest Descent)

x_0=initial guess
$r_0 = b - A \cdot x_0$
$k = 0$
while $\|r_k\|_2 > \varepsilon$ **do**
  $k = k + 1$
  $\alpha_k = \frac{r_{k-1}^T \cdot r_{k-1}}{r_{k-1}^T \cdot A \cdot r_{k-1}}$
  $x_k = x_{k-1} + \alpha_k \cdot r_{k-1}$
  $r_k = b - A \cdot x_k$
end while

Méthode à directions générales

Pour éviter les problèmes de convergence parfois rencontrés par la méthode de plus forte pente, on peut chercher à minimiser $\phi(x)$ selon des directions $P = \{p_1, p_2, ...\}$ qui ne coïncident pas nécessairement avec les résidus $R = \{r_1, r_2, ...\}$.

Dans ce cas, $\phi(x_{i-1} + \alpha \cdot p_i)$ est minimisé avec :

$$\alpha = \alpha_k = \frac{r_{k-1}^T \cdot p_k}{p_k^T \cdot A \cdot p_k} \qquad (8)$$

Méthode à directions générales

Algorithme de minimisation à directions générales

x_0=initial guess
$r_0 = b - A \cdot x_0$
$k = 0$
while $\|r_k\|_2 > \varepsilon$ **do**
  $k = k + 1$
  choisir p_k tel que $r_{k-1}^T \cdot p_k \neq 0$
  $\alpha_k = \frac{r_{k-1}^T \cdot p_k}{p_k^T \cdot A \cdot p_k}$
  $x_k = x_{k-1} + \alpha_k \cdot p_k$
  $r_k = b - A \cdot x_k$
end while

Directions A-conjuguées

Si $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ non-nuls, on dit qu'ils sont A-conjugués si :

$$x^T \cdot A \cdot y = 0 \tag{9}$$

Comme A est s.p.d, la A-conjugaison permet de définir un produit scalaire :

$$\begin{aligned} \langle x, y \rangle_A &= x^T \cdot A \cdot y = y^T \cdot A \cdot x \\ \langle x, x \rangle_A &> 0 \end{aligned} \tag{10}$$

Si $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ non-nuls vérifient $x^T \cdot A \cdot y = 0$, alors ils sont orthogonaux au sens du produit scalaire $\langle \square, \square \rangle_A$.

Directions A-conjuguées

On suppose connaître une famille P de n vecteurs de $\mathbb{R}^n$ A-conjugués 2 à 2 ($P = \{p_1, ..., p_n\}$). On a donc par définition :

$$\begin{aligned} \langle p_i, p_j \rangle_A &= 0 \quad && , \; i \neq j \\ &= p_i^T \cdot A \cdot p_i \quad && , \; i = j \end{aligned} \tag{11}$$

P est une famille de $\mathbb{R}^n$ libre (par la A-conjugaison) et génératrice (n termes). P est donc une base de $\mathbb{R}^n$.

Si x^* est la solution de $A \cdot x = b$, il existe alors n coefficients α_i tels que :

$$\begin{aligned} x^* &= \sum_{i=1}^{n} \alpha_i \cdot p_i \\ x^* &= \sum_{i=1}^{n} \frac{\langle x^*, p_i \rangle_A}{\langle p_i, p_i \rangle_A} \cdot p_i \end{aligned} \tag{12}$$

Directions A-conjuguées

On a alors

$$\begin{aligned}
x^* &= \sum_{i=1}^{n} \alpha_i \cdot p_i \\
A \cdot x^* &= \sum_{i=1}^{n} \alpha_i \cdot A \cdot p_i \\
b &= \sum_{i=1}^{n} \alpha_i \cdot A \cdot p_i \\
p_j^T \cdot b &= \sum_{i=1}^{n} \alpha_i \cdot p_j^T \cdot A \cdot p_i \\
p_j^T \cdot b &= \alpha_j \cdot p_j^T \cdot A \cdot p_j \\
\alpha_j &= \frac{p_j^T \cdot b}{p_j^T \cdot A \cdot p_j}
\end{aligned} \tag{13}$$

Directions A-conjuguées

Algorithme de minimisation à directions A-conjuguées

x_0=initial guess
$r_0 = b - A \cdot x_0$
$k = 0$
while $\|r_k\|_2 > \varepsilon$ **do**
  $k = k + 1$
  choisir p_k tel que $r_{k-1}^T \cdot p_k \neq 0$ et $p_k \in \text{span}\{A \cdot p_1, ..., A \cdot p_{k-1}\}^\perp$
  $\alpha_k = \frac{b^T \cdot p_k}{p_k^T \cdot A \cdot p_k}$
  $x_k = x_{k-1} + \alpha_k \cdot p_k$
  $r_k = b - A \cdot x_k$
end while

Gradient conjugué 1

Algorithme de minimisation du gradient conjugué 1 (convergence assurée en au plus n itérations)

x_0=initial guess, $r_0 = b - A \cdot x_0$, $k = 0$
while $\|r_k\|_2 > \varepsilon$ **do**
  $k = k + 1$
  if k=1 **then**
    $p_1 = r_0$
  else
    $p_k = r_{k-1} - \sum_{i=1}^{k-1} \frac{r_{k-1}^T \cdot A \cdot p_i}{p_i^T \cdot A \cdot p_i} \cdot p_i$
  end if
  $\alpha_k = \frac{b^T \cdot p_k}{p_k^T \cdot A \cdot p_k}$
  $x_k = x_{k-1} + \alpha_k \cdot p_k$
  $r_k = b - A \cdot x_k$
end while

Gradient conjugué 2

Algorithme de minimisation du gradient conjugué 2 : on ne projette la direction de plus forte pente r_{k-1} que sur la dernière direction p_{k-1} au lieu de $\{p_{k-1}, p_{k-2}, ..., p_1\}$ pour définir la nouvelle direction p_k (convergence légèrement plus lente et non nécessairement uniforme mais moins de mémoire requise)

$$p_k = r_{k-1} - \sum_{i=1}^{k-1} \frac{r_{k-1}^T \cdot A \cdot p_i}{p_i^T \cdot A \cdot p_i} \cdot p_i \rightarrow p_k = r_{k-1} - \frac{r_{k-1}^T \cdot A \cdot p_{k-1}}{p_{k-1}^T \cdot A \cdot p_{k-1}} \cdot p_{k-1} \quad (14)$$

$$p_k = r_{k-1} - \sum_{i=1}^{k-1} \frac{\langle r_{k-1}, p_i \rangle_A}{\langle p_i, p_i \rangle_A} \cdot p_i \rightarrow p_k = r_{k-1} - \frac{\langle r_{k-1}, p_{k-1} \rangle_A}{\langle p_{k-1}, p_{k-1} \rangle_A} \cdot p_{k-1} \quad (15)$$

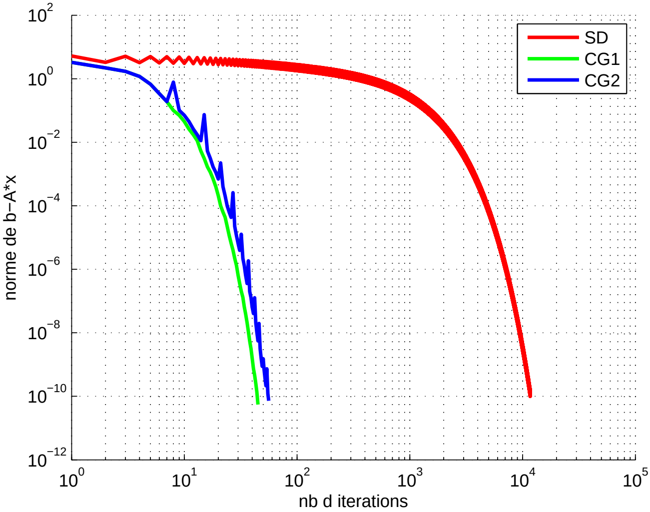
Gradient conjugué 2

Algorithme de minimisation des gradient conjugué 2

x_0=initial guess, $r_0 = b - A \cdot x_0$, $k = 0$
while $\|r_k\|_2 > \varepsilon$ **do**
  $k = k + 1$
  if k=1 **then**
    $p_1 = r_0$
  else
    $p_k = r_{k-1} - \frac{r_{k-1}^T \cdot A \cdot p_{k-1}}{p_{k-1}^T \cdot A \cdot p_{k-1}} \cdot p_{k-1}$
  end if
  $\alpha_k = \frac{r_{k-1}^T \cdot p_k}{p_k^T \cdot A \cdot p_k}$
  $x_k = x_{k-1} + \alpha_k \cdot p_k$
  $r_k = b - A \cdot x_k$
end while

Exemple 1

A matrice pleine s.p.d., $A \in \mathbb{R}^{100\times 100}$, $\kappa(A) = \left|\frac{\lambda_{max}}{\lambda_{min}}\right| \approx 1249$, $\varepsilon \approx 10^{-10}$



Nb Itérations	
SD	11688
CG1	46
CG2	57