Chapitre 4

LA MÉTHODE DU GRADIENT CONJUGUÉ

Chapitre 4

LA MÉTHODE DU GRADIENT CONJUGUÉ

4.1 Introduction

La méthode du gradient conjugué permet de résoudre les systèmes linéaires dont la matrice est symétrique définie positive.

Il s'agit d'une méthode qui consiste à partir d'un vecteur donné x^0 et à déterminer à chaque étape un vecteur p^k et un scalaire α^k permettant de calculer x^{k+1} à partir de x^k par

$$x^{k+1} = x^k + r^k p^k \tag{4.1}$$

avec l'objectif de minimiser une fonction (descendre un potentiel).

Avant d'introduire la méthode du gradient conjugué pour résoudre un système linéaire, il est utile d'aborder sommairement les méthodes du gradient.

4.2 Les méthodes du gradient

On considère le problème de minimisation :

$$\begin{cases} \text{Trouver } \widetilde{x} \in I\!R^n \\ J(\widetilde{x}) \leq J(x) \quad \text{pour tout } x \in I\!R^n \end{cases}$$

où J est une fonction donnée de $I\!R^n$ dans $I\!R$.

Dans ce paragraphe, nous allons choisir dans (4.1) le gradient de la fonction à minimiser comme direction de descente :

$$p^k = \nabla J(x^k) \tag{4.2}$$

où $J : x \in I\!R^n \mapsto J(x) \in I\!R$ est la fonction à minimiser. Le paramètre α^k sera d'abord choisi constant, ensuite optimal.

Le but étant de résoudre le système linéaire :

$$Ax = b, \tag{4.3}$$

où A est une matrice d'ordre n symétrique définie positive et $b \in I\!R^n$. Nous considérons la fonction J suivante

$$J(x) \equiv \frac{1}{2}(Ax, x) - (b, x) \tag{4.4}$$

Ce choix est motivé par le résultat de la proposition suivante.

Proposition 4.1

Soient $A \in \mathcal{M}_n(I\!R)$ une matrice symétrique définie positive et $b \in I\!R^n$. Alors $\widetilde{x} \in I\!R^n$ est solution du système linéaire (3.3) si et seulement si

$$\forall x \in I\!R^n, \qquad J(\widetilde{x}) \leq J(x) \tag{4.5}$$

où J est la fonction définie par (4.4).

Démonstration

Supposons que $\widetilde{x}$ est solution de (4.3), on a alors pour tout $x \in I\!R^n$

$$\begin{aligned} J(x) - J(\widetilde{x}) &= \frac{1}{2}(Ax,x) - (b,x) - \frac{1}{2}(A\,\widetilde{x},\widetilde{x}) + (b,\widetilde{x}) \\ &= \frac{1}{2}(Ax,x) - \frac{1}{2}(A\,\widetilde{x},\widetilde{x}) - (b,x-\widetilde{x}) \\ &= \text{ (car } \widetilde{x} \text{ est solution de (4.3))} \\ &= \frac{1}{2}(Ax,x) - \frac{1}{2}(A\,\widetilde{x},\widetilde{x}) - (A\,\widetilde{x},x-\widetilde{x}) \\ &= \frac{1}{2}(Ax,x) + \frac{1}{2}(A\,\widetilde{x},\widetilde{x}) - (A\,\widetilde{x},x) \\ &= \frac{1}{2}(A(x-\widetilde{x}),x-\widetilde{x}) \text{ (car } A \text{ est symétrique)} \\ &\geq 0 \text{ (car } A \text{ est positive)} \end{aligned}$$

Réciproquement, supposons que (4.5) est vérifiée. On introduit, pour z donné dans $I\!R^n$, la fonction $f : I\!R \to I\!R$ définie par

$$f(t) \equiv J(\widetilde{x} + tz)$$

On a d'après (4.5) :

$$\forall t \in I\!R, \qquad f(0) \leq f(t)$$

et par conséquent, zéro est un minimum de la fonction dérivable f, d'où

$$f'(0) = 0.$$

On a par ailleurs

$$\begin{aligned} f'(0) &= \lim_{t\to 0} \frac{f(t)-f(0)}{t} \\ &= \lim_{t\to 0} \frac{\frac{1}{2}(A(\widetilde{x}+tz),\widetilde{x}+tz) - (b,\widetilde{x}+tz) - \frac{1}{2}(A\,\widetilde{x},\widetilde{x}) + (b,\widetilde{x})}{t} \\ &= \lim_{t\to 0} \frac{t\left((A\,\widetilde{x},z) - (b,z)\right) + \frac{1}{2}t^2(Az,z)}{t} = (A\,\widetilde{x} - b,z) \end{aligned}$$

Cela entraine que

$$\forall z \in I\!R^n, \qquad (A\,\widetilde{x} - b,z) = 0,$$

c'est-à-dire, que $\widetilde{x}$ est solution du système linéaire (4.3). ■

Remarques :
1) On rappelle que le gradient d'une fonction $G : I\!R^n \longrightarrow I\!R$ au point $x \in I\!R^n$ vérifie

$$\forall z \in I\!R^n, \qquad (\nabla G(x),z) = \lim_{t\to 0} \frac{G(x+tz) - G(x)}{t}$$

Il en résulte, en considérant la fonction définie par (4.4), que

$$\forall z \in I\!R^n, \qquad (\nabla J(x),z) = (Ax - b,z)$$

et que

$$\nabla J(x) = Ax - b$$

(Il suffit de considérer la fonction $g : t \in \mathbb{R} \to J(x + tz)$ et de calculer $g'(0)$)
2) On peut montrer en utilisant la définition suivante du gradient d'une fonction $G : x \in \mathbb{R}^n \to \mathbb{R}$ au point $a \in \mathbb{R}^n$:

$$\nabla G(a) \equiv \begin{pmatrix} \dfrac{\partial G}{\partial x_1}(a) \\ \cdot \\ \cdot \\ \dfrac{\partial G}{\partial x_n}(a) \end{pmatrix}$$

que $\nabla J(a) = Aa - b$. ■

Notation :

Lorsqu'il existe $\widetilde{x} \in \mathbb{R}^n$ vérifiant (4.5), on dit que $\widetilde{x}$ est solution du problème de minimisation suivant :

$$\inf_{x \in \mathbb{R}^n} J(x)$$

ou que $\widetilde{x}$ est solution du problème suivant

$$J(\widetilde{x}) = \min_{x \in \mathbb{R}^n} J(x). \tag{4.6}$$

Résoudre (4.3) est donc équivalent, d'après la proposition 4.1, à trouver $\widetilde{x}$ solution de (4.6). ■

Remarque :

Comment choisir $z \in \mathbb{R}^n$, pour que

$$J(x + tz) \leq J(x)$$

pour t "petit" ?

La réponse à cette question peut être obtenue en effectuant un développement limité de $J(x + tz)$ par rapport à t au voisinage de zéro :

$$J(x + tz) = J(x) + t\,(\nabla J(x), z) + O(t^2).$$

En choisissant de descendre selon "la plus grande pente" $z_0 = -\nabla J(x)$, on a

$$J(x + tz_0) = J(x) - t\,\|\nabla J(x)\|_2^2 + O(t^2) \leq J(x)$$

pour t suffisament "petit" et positif.

4.2.1 La méthode du gradient à pas fixe

La méthode du gradient à pas fixe est définie par l'algorithme suivant

$$\begin{cases} x^0 \text{ donné} \\ x^{k+1} = x^k + r\nabla J(x^k) = x^k + r(Ax^k - b) \end{cases} \tag{4.7}$$

Théorème 4.1

Si $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique définie positive, alors la méthode du gradient à pas fixe r, converge pour

$$r \in \left]-\frac{2}{\rho(A)}, 0\right[$$

où $\rho(A)$ est le rayon spectral de la matrice A. C'est-à-dire, que la suite générée par l'algorithme (4.7) converge, pour tout choix de x^0, vers l'unique solution du système linéaire :

$$Ax = b$$

Démonstration :

La matrice de la méthode itérative (4.7) est donnée par

$$B = I + rA \tag{4.8}$$

On sait d'après le théorème 3.1 que la suite définie par (4.7) converge si et seulement si

$$\rho(B) < 1 \tag{4.9}$$

Cela entraine que $r \neq 0$. Par ailleurs, on a en désignant par Sp(A) et Sp(B), respectivement les ensembles de valeurs propres de A et de B, que

$$Sp(B) = \{1 + r\lambda / \quad \lambda \in Sp(A)\} \tag{4.10}$$

En effet, il est clair que si $\lambda \in Sp(A)$ alors $1 + r\lambda \in Sp(B)$. Inversement, si $\delta \in Sp(B)$ alors il existe $u \in I\!R^n (u \neq 0)$ tel que

$$(I + rA)u = \delta u$$

d'où pour $r \neq 0$

$$Au = \frac{\delta - 1}{r} u$$

et $\delta = 1 + r\dfrac{\delta - 1}{r}$ avec $\dfrac{\delta - 1}{r} \in Sp(A)$. On déduit de (4.10), que la condition de convergence (4.9) est satisfaite si et seulement si

$$\forall \lambda \in Sp(A) \quad 1 + r\lambda \in]-1, 1[$$

On en déduit puisque A est définie positive et donc $Sp(A) \subset I\!R_+^*$, que

$$\forall \lambda \in Sp(A) \quad r \in \left]-\frac{2}{\lambda}, 0\right[$$

où de manière équivalente

$$r \in \left]-\frac{2}{\rho(A)}, 0\right[$$

■

Remarque :

Dans l'algorithme du gradient décrit par (4.7), le paramètre r est appelé le pas de la méthode.

4.2.2 La méthode du gradient à pas optimal

La méthode du gradient à pas optimal consiste à changer le paramètre r intervenant dans (4.7) à chaque itération k par un pas r^k, dit pas optimal, dont la construction est basée sur le résultat suivant.

Proposition 4.2

Soient J la fonction définie sur $I\!R^n$ par (4.4), x un vecteur de $I\!R^n$ et w un vecteur non nul de $I\!R^n$. On considère le problème

$$\begin{cases} \text{Trouver} \quad r \in I\!R \quad \text{tel que} \\ \forall t \in I\!R, \qquad J(x + rw) \leq J(x + tw) \end{cases} \tag{4.11}$$

Si la matrice A intervenant dans (4.4) est symétrique définie positive, alors (4.11) admet une solution unique donnée par

$$r = -\frac{(Ax - b, w)}{(Aw, w)} \tag{4.12}$$

Démonstration :

On considère la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(t) \equiv J(x + tw)$$

On a

$$\begin{aligned} f(t) &= J(x+tw) = \frac{1}{2}\left(A(x+tw), x+tw\right) - (b, x+tw) \\ &= \frac{1}{2}\left[(Ax,x) + 2t(Ax,w) + t^2(Aw,w)\right] - (b,x) - t(b,w) \\ &= \frac{1}{2}t^2(Aw,w) + t(Ax-b,w) + \frac{1}{2}(Ax,x) - (b,x). \end{aligned}$$

Puisque $(Aw, w) > 0$ (car A est définie positive et $w \neq 0$ par hypothèse), on en déduit que la fonction f admet un minimum global au point $t = r$ donné par

$$r = -\frac{(Ax-b,w)}{(Aw,w)}$$

■

Commentaire :

Le résultat de la proposition 4.2 justifie le choix suivant de r^k dans l'algorithme du gradient à pas optimal :

$$\left\{ \begin{array}{l} x^0 \text{ donné} \\ g^k = Ax^k - b \\ r^k = -\dfrac{\|g^k\|_2^2}{(Ag^k, g^k)} \\ x^{k+1} = x^k + r^k g^k \end{array} \right. \tag{4.13}$$

En effet la valeur de r^k calculée à la k^{ieme} itération de l'algorithme est optimale, car pour tout autre choix $t \in \mathbb{R}$, on a d'après la proposition 4.2 :

$$J(x^{k+1}) \leq J(x^k + tg^k)$$

Ce choix "permet de s'approcher" plus rapidement de $\widetilde{x} \in \mathbb{R}^n$ vérifiant

$$\forall x \in \mathbb{R}^n, \qquad J(\widetilde{x}) \leq J(x)$$

que par la stratégie du pas fixe décrite par (4.7). On a le résultat de convergence de la méthode du gradient à pas optimal suivant :

Théorème 4.2

Si $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique définie positive, alors la méthode du gradient à pas optimal converge. C'est-à-dire, que la suite générée par l'algorithme (4.13) converge, pour tout choix de x^0, vers l'unique solution du système linéaire :

$$Ax = b$$

Démonstration : *voir Lascaux et Théodore* [1987].

4.3 La méthode du gradient conjugué

On rappelle que la méthode du gradient à pas optimal est une méthode itérative de descente suivant la plus grande pente dans laquelle le passage de x^k à x^{k+1} se fait selon la formule

$$x^{k+1} = x^k + r^k g^k \tag{4.14}$$

où $g^k = \nabla J(x^k)$ et r^k est le pas à l'étape k. La méthode du gradient conjugué consiste à prendre comme première direction de descente la direction $w^0 = \nabla J(x^0)$ et à choisir à l'étape $k+1$ une direction w^{k+1} telle que

$$(w^{k+1}, Aw^k) = 0 \tag{4.15}$$

et un pas optimal r^k défini par

$$\forall t \in I\!R, \qquad J(x^k + r^k w^k) \leq J(x^k + tw^k) \tag{4.16}$$

Ainsi les directions de descente w^k et w^{k+1} sont A-conjuguées ou orthogonales au sens du produit scalaire défini par la matrice symétrique définie positive A.

Proposition 4.3

Pour x^k donné et pour tout choix de $w^k \neq 0$, le paramètre optimal r^k défini par (4.16) a pour expression

$$r^k = -\frac{(g^k, w^k)}{(Aw^k, w^k)}$$

et on a les deux relations suivantes

$$\forall k \in I\!N, \qquad g^{k+1} = g^k + r^k Aw^k \tag{4.17}$$

$$\forall k \in I\!N, \qquad (w^k, g^{k+1}) = 0 \tag{4.18}$$

où $g^k = Ax^k - b$ est le vecteur résidu à l'itération k où encore le gradient de J au point x^k.

Démonstation :

L'expression de r^k découle immédiatement de la proposition 4.2. Par ailleurs, on a

$$g^{k+1} = Ax^{k+1} - b = A\left(x^k + r^k w^k\right) - b = g^k + r^k Aw^k$$

$$(w^k, g^{k+1}) = (w^k, g^k) + r^k (Aw^k, w^k) = (w^k, g^k) - \frac{(g^k, w^k)}{(Aw^k, w^k)}(Aw^k, w^k) = 0$$

■

Remarque :

En fait, dans la méthode du gradient conjugué, on suppose déjà calculés les vecteurs $x^1, x^2, \ldots, x^{k-1}$ et $w^1, w^2, \ldots, w^{k-1}$. Alors si $w^{k-1} \neq 0$, on cherche w^k dans le plan formé par les deux directions orthogonales g^k et w^{k-1}, en posant

$$w^k = g^k + \alpha^k w^{k-1} \tag{4.19}$$

On a alors d'après (4.15) :

$$0 = (w^k, Aw^{k-1}) = (g^k + \alpha^k w^{k-1}, Aw^{k-1}) \tag{4.20}$$

d'où l'on déduit que

$$\alpha^k = -\frac{(g^k, Aw^{k-1})}{(Aw^{k-1}, w^{k-1})} \tag{4.21}$$

4.3.1 Description de l'algorithme du gradient conjugué

On initialise l'algorithme en choisissant un vecteur x^0. Si $g^0 \equiv Ax^0 - b = 0$, on arrête les calculs. Si $g^0 \neq 0$, on choisit $w^0 = g^0$ et on effectue les calculs suivants :

$$\left\{\begin{array}{l} x^0 \text{donné} \\ w^0 = g^0 = Ax^0 - b \\ r^0 = -\dfrac{\|g^0\|_2^2}{(Ag^0, g^0)} \\ x^1 = x^0 + r^0 g^0 \end{array}\right. \tag{4.22}$$

Pour $k \geq 1$, on prend

$$\left\{\begin{array}{l} g^k = Ax^k - b \\ \alpha^k = -\dfrac{(Ag^k, w^{k-1})}{(Aw^{k-1}, w^{k-1})} \\ w^k = g^k + \alpha^k w^{k-1} \\ r^k = -\dfrac{(g^k, w^k)}{(Aw^k, w^k)} \\ x^{k+1} = x^k + r^k w^k \end{array}\right. \tag{4.23}$$

Les dénominateurs des expressions apparaissant dans (4.23) sont non-nuls comme le montre le lemme suivant.

Lemme 4.1

Soit $p \in \{1, 2, .., n\}$, on suppose que g^0, g^1 ,.....,g^{p-1} sont tous non nuls. Alors $w^0, w^1, ..., w^{p-1}$ sont non nuls et les quantités :

$$r^j = -\frac{(g^j, w^j)}{(Aw^j, w^j)} \qquad pour \quad j \in \{0, .., p-1\} \tag{4.24}$$

$$\alpha^k = -\frac{(Ag^k, w^{k-1})}{(Aw^{k-1}, w^{k-1})} \qquad pour \quad k \in \{1, .., p\} \tag{4.25}$$

sont bien définies et, on a $r^k \neq 0$ pour tout $k \in \{0, .., p-1\}$.

Démonstration :

Pour montrer que les quantités (4.24) et (4.25) sont bien définies, il suffit de montrer que $w^j \neq 0$ pour tout $j \in \{0, .., p-1\}$. En effet A étant définie positive cela entraine que

$$(Aw^j, w^j) \neq 0 \text{ pour} \quad j \in \{0, .., p-1\}$$

et donc les quantités r^j et α^k sont bien définies pour $j \in \{0, .., p-1\}$ et $k \in \{1, .., p\}$. Pour montrer que $w^j \neq 0$ pour tout $j \in \{0, .., p-1\}$, on procède par récurrence sur j :
Pour $j = 0$

$$w^0 = g^0 \neq 0$$

Supposons que $w^0, w^1, ..., w^{j-1}$ sont non nuls pour $j \in \{1, .., p-1\}$, alors α^j est bien définie et on a

$$w^j = g^j + \alpha^j w^{j-1} \qquad \text{(d'après (4.19))}$$

$$(g^j, w^{j-1}) = 0 \qquad \text{(d'après la proposition 4.3)}$$

Par conséquent, on a

$$\|w^j\|_2^2 = (g^j + \alpha^j w^{j-1}, g^j + \alpha^j w^{j-1}) = \|g^j\|_2^2 + \left(\alpha^j\right)^2 \|w^{j-1}\|_2^2$$

et puisque par hypothèse $g^j \neq 0$ pour tout $j \in \{0, \ldots, p-1\}$, on a $w^j \neq 0$ pour tout $j \in \{0, .., p-1\}$. On en déduit que α^k est bien définie pour tout $k \in \{1, \ldots, p\}$ et par conséquent on obtient grâce à la proposition 4.3 :

$$r^k = -\frac{(g^k, w^k)}{(Aw^k, w^k)} = -\frac{(g^k, g^k + \alpha^k w^{k-1})}{(Aw^k, w^k)} = -\frac{\|g^k\|_2^2}{(Aw^k, w^k)} \neq 0$$

pour tout $k \in \{1, .., p-1\}$. ■

Lemme 4.2

Soit $p \in \{1, 2, .., n\}$, on suppose que g^0, g^1 ,....,g^{p-1} sont tous non nuls. Alors on a pour tout $k \in \{1, 2, .., p\}$:

(4.26) $$(g^k, w^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k-1\}$$

(4.27) $$(g^k, g^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k-1\}$$

(4.28) $$(w^k, Aw^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k-1\}$$

Démonstration :

On va procéder par récurrence sur k. Pour $k = 1$, on a

$$(g^1, w^0) = 0 \qquad \text{(d'après (4.18))}$$

$$(g^1, g^0) = (g^1, w^0) = 0$$

$$(w^1, Aw^0) = 0 \qquad \text{(d'après (4.15))}$$

Supposons que la propriété reste vraie jusqu'à l'ordre $k \leq p-1$ et montrons qu'elle est vraie pour l'ordre $k+1$, c'est-à-dire que

(4.29) $$(g^{k+1}, w^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k\}$$

(4.30) $$(g^{k+1}, g^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k\}$$

(4.31) $$(w^{k+1}, Aw^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k\}$$

* Si $j = k$, on a

$$(g^{k+1}, w^k) = 0 \qquad \text{(d'après (4.18))}$$

* Si $j \leq k-1$, on a d'après (4.17), le lemme 4.1 et l'hypothèse de récurrence :

$$(g^{k+1}, w^j) = (g^k + r^k Aw^k, w^j) = (g^k, w^j) + r^k (Aw^k, w^j) = 0$$

Ainsi (4.29) est prouvée.

Par ailleurs, on a d'après (4.19),(4.29) et le lemme 4.1, que pour tout $j \in \{1, .., k\}$

$$\begin{aligned} (g^{k+1}, g^j) &= (g^{k+1}, w^j - \alpha^j w^{j-1}) \\ &= (g^{k+1}, w^j) - \alpha^j (g^{k+1}, w^{j-1}) \\ (g^{k+1}, g^0) &= (g^{k+1}, w^0) = 0 \end{aligned}$$

D'où, en utilisant (4.29) qui est déjà prouvée :

$$(g^{k+1}, g^j) = 0 \qquad \textit{pour} \quad j \in \{0, 1, .., k\}$$

Il reste à examiner (4.31).

* Si $j = k$, on a

$$(w^{k+1}, Aw^k) = 0 \qquad \text{(d'après (4.15))}$$

* Si $j \leq k-1$, on a

$$\begin{array}{lll}(w^{k+1}, Aw^j) & = (g^{k+1} + \alpha^{k+1} w^k, Aw^j) & \text{(d'après (4.19) et le lemme 4.1)} \\ & = (g^{k+1}, Aw^j) + \alpha^{k+1}(w^k, Aw^j) & \\ & = (g^{k+1}, Aw^j) & \text{(d'après l'hypothèse de récurrence)} \\ & = (g^{k+1}, \dfrac{g^{j+1} - g^j}{r^j}) & \text{(d'après (4.17))}\end{array}$$

puisque $r^j \neq 0$ d'après le lemme 4.1. D'où finalement, d'après (4.30) qui est déjà démontrée :

$$(w^{k+1}, Aw^j) = \frac{1}{r^j}\left[(g^{k+1}, g^{j+1}) - (g^{k+1}, g^j)\right] = 0$$

■

Théorème 4.3

L'algorithme du gradient conjugué converge en au plus n itérations. Autrement dit pour tout choix de x^0, si $Ax^0 - b \neq 0$, alors

$$\exists\ k \in \{1, 2, .., n\} \qquad \textit{tel que} \qquad Ax^k = b.$$

Démonstration :

Si $Ax^0 - b = 0$, on n'a rien à démontrer. Supposons donc que $Ax^0 - b \neq 0$.
1^{er}_{-} cas : Il existe $k_0 \in \{1, 2, .., n-1\}$ tel que $g^{k_0} = 0$. Dans ce cas l'algorithme converge visiblement en moins de n itérations.
2^{eme}_{-} cas : Si $g^k \neq 0$ pour tout $k \in \{1, 2, ..., n-1\}$, on a d'après (4.19), (4.26) et le lemme 4.1 que :

$$\left\|w^k\right\|_2^2 = (g^k + \alpha^k w^{k-1}, g^k + \alpha^k w^{k-1}) = \left\|g^k\right\|_2^2 + \left(\alpha^k\right)^2 \left\|w^{k-1}\right\|_2^2$$

On en déduit que $w^k \neq 0$ pour tout $k \in \{1, 2, .., n-1\}$. La famille $(w^0, w^1, ..., w^{n-1})$ forme d'après (4.28) une base de $\mathbb{R}^n$, orthogonale pour le produit scalaire :

$$(x, y)_A \equiv (Ax, y)$$

En utilisant le lemme 4.2, on a alors pour tout $j \in \{0, 1, 2, .., n-1\}$:

$$\left(w^n, w^j\right)_A = 0$$

et par conséquent $w^n = 0$. On a d'après (4.19), (4.26) et le lemme 4.1

$$\|w^n\|_2^2 = \|g^n\|_2^2 + (\alpha^n)^2 \left\|w^{n-1}\right\|_2^2$$

D'où puisque $w^n = 0$:

$$g^n = Ax^n - b = 0.$$

L'algorithme converge donc à la n^{ieme}_{-} itération. ■

Proposition 4.4

Pour $k \geq 1$ et si $g^i \neq 0$ pour $i \in \{0, 1, 2, .., k-1\}$, on a la relation suivante

$$\alpha^k = \frac{\left\|g^k\right\|_2^2}{\|g^{k-1}\|_2^2} \tag{4.32}$$

Démonstration :

On a d'après le lemme 4.1 :

$$\alpha^k = -\frac{(Ag^k, w^{k-1})}{(Aw^{k-1}, w^{k-1})}. \tag{4.33}$$

Or, grâce à (4.17), le lemme 4.1 et (4.27) on a

$$(Ag^k, w^{k-1}) = (g^k, Aw^{k-1}) = (g^k, \frac{g^k - g^{k-1}}{r^{k-1}}) = \frac{\left\|g^k\right\|_2^2}{r^{k-1}}. \tag{4.34}$$

Par ailleurs, on a d'après (4.17) et (4.26) et le lemme 4.1 :

$$(Aw^{k-1}, w^{k-1}) = (\frac{g^k - g^{k-1}}{r^{k-1}}, w^{k-1}) = -\frac{(g^{k-1}, w^{k-1})}{r^{k-1}}. \tag{4.35}$$

Mais d'après (4.19) et le lemme 4.1, on a

$$w^0 = g^0$$

et

$$w^{k-1} = g^{k-1} + \alpha^{k-1} w^{k-2} \qquad \text{pour} \quad k \geq 2.$$

D'où en utilisant le lemme 4.1, (4.35) et (4.26) :

$$(Aw^{k-1}, w^{k-1}) = -\frac{1}{r^{k-1}}(g^{k-1}, g^{k-1} + \alpha^{k-1} w^{k-2}) = -\frac{\left\|g^{k-1}\right\|_2^2}{r^{k-1}} \tag{4.36}$$

et par conséquent, en regroupant (4.33), (4.34) et (4.36), on obtient

$$\alpha^k = \frac{\left\|g^k\right\|_2^2}{\|g^{k-1}\|_2^2}.$$

■

Le résultat de la proposition 4.4 et la propriété (4.17) permettent décrire l'algorithme du gradient conjugué (4.22) et (4.23) sous la forme suivante

$$\begin{cases} x^0 \\ w^0 = g^0 = Ax^0 - b \\ r^0 = -\dfrac{\left\|g^0\right\|_2^2}{(Ag^0, g^0)} \\ x^1 = x^0 + r^0 w^0 \\ g^1 = g^0 + r^0 A w^0 \end{cases} \tag{4.37}$$

Pour $k \geq 1$, on prend

$$\begin{cases} \alpha^k = \dfrac{\left\|g^k\right\|_2^2}{\left\|g^{k-1}\right\|_2^2} \\ w^k = g^k + \alpha^k w^{k-1} \\ r^k = -\dfrac{(g^k, w^k)}{(Aw^k, w^k)} \\ x^{k+1} = x^k + r^k w^k \\ g^{k+1} = g^k + r^k A w^k \end{cases} \tag{4.38}$$

4.3.2 Choix d'un test d'arrêt

Le test d'arrêt usuel consiste à s'arrêter à la k^{ieme} itération, si

$$\frac{\left\|g^k\right\|}{\|b\|} < \varepsilon \tag{4.39}$$

pour ε choisi "petit". Ce choix doit tenir compte de la précision de l'ordinateur sur lequel les calculs sont effectués.

Si la relation (4.39) est vérifiée, on a

$$\frac{\left\|x^k - A^{-1}b\right\|}{\|x\|} < \varepsilon \ \mathrm{Cond}(A)$$

où x est la solution de (4.3). En effet comme

$$\left\|x^k - A^{-1}b\right\| = \left\|A^{-1}g^k\right\| \leq \left\|A^{-1}\right\| \ \left\|g^k\right\|$$

où $g^k = Ax^k - b$. Si (4.39) est vérifiée, alors

$$\left\|x^k - x\right\| \leq \varepsilon \ \left\|A^{-1}\right\| \ \|b\| \leq \varepsilon \ \left\|A^{-1}\right\| \ \|A\| \ \|x\| = \varepsilon \ \mathrm{Cond}(A)\,\|x\|$$

On utilise également un autre test qui consiste à arrêter les itérations lorsque

$$\left\|x^{k+1} - x^k\right\| \leq \varepsilon \ \left\|x^k\right\|$$

Ce test présente l'inconvénient dans certains cas d'être vérifié sans que x^k ne soit suffisamment proche de la solution.

4.3.3 Nombre d'opérations

Soit c le nombre maximal de coefficients non nuls par ligne de la matrice A. À chaque itération de l'algorithme du gradient conjugué, le nombre d'opérations est majoré comme l'indique le tableau suivant

calcul de	**mult/div**	**add/soust**	**commentaire**
α^k	$n+1$	$n-1$	$\left\|g^{k-1}\right\|_2^2$ est gardé
w^k	n	n	en mémoire
Aw^k	nc	$n(c-1)$	
r^k	$2n+1$	$2(n-1)$	
x^{k+1}	n	n	
g^{k+1}	n	n	
$\left\|g^{k+1}\right\|_2^2$	n	$n-1$	

Soit au total $(c+7)\ n+2$ multiplications et divisions et $(c+6)\ n-4$ additions et soustractions. Donc un nombre d'opérations voisin de $(2c+13)n$ par itération. Pour un nombre d'itérations de l'ordre de n, on aboutit à un nombre d'opérations voisin de $(2c+13)n^2$, ce qui est relativement important lorsque c est grand (si $c = n$, on obtient $2n^3$ opérations, alors que la méthode de Cholesky n'en nécessite que $\dfrac{n^3}{3}$ opérations)

En fait, La méthode du gradient conjugué est l'une des mieux adaptées à la résolution de systèmes linéaires dont la matrice est symétrique, définie positive et creuse. En outre grâce aux techniques de préconditionnement du système linéaire (4.3), le nombre d'opérations peut être sérieusement plus réduit.

4.3.4 Préconditionnement

L'une des techniques de préconditionnement du système linéaire (4.3) consiste à remplacer la résolution de celui-ci par celle du système équivalent

$$\left\{ \begin{array}{ll} L^{-T}AL^{-1}y & = L^{-T}b \\ Lx & = y \end{array} \right. \tag{4.40}$$

où L est une matrice inversible telle que $L^{-T}AL^{-1}$ soit bien conditionnée, c'est-à-dire que

$$\mathrm{Cond}(L^{-T}AL^{-1}) \simeq 1 \tag{4.41}$$

On peut facilement vérifier que la matrice $L^{-T}AL^{-1}$ est symétrique définie positive dès que A l'est et que L est inversible.

La méthode du gradient conjugué appliquée à la résolution du système linéaire

$$L^{-T}AL^{-1}y = L^{-T}b \tag{4.42}$$

peut être interprétée comme étant une méthode de minimisation de la fonction $F : \mathbb{R}^n \to \mathbb{R}$ définie par

$$F(y) \equiv \frac{1}{2}(L^{-T}AL^{-1}y, y) - (L^{-T}b, y). \tag{4.43}$$

Le gradient de la fonction F est donné, comme on peut le vérifier à titre d'exercice par

$$\nabla F(y) = L^{-T}AL^{-1}y - L^{-T}b. \tag{4.44}$$

En introduisant la notation $G^k \equiv \nabla F(y^k)$ et en désignant par W^k la direction de descente numéro k, l'itération courante de l'algorithme du gradient conjugué appliquée à la résolution de (4.42) s'écrit alors, d'après (4.38) comme suit :

$$\left\{ \begin{array}{l} \beta^k = \dfrac{\left\|G^k\right\|_2^2}{\|G^{k-1}\|_2^2} \\ \\ W^k = G^k + \beta^k W^{k-1} \\ \\ R^k = -\dfrac{(G^k, W^k)}{(L^{-T}AL^{-1}W^k, W^k)} \\ \\ y^{k+1} = y^k + R^k W^k \\ \\ G^{k+1} = G^k + R^k L^{-T}AL^{-1}W^k \end{array} \right. \tag{4.45}$$

En revenant à l'inconnue initiale $x = L^{-1}y$ et en adoptant les notations

$$\left\{ \begin{array}{l} x^k = L^{-1}y^k \\ w^k = W^k \\ \alpha^k = \beta^k \\ r^k = R^k \\ g^k = Ax^k - b \end{array} \right. \tag{4.46}$$

il est clair que l'on a

$$G^k = L^{-T}Ax^k - L^{-T}b = L^{-T}g^k \tag{4.47}$$

et l'on peut écrire (4.45) comme suit

$$
\left\{
\begin{array}{l}
\alpha^k = \dfrac{\left\|L^{-T}g^k\right\|_2^2}{\left\|L^{-T}g^{k-1}\right\|_2^2} \\
w^k = L^{-T}g^k + \alpha^k w^{k-1} \\
r^k = -\dfrac{(L^{-T}g^k, w^k)}{(AL^{-1}w^k, L^{-1}w^k)} \\
x^{k+1} = x^k + r^k L^{-1}w^k \\
g^{k+1} = g^k + r^k AL^{-1}w^k
\end{array}
\right. \tag{4.48}
$$

L'algorithme dont l'itération courante est ainsi définie s'avère plus efficace que celui défini par (4.37) et (4.38) lorsque (4.41) est vérifiée, c'est-à-dire lorsque $L^{-T}AL^{-1}$ est bien conditionnée.

Remarques :

1) Soit $A = BB^T$ la factorisation Cholesky de A, en choisissant

$$L = B^T$$

on a alors

$$L^{-T}AL^{-1} = B^{-1}AB^{-T} = I$$

d'où

$$\mathrm{Cond}(L^{-T}AL^{-1}) = 1$$

et l'algorithme du gradient conjugué appliqué à $L^{-T}AL^{-1}$ converge en une itération, comme on peut le vérifier à titre d'exercice.

2) Dans la pratique, lorsque la matrice A est creuse et n est grand, la factorisation de Cholesky est coûteuse en stockage. On peut alors choisir comme matrice de préconditionnement une matrice L voisine de B^T en effectuant ce que l'on appelle une décomposition de Cholesky incomplète qui consiste à ne calculer dans L que les termes correspondants aux emplacements des termes non-nuls de la matrice A.

Les techniques de préconditionnement font encore l'objet de travaux de recherche en cours de réalisation. ■