M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

P. JOLY

# Présentation de synthèse des méthodes de gradient conjugué

*M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique*, tome 20, nº 4 (1986), p. 639-665

<http://www.numdam.org/item?id=M2AN_1986__20_4_639_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 20, n° 4, 1986, p 639 à 665)

# PRÉSENTATION DE SYNTHÈSE DES MÉTHODES DE GRADIENT CONJUGUÉ (*)

par P. JOLY [1]

Communiqué par P G CIARLET

Résumé — *Trois formulations d'algorithmes de gradient conjugué sont présentées. La plupart des algorithmes connus s'en déduisent comme cas particuliers L'étude des propriétés de la méthode générale permet de comparer leur vitesse de convergence*

Abstract. — *General formulations of conjugate gradient methods are introduced, which contain most of the wellknown algorithms Their properties are deduced from the study of the general formulation*

## HYPOTHÈSES ET NOTATIONS

Soit à résoudre le système linéaire $A.x = b$, où $A \in R^{n \times n}$ est une matrice régulière, $x$ et $b \in R^n$.

On note ${}^sA = \frac{1}{2}(A + {}^tA)$ la partie symétrique de $A$

et ${}^{as}A = \frac{1}{2}(A - {}^tA)$ la partie antisymétrique de $A$.

Soient $H$ et $K \in R^{n \times n}$ deux matrices dont la partie symétrique est définie positive, on introduit

$N = {}^tA\, {}^sH\, A$ matrice symétrique définie positive

et $M = {}^tL\, N\, L$ matrice symétrique définie positive

où ${}^sK = L\, {}^tL$ est la factorisation de Cholesky de la partie symétrique de $K$.

Enfin pour tout $r \in R^n$, on définit $E(r) = (r, Hr) = (r, {}^sHr)$.

$E$ est une fonction strictement convexe puisque ${}^sH$ est définie positive. $(.,.)$ désigne le produit scalaire usuel de $R^n$.

(*) Reçu en janvier 1985

[1] Université Pierre et Marie Curie, Laboratoire d'Analyse Numérique, Tour 55-65, 5e étage, 4, place Jussieu, 75230 Paris Cedex 05

L'objet de ce travail est l'étude d'algorithmes de minimisation sur $R^n$ de la fonction $r \to E(r)$, qui admet un minimum unique 0, pour $r = 0$. La convergence de ces algorithmes est liée aux propriétés des matrices $A$, $H$ et $K$.

**PREMIÈRE PARTIE : UN ALGORITHME D'ORTHOGONALISATION**

## I. Définition de l'algorithme d'orthogonalisation

On considère l'algorithme suivant

*Initialisation*

$$\begin{aligned} x^0 &\in \mathbb{R}^n \text{ quelconque} \\ r^0 &= b - Ax^0 = A(x - x^0) \\ g^0 &= {}^tA\,{}^sHr^0 = N(x - x^0) \\ p^0 &= Kg^0 . \end{aligned} \tag{A.1}$$

*Itérations*

$$\begin{aligned} \alpha^k &= (g^k, p^k)/(p^k, Np^k) \\ x^{k+1} &= x^k + \alpha^k p^k \\ g^{k+1} &= g^k - \alpha^k Np^k = {}^tA\,{}^sH\,r^{k+1} \\ p^{k+1} &= Kg^{k+1} + \sum_{l=0}^{k} \beta_l^{k+1} p^l \end{aligned}$$

*avec* $$\beta_l^{k+1} = -(Kg^{k+1}, Np^l)/(p^l, Np^l) \quad 0 \leqslant l \leqslant k .$$

## II. Propriétés de l'algorithme d'orthogonalisation

LEMME 1 :

1) $(p^k, Np^l) = 0 \qquad \forall k \neq l$
2) $(g^k, p^l) = 0 \qquad 0 \leqslant l \leqslant k$
3) $(g^l, p^k) = (g^0, p^k) \qquad 0 \leqslant l < k$
4) $(g^k, p^k) = (g^k, Kg^k)$
5) $(Kg^k, Np^k) = (p^k, Np^k)$
6) $(g^k, Kg^l) = 0 \qquad 0 \leqslant l < k$
7) $\exists k \leqslant n, g^k = 0$

*Démonstration* : les propriétés sont établies par récurrence

1) $$(p^1, Np^0) = (Kg^1, Np^0) + \beta_0^1(p^0, Np^0) = 0$$

supposons la propriété vraie jusqu'à l'ordre $k - 1$, alors

$$\begin{aligned}(p^k, Np^l) &= (Kg^k, Np^l) + \sum_{i=0}^{k-1} \beta_i^k(p^i, Np^l) \qquad 0 \leqslant l < k \\ &= (Kg^k, Np^l) + \beta_l^k(p^l, Np^l) \qquad 0 \leqslant l < k \\ &= 0 \text{ par définition de } \beta_l^k\end{aligned}$$

on en déduit 1) en utilisant la symétrie de $N$.

2) $$g^k = g^l - \sum_{i=l}^{k-1} \alpha^i Np^i \qquad 0 \leqslant l < k$$

donc $$\begin{aligned}(g^k, p^l) &= (g^l, p^l) - \alpha^l(p^l, Np^l) \qquad 0 \leqslant l < k \\ &= 0 \text{ par définition de } \alpha^l\end{aligned}$$

3) $k$ étant fixé, la propriété est vraie pour $l = 0$, on la suppose vraie jusqu'à l'ordre $l$, alors

$$(g^{l+1}, p^k) = (g^l, p^k) - \alpha^l(Np^l, p^k) = (g^0, p^k)$$

4) $$(g^k, p^k) = (g^k, Kg^k) + \sum_{l=0}^{k-1} \beta_l^k(g^k, p^l) = (g^k, Kg^k)$$

5) $$(p^k, Np^k) = (Kg^k, Np^k) + \sum_{l=0}^{k-1} \beta_l^k(p^l, Np^k) = (Kg^k, Np^k)$$

6) $$\begin{aligned}(g^k, Kg^l) &= \left(g^k, p^l - \sum_{i=0}^{l-1} \beta_i^l p^i\right) \\ &= 0 \quad \text{pour} \quad 0 \leqslant l < k \quad \text{d'après (2)}\end{aligned}$$

7) $$\text{soit } G^k = \{ g \in \mathbb{R}^n, \quad \forall l \quad 0 \leqslant l \leqslant k \quad (g, Kg^l) = 0 \}$$

Par construction de la suite $(g^k)_{k \in \mathbb{N}}$ $g^{k+1} \in G^k$, donc $G^k$ n'est pas réduit à $\{ 0 \}$ (sinon $g^{k+1} = 0$ et c'est terminé).

Par définition $G^{k+1} \subset G^k$.

Supposons qu'il existe $k$, $0 \leqslant k < n$ tel que $G^{k+1} = G^k$ alors d'après ce qui précède $(g^{k+1}, Kg^{k+1}) = 0$, soit $g^{k+1} = 0$ puisque la partie symétrique de $K$ est supposée définie positive.

S'il n'existe pas de $k$, $0 \leqslant k < n$ tel que $G^{k+1} = G^k$, alors on a la suite d'inclusions strictes :

$$G^{n-1} \underset{\neq}{\subset} G^{n-2} \dots \underset{\neq}{\subset} G^1 \underset{\neq}{\subset} G^0$$

soit $\quad \dim G^{n-1} < \dim G^{n-2} \cdots < \dim G^1 < \dim G^0 = n - 1\,.$

D'où $\dim G^{n-1} = 0$ ; on en déduit $g^n = 0$.

L'algorithme converge donc en au plus $n$ itérations. □

THÉORÈME 1 : *L'algorithme* (A.1) *converge en au plus n itérations, vers la solution du système* $Ax = b$.

*Remarque* 1 : Des relations $(Kg^k, Np^k) = (p^k, Np^k)$
et $\quad (g^k, p^k) = (g^k, Kg^k)$

on déduit que l'algorithme ne peut dégénérer.

En effet si $(p^k, Np^k) = 0$, alors $p^k = 0$ ce qui entraîne
$(g^k, Kg^k) = 0$ et $\quad g^k = 0$.

Donc on peut toujours calculer les coefficients $\alpha^k$ et $\beta_l^{k+1}$.

De même si $\alpha^k = 0$, alors $(g^k, p^k) = (g^k, Kg^k) = 0$, soit $g^k = 0$. L'algorithme ne s'arrête que s'il a convergé !

*Remarque* 2 : Si ${}^sK$ n'est pas définie positive, l'algorithme (A.1) peut être modifié pour minimiser $E$ sur l'orthogonal dans $\mathbb{R}^n$ de l'ensemble des vecteurs $g \neq 0$, tels que $(g, Kg) = 0$ (voir K. Ito [11]).

*Remarque* 3 : Le signe de $(p^k, Np^k)$ (respectivement $(g^k, Kg^k)$) ne joue aucun rôle dans la démonstration de convergence, il suffit que $N$ (respectivement $K$) soit définie. Ces hypothèses seront utilisées plus loin, pour l'étude de la vitesse de convergence.

*Remarque* 4 : L'algorithme (A.1) est une généralisation de l'algorithme présenté par Eisenstat *et al.* [6], qui correspond au cas $H = I$ et $K = {}^tA^{-1}$ (en supposant ${}^sA$ définie positive).

LEMME 2 : *Si on suppose la matrice K symétrique, alors*

8) $\quad (g^k, Kg^l) = 0 \qquad \forall k \neq l$

9) $\quad (Kg^{k+1}, Np^l) = 0 \qquad 0 \leqslant l < k$

10) $\quad (Kg^{k+1}, Np^k) = -(g^{k+1}, Kg^{k+1})/\alpha^k$

11) $\quad \beta_l^{k+1} = 0 \qquad 0 \leqslant l < k$

$$\beta_k^{k+1} = (g^{k+1}, Kg^{k+1})/(g^k, Kg^k)$$

*Démonstration* :

8) D'après 6) en utilisant la symétrie de $K$.

9) $$(Kg^{k+1}, Np^l) = \left(Kg^{k+1}, \frac{1}{\alpha^l}(g^l - g^{l+1})\right)$$
$$= 0 \quad \text{pour} \quad 0 \leqslant l < k$$

10) $$(Kg^{k+1}, Np^k) = -(g^{k+1}, Kg^{k+1})/\alpha^k$$

11) $$\beta_l^{k+1} = -(Kg^{k+1}, Np^l)/(p^l, Np^l)$$
$$= 0 \quad \text{pour} \quad 0 \leqslant l < k$$

et $$\beta_k^{k+1} = \frac{(g^{k+1}, Kg^{k+1})}{(p^k, Np^k)} \times \frac{(p^k, Np^k)}{(g^k, Kg^k)}.$$

□

Si la matrice $K$ est symétrique, l'algorithme (A.1) s'écrit donc :

*Initialisation* :

$$\begin{aligned} x^0 &\in \mathbb{R}^n \textit{ quelconque} \\ r^0 &= b - Ax^0 = A(x - x^0) \\ g^0 &= {}^tA\,{}^sHr^0 = N(x - x^0) \\ p^0 &= Kg^0 . \end{aligned} \tag{A.2}$$

*Itérations* :

$$\begin{aligned} \alpha^k &= (g^k, Kg^k)/(p^k, Np^k) \\ x^{k+1} &= x^k + \alpha^k p^k \\ g^{k+1} &= g^k - \alpha^k Np^k = {}^tA\,{}^sHr^{k+1} \\ \beta^{k+1} &= (g^{k+1}, Kg^{k+1})/(g^k, Kg^k) \\ p^{k+1} &= Kg^{k+1} + \beta^{k+1} p^k . \end{aligned}$$

On reconnaît sous cette forme l'algorithme du gradient conjugué généralisé de Daniel [4], et la méthode des directions $M1$-$M2$ conjuguées de Il'In [10].

## III. Vitesse de convergence

On pose $$E^k = E(r^k) = (b - Ax^k, {}^sH(b - Ax^k))$$

et on note $\langle p^0, p^1, ..., p^k \rangle$ l'espace vectoriel engendré par $p^0, p^1, ..., p^k$.

LEMME 3 : *$x^{k+1}$ réalise le minimum de $E^{k+1}$ sur l'espace affine*

$$x^0 + \langle p^0, p^1, ..., p^k \rangle .$$

*Démonstration* : Il suffit de montrer que $x^{l+1}$ réalise le minimum de $E^{l+1}$ dans la direction $p^l$, pour $0 \leqslant l \leqslant k$.

$$\begin{aligned} x(\alpha) &= x^l + \alpha p^l \\ r(\alpha) &= r^l - \alpha A p^l \\ E^{l+1} &= (r^l, Hr^l) - 2\,\alpha(g^l, p^l) + \alpha^2(p^l, Np^l) \end{aligned}$$

la valeur $\alpha^l = (g^l, p^l)/(p^l, Np^l)$ réalise le minimum de $E^{l+1}$ dans la direction $p^l$, et l'on pose $x^{l+1} = x(\alpha^l)$, $r^{l+1} = r(\alpha^l)$.

De manière générale

$$E^{k+1} = E^0 - \sum_{l=0}^{k} 2\,\alpha^l(g^l, p^l) + \sum_{l=0}^{k} (\alpha^l)^2(p^l, Np^l) . \qquad \square$$

LEMME 4 : $E^{k+1} = E^k - (g^k, Kg^k)^2/(p^k, Np^k)$.

*Démonstration* : en utilisant la définition de $\alpha^k$ et le lemme 1.

On en déduit que la convergence de l'algorithme (A.1) est monotone, le cas $E^{k+1} = E^k$ entraînant $(g^k, Kg^k) = 0$, soit $g^k = 0$.

LEMME 5 : $(p^k, Np^k) \leqslant (Kg^k, NKg^k)$.

*Démonstration* :

$$p^k = Kg^k + \sum_{l=0}^{k-1} \beta_l^k\, p^l$$

soit $(p^k, Np^k) = (Kg^k, NKg^k) + 2\sum_{l=0}^{k-1} \beta_l^k(Kg^k, Np^l) + \sum_{l=0}^{k-1} (\beta_l^k)^2\,(p^l, Np^l)$

en utilisant la définition de $\beta_l^k$, on obtient

$$(p^k, Np^k) = (Kg^k, NKg^k) - \sum_{l=0}^{k-1} (Kg^k, Np^l)^2/(p^l, Np^l) . \qquad \square$$

THÉORÈME 2 :

2.1) $$E^k \leqslant E^0\left(1 - \frac{\lambda_{\min}({}^tL^s(K^{-1})\,L)}{\mathrm{cond}\,(M)}\right)^k.$$

2.2) *Si la matrice K est symétrique*

$$E^k \leqslant E^0\left(\frac{\mathrm{cond}\,(M) - 1}{\mathrm{cond}\,(M) + 1}\right)^{2k}$$

*avec* $$\mathrm{cond}\,(M) = \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}.$$

*Démonstration* : D'après le lemme 4

$$E^{k+1} = E^k - (g^k, Kg^k)^2/(p^k, Np^k)$$

d'autre part

$$E^k = (r^k, Hr^k) = (g^k, N^{-1}\,g^k)$$

d'où

$$\frac{E^{k+1}}{E^k} = 1 - \frac{(g^k, Kg^k)}{(g^k, N^{-1}\,g^k)} \times \frac{(g^k, Kg^k)}{(p^k, Np^k)}$$

soit encore d'après le lemme 5

$$\frac{E^{k+1}}{E^k} \leqslant 1 - \frac{(g^k, Kg^k)}{(g^k, N^{-1}\,g^k)} \times \frac{(g^k, Kg^k)}{(Kg^k, NKg^k)}.$$

1) *K n'est pas symétrique*

$$K = L.{}^tL + {}^{as}K \quad \text{avec} \quad {}^sK = L.{}^tL \quad \text{définie positive}$$

alors $$\frac{(g^k, Kg^k)}{(g^k, N^{-1}\,g^k)} = \frac{(h, h)}{(h, M^{-1}\,h)} \geqslant \lambda_{\min}(M)$$

où $M = {}^tLNL$ est une matrice symétrique définie positive
et $h = {}^tLg^k$

$\lambda_{\min}(M)$ est la plus petite valeur propre de $M$.

Soit $e \in \mathbb{R}^n$, tel que $Le = Kg^k = Lh + {}^{as}Kg^k$

alors $$(Kg^k, NKg^k) = (e, Me)$$
$$(g^k, Kg^k) = (e, {}^tL^s(K^{-1})\,Le)$$

donc

$$\frac{(g^k, Kg^k)}{(Kg^k, NKg^k)} = \frac{(e, e)}{(e, Me)} \times \frac{(e, {}^tL^s(K^{-1})\, Le)}{(e, e)} \geqslant \frac{\lambda_{\min}({}^tL^s(K^{-1})\, L)}{\lambda_{\max}(M)}$$

en déduit

$$\frac{E^{k+1}}{E^k} \leqslant 1 - \frac{\lambda_{\min}(M)}{\lambda_{\max}(M)} \times \lambda_{\min}({}^tL^s(K^{-1})\, L)$$

$\lambda_{\max}(M)$ est la plus grande valeur propre de $M$.

2) *K est symétrique*

$$K = L.{}^tL$$

alors
$$\frac{(g^k, Kg^k)}{(g^k, N^{-1}\, g^k)} = \frac{(h, h)}{(h, M^{-1}\, h)}$$

et
$$\frac{(g^k, Kg^k)}{(Kg^k, NKg^k)} = \frac{(h, h)}{(h, Mh)}$$

avec
$$h = {}^tLg^k$$

et $M = {}^tLNL$ matrice symétrique définie positive en utilisant l'inégalité de KANTOROVITCH [8], on obtient

$$\frac{(g^k, Kg^k)}{(g^k, N^{-1}\, g^k)} \times \frac{(g^k, Kg^k)}{(Kg^k, NKg^k)} \geqslant 4\, \frac{\lambda_{\min}(M)\, \lambda_{\max}(M)}{(\lambda_{\min}(M) + \lambda_{\max}(M))^2}$$

soit encore

$$\frac{E^{k+1}}{E^k} \leqslant 1 - 4\, \frac{\lambda_{\min}(M)\, \lambda_{\max}(M)}{(\lambda_{\min}(M) + \lambda_{\max}(M))^2} = \left[\frac{\mathrm{cond}\,(M) - 1}{\mathrm{cond}\,(M) + 1}\right]^2$$

avec
$$\mathrm{cond}\,(M) = \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}. \qquad \square$$

*Remarque* 5 : De l'identité ${}^s(K^{-1}) = [{}^sK + {}^{tas}K({}^sK)^{-1}\, {}^{as}K]^{-1}$ (voir Eisenstat *et al.* [6]) on tire

$${}^tL\, {}^s(K^{-1})\, L = [I + L^{-1}\, {}^{tas}K({}^sK)^{-1}\, {}^{as}K^t\, L^{-1}]^{-1}$$

d'où $\lambda_{\min}({}^tL\, {}^s(K^{-1})\, L) = 1/(1 + \lambda_{\max}(S))$ avec $S = L^{-1}\, {}^{tas}K({}^sK^{-1})\, {}^{as}K\, {}^tL^{-1}$ matrice symétrique définie positive.

Donc $\lambda_{\min}({}^tL\, {}^s(K^{-1})\, L) < 1$.

*Remarque* 6 : Si la matrice $K$ est symétrique, la majoration 2.2 peut être améliorée suivant :

$$E^k \leqslant 4\, E^0 \left[\frac{\sqrt{\operatorname{cond}(M)} - 1}{\sqrt{\operatorname{cond}(M)} + 1}\right]^{2k}$$

en utilisant les propriétés des polynômes de Tchebycheff (voir par exemple G. H. Golub, G. Meurant [8]).

## IV. Optimalité de l'algorithme (A.1)

LEMME 6 :

$$\langle p^0, p^1, \ldots, p^k \rangle = \langle Kg^0, Kg^1, \ldots, Kg^k \rangle = \langle p^0, (KN)\, p^0, \ldots, (KN)^k\, p^0 \rangle .$$

*Démonstration* :

1) $\langle p^0, p^1, \ldots, p^k \rangle = \langle Kg^0, Kg^1, \ldots, Kg^k \rangle$.

On établit la propriété par récurrence :

— elle est vraie pour $k = 0$, puisque $p^0 = Kg^0$,

— supposons la vraie pour tout $l$, $0 \leqslant l < k$.

De la relation $p^k = Kg^k + \sum_{l=0}^{k-1} \beta^k\, p^l$, on déduit l'inclusion

$$\langle p^0, p^1, \ldots, p^k \rangle \subset \langle Kg^0, Kg^1, \ldots, Kg^k \rangle .$$

Mais d'après le lemme 1, les vecteurs $p^0, p^1, \ldots, p^k$ sont linéairement indépendants, soit

$$\dim \langle p^0, p^1, \ldots, p^k \rangle = k + 1 \leqslant \dim < Kg^0, Kg^1, \ldots, Kg^k \rangle \leqslant k + 1 .$$

Donc $\langle p^0, p^1, \ldots, p^k \rangle = \langle Kg^0, Kg^1, \ldots, Kg^k \rangle$.

2) $\langle p^0, p^1, \ldots, p^k \rangle = \langle p^0, (KN)\, p^0, \ldots, (KN)^k\, p^0 \rangle$

— la relation est vraie pour $k = 0$,

— supposons la vraie pour tout $l$, $0 \leqslant l < k$,

par définition $p^k = K(g^{k-1} - \alpha^{k-1}\, Np^{k-1}) + \sum_{l=0}^{k-1} \beta_l^k\, p^l$

soit $$p^{k-1} = \left(Kg^{k-1} + \sum_{l=0}^{k-1} \beta_l^k\, p^l\right) - \alpha^{k-1}\, KNp^{k-1}$$

d'après l'hypothèse de récurrence et la première égalité du lemme 6, le premier terme appartient au sous-espace $\langle p^0, (KN)\,p^0, \ldots, (KN)^{k-1}\,p^0 \rangle$. On en déduit l'inclusion

$$\langle p^0, p^1, \ldots, p^k \rangle \subset \langle p^0, (KN)\,p^0, \ldots, (KN)^k\,p^0 \rangle$$

et on conclut comme précédemment. □

*Remarque* 7 : L'étude de la vitesse de convergence montre que l'algorithme idéal correspond au cas $KN = NK = I$ (en supposant $K$ symétrique). Dans ces conditions $\operatorname{cond}(M) = 1$, mais d'après le lemme 6 :

$$\forall k \ \langle p^0, p^1, \ldots, p^k \rangle = \langle p^0, (NK)\,p^0, \ldots, (NK)^k\,p^0 \rangle = \langle p^0 \rangle .$$

La méthode converge donc en 1 itération ! Il s'agit en fait d'une méthode directe, car l'initialisation d'un tel algorithme nécessite le calcul de

$$p^0 = A^{-1}\,r^0 = KNA^{-1}\,r^0 .$$

LEMME 7 : *Si $B \in \mathbb{R}^{n \times n}$ est symétrique définie positive,*

$$\forall y \in \mathbb{R}^n, \quad \forall A \in \mathbb{R}^{n \times n}, \quad (Ay, BAy) \leqslant \operatorname{cond}(B)\, \| A \|_2^2\, (y, By)$$

*où*

$$\| A \|_2 = \sup_{y \neq 0} \frac{\| Ay \|_2}{\| y \|_2} = \sup_{y \neq 0} \frac{(Ay, Ay)^{1/2}}{(y, y)^{1/2}} .$$

*Démonstration* :

$$\frac{(Ay, BAy)}{(y, By)} = \frac{(Ay, BAy)}{(Ay, Ay)} * \frac{(Ay, Ay)}{(y, y)} * \frac{(y, y)}{(y, By)}$$

$$\frac{(Ay, BAy)}{(Ay, Ay)} \leqslant \lambda_{\max}(B)$$

$$\frac{(Ay, Ay)}{(y, y)} \leqslant \| A \|_2^2$$

$$\frac{(y, y)}{(y, By)} \leqslant \frac{1}{\lambda_{\min}(B)}$$

$$\Rightarrow \frac{(Ay, BAy)}{(y, By)} \leqslant \operatorname{cond}(B)\, \| A \|_2^2 .$$

THÉORÈME 3 :

3.1) $$E^k \leqslant E^0 \operatorname{cond}(N^{-1}) \min_{\substack{q \in P_k \\ q(0) = 1}} \| q(NK) \|_2^2 .$$

3.2) *Si la matrice K est symétrique*

$$E^k \leqslant E^0 \operatorname{cond}(M^{-1}) \min_{\substack{q \in P_k \\ q(0)=1}} \| q(M) \|_2^2$$

*où $P_k$ est l'ensemble des polynômes réels de degré inférieur ou égal à K.*

*Démonstration* :

1) D'après le lemme 6, les vecteurs $g^k$ générés par l'algorithme (A.1) sont de la forme $g^k = K^{-1}\, q_k(KN)\, Kg^0 = q_k(NK)\, g^0$ où $q_k$ est un polynôme réel de degré inférieur ou égal à $k$, tel que $q_k(0) = 1$.

Donc $\quad E^k = (g^k, N^{-1} g^k) = (q_k(NK)\, g^0, N^{-1}\, q_k(NK)\, g^0)$

soit d'après le lemme 3

$$E^k = \min_{\substack{q \in P_k \\ q(0)=1}} (q(NK)\, g^0, N^{-1}\, q(NK)\, g^0)$$

et $E^k \leqslant E^0 \operatorname{cond}(N^{-1}) \min_{\substack{q \in P_k \\ q(0)=1}} \| q(NK) \|_2^2$ d'après le lemme 7.

2) Si $K$ est symétrique : $NK = {}^tL^{-1}\, M\, {}^tL$ et $q_k(NK) = {}^tL^{-1}\, q_k(M)\, {}^tL$

soit
$$E^k = \min_{\substack{q \in p_k \\ q(0)=1}} ({}^tL^{-1}\, q(M)\, {}^tLg^0, N^{-1}\, {}^tL^{-1}\, q(M)\, {}^tLg^0)$$

$$E^k = \min_{\substack{q \in p_k \\ q(0)=1}} (q(M)\, {}^tLg^0, M^{-1}\, q(M)\, {}^tLg^0)$$

$$\Rightarrow E^k \leqslant E^0 \operatorname{cond}(M^{-1}) \min_{\substack{q \in p_k \\ q(0)=1}} \| q(M) \|_2^2 . \qquad \square$$

## V. Interprétation de l'algorithme (A.2) comme une méthode de Lanczos

On suppose dans cette partie que la matrice $K$ est symétrique. Par définition des vecteurs $p^k$ et $g^k$ :

$$Kg^k = p^k - \beta^k p^{k-1}$$

$$Np^k = (g^k - g^{k+1})/\alpha^k$$

soit
$$NKg^k = -\frac{\beta^k}{\alpha^{k-1}} g^{k-1} + \left(\frac{1}{\alpha^k} + \frac{\beta^k}{\alpha^{k-1}}\right) g^k - \frac{1}{\alpha^k} g^{k+1}$$

on définit $\quad | g^k | = (g^k, Kg^k)^{1/2}$

et $\qquad G^k = g^k/|g^k|$

alors $\qquad \beta^k = |g^k|^2/|g^{k-1}|^2$

et $\qquad NKG^k = \dfrac{\sqrt{\beta^k}}{\alpha^{k-1}}\,G^{k-1} + \left(\dfrac{1}{\alpha^k} + \dfrac{\beta^k}{\alpha^{k-1}}\right) G^k - \dfrac{\sqrt{\beta^{k+1}}}{\alpha^k}\,G^{k+1}\,.$

Enfin on note $\mathbb{G}, \mathbb{L}, \mathbb{\Delta} \in \mathbb{R}^{nn}$ les matrices définies par

$$\mathbb{G} = [G^1, G^2, \dots, G^n]$$

$$\mathbb{L} = \begin{bmatrix} -1 & & & \\ -\sqrt{\beta^2} & \ddots & & 0 \\ & \ddots & \ddots & \\ 0 & & -\sqrt{\beta^n} & 1 \end{bmatrix}$$

$$\mathbb{\Delta} = \begin{bmatrix} \frac{1}{\alpha^1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\alpha^n} \end{bmatrix}$$

L'algorithme (A.2) se résume alors dans l'équation matricielle

$$NK\mathbb{G} = \mathbb{G}(\mathbb{L}\Delta{}^t\mathbb{L})$$

mais par définition de $\mathbb{G}$ : ${}^t\mathbb{G}K\mathbb{G} = I$
soit en utilisant la factorisation de Cholesky de la matrice $K$ :

$$\boxed{({}^t\mathbb{G}L)\, M({}^tL\mathbb{G}) = \mathbb{L}\mathbb{\Delta}{}^t\mathbb{L}}$$

où $\mathbb{Q} = {}^tL\mathbb{G}$ est une matrice orthogonale : ${}^t\mathbb{Q}\mathbb{Q} = I$
et $M = {}^tLNL$ est une matrice symétrique définie positive.

L'algorithme (A.2) est donc équivalent à la méthode de Lanczos pour le calcul des valeurs propres de la matrice $M$.

*Remarque* 8 : si $K$ n'est pas symétrique, l'algorithme (A.1) peut s'interpréter comme une méthode d'Arnoldi.

## VI. Écriture d'algorithmes classiques sous la forme (A.1)

Nous présentons un tableau résumé des algorithmes classiques obtenus par un choix approprié des matrices $H$ et $K$. La matrice $M$ correspondante détermine leur vitesse de convergence.

| Nom de l'algorithme | Condition suffisante de convergence | $H$ | $K$ | $M$ |
|---|---|---|---|---|
| Gradient conjugué [8] | $A$ symétrique | $A^{-1}$ | $I$ | $A$ |
| Gradient conjugué préconditionné [8] | définie | $A^{-1}$ | $(l\,{}^tl)^{-1}$ | $l^{-1} A\,{}^tl^{-1}$ |
| Résidu conjugué [6] | positive | $I$ | $A^{-1}$ | $A$ |
| Résidu conjugué généralisé [6] | ${}^sA$ définie positive | $I$ | ${}^tA^{-1}$ | ${}^t(AL)(AL)$ |
| Équation normale [8] | $A$ régulière | $I$ | $I$ | ${}^tAA$ |
| Erreur minimale [15] | | $(A\,{}^tA)^{-1}$ | ${}^tAA$ | $A\,{}^tA$ |

Avec $l\,{}^tl$ factorisation de Cholesky incomplète de la matrice $A$ (cas $A$ symétrique définie positive).

$L\,{}^tL$ factorisation de Cholesky de la partie symétrique de $A$ (cas $A$ non symétrique).

Pour d'autres détails voir [14].

*Remarque* 9 : Les algorithmes présentés ici, correspondent aux choix des matrices $H$ et $K$ les plus intéressants ; parmi toutes les méthodes qui s'écrivent sous la forme (A.1), elles sont les moins coûteuses : les bons choix ont déjà été faits !

D'autre part, on peut montrer que l'algorithme de Concus et Golub [3] pour les systèmes non symétriques, ne peut pas s'écrire sous la forme (A.1) (voir aussi Eisenstat [5]).

*Remarque* 10 : *Préconditionnement* : Dans la méthode du gradient conjugué préconditionné, il est naturel de considérer la matrice $K$ comme une matrice de préconditionnement. Cette interprétation n'est plus valable pour les autres méthodes, dans lesquelles la technique de préconditionnement se met en œuvre en substituant au système initial $Ax = b$, le système équivalent $\hat{A}\hat{x} = \hat{b}$, avec $\hat{A} = L^{-1} AU^{-1}$, $\hat{x} = Ux$ et $\hat{b} = L^{-1} b$ ; $LU$ est une factorisation de Gauss incomplète de la matrice $A$. L'algorithme choisi est alors appliqué à la résolution du nouveau système linéaire.

## VII. Une variante de l'algorithme (A.1) : la méthode ORTHOMIN ($m$)

1) *Définition de l'algorithme*

Dans le cas où la matrice $K$ n'est pas symétrique, le calcul des coefficients $\beta_l^k$ est d'autant plus coûteux qu'il nécessite le stockage de tous les vecteurs $p^l$ déjà utilisés !

Vinsome [17] propose une méthode qui ne s'appuie que sur les $m$ ($m \geqslant 1$) dernières directions calculées. Il s'agit de la méthode ORTHOMIN ($m$) qui s'écrit :

*Initialisation*

$$\begin{aligned} & x^0 \in R^n \text{ quelconque} \\ & r^0 = b - Ax^0 \\ & g^0 = {}^tA\,{}^sHr^0 \\ & p^0 = Kg^0 . \end{aligned} \tag{A.3}$$

*Itérations*

$$\begin{aligned} \alpha^k &= (g^k, p^k)/(p^k, Np^k) \\ x^{k+1} &= x^k + \alpha^k p^k \\ g^{k+1} &= g^k - \alpha^k Np^k = {}^tA\,{}^sHr^{k+1} \\ p^{k+1} &= Kg^{k+1} + \sum_{l=k-D+1}^{k} \beta_l^{k+1} p^l \end{aligned}$$

avec $\quad \beta_l^{k+1} = -(Kg^{k+1}, Np^l)/(p^l, Np^l) \quad k - m + 1 \leqslant l \leqslant k .$

Cet algorithme est identique à l'algorithme (A.1), mais l'orthogonalisation n'a lieu que sur les $m$ derniers vecteurs : $p^{k-m+1}, \ldots, p^k$.

2) *Propriétés de l'algorithme*

Les propriétés de cet algorithme sont résumées dans le

LEMME 8 :

12) $\quad (p^k, Np^l) = (p^l, Np^k) = 0 \qquad k - m \leqslant l < k$

13) $\quad (g^k, p^l) = 0 \qquad k - m - 1 \leqslant l < k$

14) $\quad (g^l, p^k) = (g^{k-m}, p^k) \qquad k - m \leqslant l \leqslant k$

15) $\quad (g^k, p^k) = (g^k, Kg^k)$

16) $\quad (Kg^k, Np^k) = (p^k, Np^k)$

17) $\quad (g^k, Kg^l) = 0 \qquad k - m \leqslant l < k$

*Démonstration* : Analogue à celle du lemme 1.

LEMME 9 : *$x^{k+1}$ réalise le minimum de $E^{k+1}$ sur l'espace affine*

$$x^{k-m} + \langle p^{k-m}, \dots, p^k \rangle .$$

*Démonstration* : Analogue à celle du lemme 3.

THÉORÈME 4 :

4.1) $$\lim_{k \to +\infty} r^k = 0 \qquad (\textit{l'algorithme } (\mathrm{A}.3) \textit{ converge})$$

4.2) $$E^k \leqslant E^0 \left(1 - \frac{\lambda_{\min}({}^tL^s(K^{-1})\,L)}{\mathrm{cond}\,(M)}\right)^k$$

4.3) $$E^k \leqslant E^0 \left[\mathrm{cond}\,(N^{-1}) \min_{\substack{q \in P_m \\ q(0)=1}} \| q(NK) \|_2^2 \right]^{k/m} .$$

*Démonstration* :

4.1) D'après le lemme 4 : $E^{k+1} = E^k - (g^k, Kg^k)^2/(p^k, Np^k)$ la suite $(E^k)_{k \in \mathbb{N}}$ est donc telle que l'on a l'alternative :

— Soit $\exists k \in \mathbb{N}$, tel que $E^{k+1} = E^k$, et alors

$$(g^k, Kg^k) = 0 , \quad \text{soit} \quad g^k = 0 \quad \text{et} \quad r^k = 0 .$$

— Soit $\nexists k \in \mathbb{N}$, tel que $E^{k+1} = E^k$.

Dans ce dernier cas, la suite $(E^k)_{k \in \mathbb{N}}$ est strictement décroissante, bornée inférieurement par 0. Elle converge vers une limite finie : $E^\infty \geqslant 0$. On en déduit que

$$\lim_{k \to \infty} (E^{k+1} - E^k) = 0 = \lim_{k \to \infty} \{ (g^k, Kg^k)^2/(p^k, Np^k) \}$$

i.e. $\forall \varepsilon > 0$, $\exists k_0 \in \mathbb{N}$ tel que

$$\forall k > k_0 \qquad (g^k, Kg^k)^2 \leqslant \varepsilon (p^k, Np^k)$$

mais $(g^k, Kg^k) \geqslant \lambda_{\min}({}^sK) \| g^k \|_2^2$

et $(p^k, Np^k) \leqslant (Kg^k, NKg^k) \leqslant \lambda_{\max}(N) \| K \|_2^2 \| g^k \|_2^2 .$

Soit encore : $\forall \varepsilon > 0$, $\exists k_0 \in \mathbb{N}$ tel que

$$\forall k > k_0 \qquad \| g^k \|_2^2 \leqslant \varepsilon \frac{\lambda_{\max}(N)}{\lambda_{\min}^2({}^sK)} \| K \|_2^2$$

ou encore $$\lim_{k \to +\infty} g^k = 0 .$$

4.2) Même démonstration que pour le théorème 2.

4.3) Même raisonnement que pour le théorème 2, on obtient

$$E^k \leqslant E^{k-m}\left[\operatorname{cond}(N^{-1}) \min_{\substack{q \in p_m \\ q(0)=1}} \| q(NK) \|_2^2\right]$$

soit

$$E^k \leqslant E^0\left[\operatorname{cond}(N^{-1}) \min_{\substack{q \in p_m \\ q(0)=1}} \| q(NK) \|_2^2\right]^{k/m}.$$

□

*Remarque* 11 : Au lieu de l'algorithme (A.3), on peut envisager de réinitialiser l'algorithme (A.1) toutes les $m$ itérations. On obtient ainsi un algorithme assez proche de (A.3), la seule différence se trouve dans la définition des vecteurs $p^{jm}$ (pour $j \in \mathbb{N}$) qui se réduit à $p^{jm} = Kg^{jm}$ (soit $\beta_l^{jm} = 0$ pour $(j-1) \times m + 1 \leqslant l \leqslant jm$). On vérifie facilement que le théorème 4 est encore valable pour cet algorithme.

## VIII. Une autre formulation de l'algorithme (A.1)

1) *Définition de l'algorithme*

On considère l'algorithme suivant :

*Initialisation*

$$\begin{aligned} x^0 &\in R^n \textit{ quelconque} \\ r^0 &= b - Ax^0 = A(x - x^0) \\ g^0 &= {}^tA\,{}^sHr^0 = N(x - x^0) \\ p^0 &= Kg^0 . \end{aligned}$$

*Itérations* (A.4)

$$\begin{aligned} \alpha^k &= (g^k, p^k)/(p^k, Np^k) \\ x^{k+1} &= x^k + \alpha^k p^k \\ g^{k+1} &= g^k - \alpha^k Np^k = {}^tA\,{}^sHr^{k+1} \\ p^{k+1} &= KNp^k + \sum_{l=0}^{k} \beta_l^{k+1} p^l \end{aligned}$$

*avec*

$$\beta_l^{k+1} = -(KNp^k, Np^l)/(p^l, Np^l) .$$

Soient

$$\mathcal{E}_1^k = \langle p_1^0, p_1^1, \ldots, p_1^k \rangle$$

le sous-espace des directions générées par l'algorithme (A. 1)

et
$$\mathcal{E}_2^k = \langle p_2^0, p_2^1, \ldots, p_2^k \rangle$$

le sous-espace des directions générées par l'algorithme (A.4).

On va montrer que $\mathcal{E}_1^k = \mathcal{E}_2^k \ \forall k \geqslant 0$, si on prend le même vecteur initial $x^0$ pour les deux algorithmes.

Plus précisément :

LEMME 10 : $\forall k \geqslant 0,\ \exists \gamma^k \neq 0,\ p_1^k + \gamma^k p_2^k = 0.$

*Démonstration* : On procède par récurrence :

- pour $k = 0$ la relation est vraie avec $\gamma^0 = -1$,
- supposons la relation vraie pour tout $l : 0 \leqslant l \leqslant k$.

Alors
$$p_1^{k+1} = Kg_1^{k+1} + \sum_{l=0}^{k} \beta_{l,1}^{k+1} p_1^l$$

avec
$$\beta_{l,1}^{k+1} = -(Kg_1^{k+1}, Np_1^l)/(p_1^l, Np_1^l)$$

soit encore

$$p_1^{k+1} = Kg_1^k - \alpha_1^k KNp_1^k + \sum_{l=0}^{k} \beta_{l,1}^{k+1} p_1^l .$$

D'après le lemme 6, $Kg_1^k \in \mathcal{E}_1^k$, soit

$$Kg_1^k = \sum_{l=0}^{k} \lambda_{l,1}^k p_1^l$$

avec $\lambda_{l,1}^k = (Kg_1^k, Np_1^l)/(p_1^l, Np_1^l)$ par application du lemme 1 .

Ainsi

$$p_1^{k+1} = -\alpha_1^k KNp_1^k + \sum_{l=0}^{k} (\lambda_{l,1}^k + \beta_{l,1}^{k+1}) p_1^l$$

et
$$\lambda_{l,1}^k + \beta_{l,1}^{k+1} = (Kg_1^k - Kg_1^{k+1}, Np_1^l)/(p_1^l, Np_1^l)$$
$$= \alpha_1^k (KNp_1^k, Np_1^l)/(p_1^l, Np_1^l)$$

soit
$$p_1^{k+1} = -\alpha_1^k \left\{ KNp_1^k + \sum_{l=0}^{k} \delta_{l,1}^{k+1} p_1^l \right\}$$

en posant
$$\delta_{l,1}^{k+1} = -(KNp_1^k, Np_1^l)/(p_1^l, Np_1^l) .$$

On utilise maintenant l'hypothèse de récurrence, alors

$$p_1^{k+1} = \alpha_1^k \gamma^k \left\{ KNp_2^k + \sum_{l=0}^{k} \beta_{l,2}^{k+1} p_2^l \right\}$$

avec $\beta_{l,2}^{k+1} = -(KNp_2^k, Np_2^l)/(p_2^l, Np_2^l)$.

Par définition de $p_2^{k+1}$, on a donc

$$p_1^{k+1} = \alpha_1^k \gamma^k p_2^{k+1}.$$

Soit encore $\gamma^{k+1} + \alpha_1^k \gamma^k = 0$. □

*Remarque* 12 : D'une manière générale

$$\forall k \geqslant 0 \qquad \gamma^{k+1} = (-1)^k \prod_{l=0}^{k} \alpha_1^l \qquad (\gamma^0 = -1).$$

THÉORÈME 5 : *Les algorithmes* (A.1) *et* (A.4) *sont équivalents, i.e. à partir du même $x^0$ ils calculent le même $x^k$.*

*Démonstration* : La fonction $E(r^k) = (r^k, Hr^k)$ est strictement convexe, elle admet donc un minimum unique sur $\mathscr{E}_1^k$. Comme $\mathscr{E}_1^k = \mathscr{E}_2^k$ à chaque itération, $x_1^k = x_2^k$. □

En particulier, on note que

$$\forall k \quad \alpha_2^k + \alpha_1^k \gamma^k = 0$$

puisque $g_1^k = g_2^k$

et
$$\alpha_1^k = (g_1^k, p_1^k)/(p_1^k, Np_1^k)$$
$$\alpha_2^k = (g_2^k, p_2^k)/(p_2^k, Np_2^k).$$

Soit $$\boxed{\gamma^{k+1} = \alpha_2^k \quad \forall k \geqslant 0}.$$

En conséquence, toutes les propriétés de l'algorithme (A.1), se transmettent à l'algorithme (A.4), en particulier les résultats du théorème 2 sur la vitesse de convergence.

*Remarque* 13 : En éliminant $g^k$ dans la définition de l'algorithme (A.1) on trouve directement la relation

$$p^{k+1} = -\alpha^k KNp^k + (1 + \beta_k^{k+1}) p^k - \sum_{l=0}^{k-1} (\beta_l^k + \beta_l^{k+1}) p^l.$$

2) *Propriétés de l'algorithme*

LEMME 11 :

17) $(p^k, Np^l) = 0 \qquad \forall k \neq l$

18) $(g^k, p^l) = 0 \qquad 0 \leqslant l < k$

19) $(g^l, p^k) = (g^0, p^k) \qquad 0 \leqslant l < k$

20) $(g^k, p^k) = -(g^k, Kg^k)/\gamma^k$

21) $(Kg^k, Np^k) = -\gamma^k(p^k, Np^k)$

22) $(g^k, Kg^l) = 0 \qquad 0 \leqslant l < k$

23) $\exists k \leqslant n \quad g^k = 0$

*avec* $\gamma^k = (g^{k-1}, p^{k-1})/(p^{k-1}, Np^{k-1})$.

*Démonstration* : En combinant les résultats du lemme 1 à ceux du lemme 10.

LEMME 12 : *Si la matrice $K$ est symétrique, alors*

$$\beta_l^{k+1} = 0 \qquad \forall l \qquad 0 \leqslant l < k - 1 .$$

*Démonstration* : Par définition de $g^{k+1}$, on a

$$KNp^k = \frac{1}{\alpha^k} K(g^k - g^{k+1})$$

d'autre part
$$Np^l = \frac{1}{\alpha^l}(g^l - g^{l+1})$$

soit
$$(KNp^k, Np^l) = \frac{1}{\alpha^k}\,\frac{1}{\alpha^l}\,(K(g^k - g^{k+1}), (g^l - g^{l+1})) .$$

D'après le théorème 5, les vecteurs $g^k$ générés par l'algorithme (A.4) sont les mêmes que ceux de l'algorithme (A.1).

En particulier, le lemme 2 est utilisable si $K$ est symétrique :

$$(Kg^k, g^l) = 0 \qquad \forall l \neq k$$

d'où
$$(KNp^k, Np^l) = 0 \qquad \forall l \qquad 0 \leqslant l \leqslant k - 1 .$$

Et
$$(KNp^k, Np^{k-1}) = -\frac{1}{\alpha^k}\,\frac{1}{\alpha^{k-1}}\,(Kg^k, g^k)$$

$$(KNp^k, Np^k) = -\frac{1}{\alpha^k}\,\frac{1}{\alpha^k}\,\{(Kg^k, g^k) + (Kg^{k+1}, g^{k+1})\} . \qquad \square$$

Si $K$ est symétrique, l'algorithme (A.4) s'écrit donc

*Initialisation*

$$\begin{aligned} x^0 &\in R^n \text{ quelconque} \\ r^0 &= b - Ax^0 = A(x - x^0) \\ g^0 &= {}^tA\, {}^sHr^0 = N(x - x^0) \\ p^0 &= Kg^0 . \end{aligned}$$

*Itérations* (A.5)

$$\begin{aligned} \alpha^k &= (g^k, p^k)/(p^k, Np^k) \\ x^{k+1} &= x^k + \alpha^k p^k \\ g^{k+1} &= g^k - \alpha^k Np^k = {}^tA\, {}^sHr^{k+1} \\ p^{k+1} &= KNp^k + \beta_{k-1}^{k+1} p^{k-1} + \beta_k^{k+1} p^k \end{aligned}$$

*avec*

$$\begin{aligned} \beta_{k-1}^{k+1} &= -(Kg^k, g^k)/(\alpha^k \alpha^{k-1})(p^{k-1}, Np^{k-1}) \\ \beta_k^{k+1} &= -\{(Kg^{k+1}, g^{k+1}) + (Kg^k, g^k)\}/((\alpha^k \alpha^k) \times (p^k, Np^k)) \end{aligned}$$

*avec* $\beta_{-1}^1 = 0$ *par convention.*

Parmi les algorithmes s'écrivant sous cette forme, citons la méthode du résidu conjugué version Young et Jea [19] ($H = I$ et $K = {}^tA^{-1}$), et la méthode de la double suite orthogonale d'Amara et Nedelec [1] ($H = (A\,{}^tA)^{-1}$ et $K = {}^tAA$).

## SECONDE PARTIE : UN ALGORITHME DE MINIMISATION

### IX. Définition de l'algorithme de minimisation

On considère l'algorithme suivant

*Initialisation*

$$\begin{aligned} x^0 &\in \mathbb{R}^n \text{ quelconque} \\ r^0 &= b - Ax^0 = A(x - x^0) \\ g^0 &= {}^tA\, {}^sHr^0 = N(x - x^0) \\ p^0 &= Kg^0 . \end{aligned}$$

*Itérations* (A.6)

$$x^{k+1} = x^k + \sum_{l=0}^{k} \alpha_l^k p^l$$

*où les $\alpha_l^k$ réalisent le minimum de la fonction*

$$\begin{aligned} E(r^{k+1}) &= (r^{k+1}, Hr^{k+1}) \\ g^{k+1} &= g^k - \sum_{l=0}^{k} \alpha_l^k \, Np^l = {}^tA\,{}^sHr^{k+1} \\ \beta^{k+1} &= -(Kg^{k+1}, Np^k)/(p^k, Np^k) \\ p^{k+1} &= Kg^{k+1} + \beta^{k+1} p^k . \end{aligned}$$

## X. Propriétés de l'algorithme de minimisation

LEMME 15 : *Si on suppose les vecteurs $p^0, p^1, \dots, p^k$ linéairement indépendants, la matrice $B^k \in \mathbb{R}^{(k+1)\times(k+1)}$ définie par*

$$B_{ij}^k = (p^i, Np^j) \qquad 0 \leqslant i \leqslant k , \qquad 0 \leqslant j \leqslant k$$

*est symétrique définie positive.*

*Démonstration* :

$$\forall u \in \mathbb{R}^{k+1} , \quad (u, B^k u) = \sum_{i,j=0}^{k} u_i \, B_{ij}^k \, u_j = \left( \sum_{i=0}^{k} u_i \, Ap^i, {}^sH \sum_{j=0}^{k} u_j \, Ap^j \right)$$

soit
$$(u, B^k u) = (U, {}^sHU) \geqslant 0$$

où
$$U = \sum_{i=0}^{k} u_i \, Ap^i .$$

Puisque ${}^sH$ est définie positive,

$$(u, B^k u) = 0 \Leftrightarrow U = 0$$

si on suppose les vecteurs $p^0, p^1, \dots, p^k$ linéairement indépendants, alors

$$U = 0 \Leftrightarrow u = 0 . \qquad \square$$

LEMME 16 : *Si on suppose les vecteurs $p^0, p^1, \dots, p^k$ linéairement indépendants, la fonction $E(r^{k+1}) = (r^{k+1}, Hr^{k+1})$ admet un minimum unique pour $\alpha^k \in \mathbb{R}^{k+1}$ solution du système linéaire*

$$B^k \alpha^k = c^k$$

*où $c^k \in \mathbb{R}^{k+1}$ est défini par $c_i^k = (g^k, p^i) \qquad 0 \leqslant i \leqslant k$.*

*Démonstration* :

$$E^{k+1} = E(r^{k+1}) = (r^{k+1}, {}^sHr^{k+1})$$
$$= (r^k, {}^sHr^k) - 2\sum_{i=0}^{k} \alpha_i^k(g^k, p^i) + \sum_{i,j=0}^{k} \alpha_i^k \alpha_j^k(p^i, Np^j)$$

$\alpha \to E(r^{k+1})$ est fonction strictement convexe qui admet un minimum unique pour $\alpha^k$ solution du système linéaire

$$\sum_{j=0}^{k} (p^i, Np^j)\, \alpha_j^k = (g^k, p^i) \quad 0 \leqslant i \leqslant k$$

or d'après le lemme 15, la matrice $B^k$ est inversible, si on suppose les vecteurs $p^0, p^1, ..., p^k$ linéairement indépendants. □

On cherche à obtenir des propriétés analogues à celles du lemme 1 :

LEMME 17 : *Si on suppose les vecteurs $p^0, p^1, ..., p^{k-1}$ linéairement indépendants, alors*

24) $$(g^k, p^l) = 0 \qquad 0 \leqslant l < k$$

25) $$(g^k, p^k) = (g^k, Kg^k)$$

26) $$(g^k, Kg^l) = 0 \qquad 0 \leqslant l < k$$

27) $$(p^k, Np^{k-1}) = 0\,.$$

*Démonstration* :

24) Par définition de $\alpha^{k-1}$ :

$$\sum_{j=0}^{k-1} (p^i, Np^j)\, \alpha_j^{k-1} = (g^{k-1}, p^i) \qquad 0 \leqslant i \leqslant k$$

soit $$\left(p^i, g^{k-1} - \sum_{j=0}^{k-1} \alpha_j^{k-1} Np^j\right) = 0 \qquad 0 \leqslant i < k$$

i.e. $$(p^i, g^k) = 0 \qquad 0 \leqslant i < k\,.$$

25) $$(g^k, p^k) = (g^k, Kg^k + \beta^k p^{k-1}) = (g^k, Kg^k)$$

26) $$(g^k, Kg^l) = (g^k, p^l - \beta p^{l-1}) = 0 \qquad 0 \leqslant l < k$$

27) $$(p^k, Np^{k-1}) = (Kg^k, Np^{k-1}) + \beta^k(p^{k-1}, Np^{k-1})$$
$$= 0 \text{ par définition de } \beta^k\,.$$ □

*Remarque* 14 : Par construction les vecteurs $p^k$ et $p^{k+1}$ sont linéairement indépendants. On peut donc envisager de réaliser l'indépendance linéaire de tous les vecteurs $p^0, p^1, \dots, p^{k+1}$ en introduisant une relation de la forme

$$p^{k+1} = Kg^{k+1} + \sum_{l=0}^{k} \beta_l^{k+1} p^l$$

avec $$\beta_l^{k+1} = -(Kg^{k+1}, Np^l)/(p^l, Np^l) \qquad 0 \leqslant l \leqslant k .$$

Dans ces conditions $(p^l, Np^{k+1}) = 0 \qquad 0 \leqslant l \leqslant k$.

Mais alors la matrice $B^{k+1}$ est diagonale et le système linéaire

$$B^{k+1}\,\alpha^{k+1} = c^{k+1}$$

a pour solution $\alpha^{k+1}$, tel que $\alpha_l^{k+1} = (g^{k+1}, p^l)/(p^l, Np^l)$

soit $$\alpha_l^{k+1} = 0 \qquad\qquad 0 \leqslant l \leqslant k$$

et $$\alpha_{k+1}^{k+1} = (g^{k+1}, p^{k+1})/(p^{k+1}, Np^{k+1}) .$$

On retrouve l'algorithme (A.1) !

## XI. Équivalence des algorithmes d'orthogonalisation et de minimisation

D'après le lemme 1, la relation $(g^{n-1}, Kg^l) = 0, 0 \leqslant l \leqslant n-2$ est suffisante pour obtenir $g^n = 0$.

D'après le lemme 17 cette relation n'est vraie que si les vecteurs $p^0, p^1, \dots p^{n-1}$ sont linéairement indépendants. On est donc amené à étudier la dimension du sous-espace

$$\mathscr{E}^k = \langle p^0, p^1, \dots, p^k \rangle ,$$

THÉORÈME 6 : *Les algorithmes* (A.1) *et* (A.6) *sont équivalents, i.e. à partir du même* $x^0$, *ils calculent le même* $x^k$.

*Démonstration* : On indice par 1 les vecteurs générés par l'algorithme (A.1) :

$$p_1^0, p_1^1, \dots, p_1^k, g_1^k, x_1^{k+1} \dots$$

et par 2 les vecteurs générés par l'algorithme (A.6)

$$p_2^0, p_2^1, \dots, p_2^k, g_2^k, x_2^{k+1} \dots$$

D'après le lemme 3 $x_1^{k+1}$ réalise le minimum de $E(r^{k+1})$ sur l'espace $\mathscr{E}_1^k + x^0$.

Par définition, $x_2^{k+1}$ réalise le minimum de $E(r^{k+1})$ sur l'espace $\mathscr{E}_2^k + x^0$.

La fonction $E(r)$ étant strictement convexe, pour montrer que $x_1^{k+1} = x_2^{k+1}$, il suffit de montrer que $\mathscr{E}_1^k = \mathscr{E}_2^k$.

On procède par récurrence :

— la propriété est vraie pour $k = 0$, si on prend pour les deux algorithmes le même vecteur initial $x^0$,

— supposons la propriété vraie pour tout $l$, $0 \leqslant l < k$, alors

$$p_1^k = Kg_1^k + \sum_{l=0}^{k-1} \beta_l^k p_1^l$$

$$p_2^k = Kg_2^k + \beta^k p_2^{k-1} .$$

D'après l'hypothèse de récurrence, $\mathscr{E}_1^{k-1} = \mathscr{E}_2^{k-1}$, soit

$$x_1^k = x_2^k \quad \text{et} \quad g_1^k = g_2^k .$$

Donc $p_1^k \in \mathscr{E}_2^k$ et $p_2^k \in \mathscr{E}_1^k$.

Soit $$\mathscr{E}_1^k = \mathscr{E}_2^k . \qquad \square$$

*Conséquences* : Les théorèmes 1, 2 et 3 s'appliquent à l'algorithme (A.6), dont on connaît ainsi toutes les propriétés.

*Remarque* 15 : Si la matrice $K$ est symétrique, on montre par le même raisonnement que les directions générées par les deux algorithmes sont *identiques.* Ainsi les $\alpha_l^k$ de l'algorithme (A.6) vérifient

$$\alpha_l^k = 0 \qquad\qquad 0 \leqslant l < k$$

$$\alpha_k^k = (g^k, Kg^k)/(p^k, Np^k) .$$

Dans ce cas, l'algorithme (A.6) s'écrit sous la forme (A.2).

## XII. Une variante de l'algorithme (A.6) : la méthode MIN ($m$)

De manière analogue à la variante ORTHOMIN ($m$) de l'algorithme (A.1), on peut envisager un algorithme (A.6) qui n'utilise que les $m$ $(1 \leqslant m \leqslant k)$ dernières directions calculées, dans le cas où la matrice $K$ n'est pas symétrique.

Cette démarche est ici justifiée par l'argument supplémentaire suivant : le calcul des $\alpha_l^k$, $0 \leqslant l \leqslant k$ nécessite la résolution d'un système linéaire de rang $k + 1$, dont la matrice est pleine !

L'algorithme MIN ($m$) s'écrit

*Initialisation*

$$\begin{aligned} & x^0 \in \mathbb{R}^n \quad quelconque \\ & r^0 = b - Ax^0 \\ & g^0 = {}^tA\,{}^sHr^0 \\ & p^0 = Kg^0 . \end{aligned}$$

*Itérations* (A.7)

$$x^{k+1} = x^k + \sum_{l=k-m+1}^{k} \alpha^k p^l$$

*où les* $\alpha_l^k$ *minimisent* $E(r^{k+1}) = (r^{k+1}, Hr^{k+1})$

$$\begin{aligned} g^{k+1} &= g^k - \sum_{l=k-m+1}^{k} \alpha_l^l\, Np^l = {}^tA\,{}^sHr^{k+1} \\ \beta^{k+1} &= -\,(Kg^{k+1}, Np^k)/(p^k, Np^k) \\ p^{k+1} &= Kg^{k+1} + \beta^{k+1} p^k . \end{aligned}$$

THÉORÈME 7 : *Les algorithmes* (A.7) *et* (A.3) *sont équivalents.*

*Démonstration* : Identique à celle du théorème 6, il suffit de vérifier que les sous-espaces

$$\mathscr{E}_1^{m,k} = \langle p_1^{k-m},\ p_1^{k-m+1}, \dots, p_1^k \rangle$$

et $\mathscr{E}_2^{m,k} = \langle p_2^{k-m},\ p_2^{k-m+1}, \dots, p_2^k \rangle$ sont égaux . □

*Conséquences* : Le théorème 4 s'applique à l'algorithme (A.7).

L'algorithme du résidu minimal a été introduit par O. Axelsson [2], sous la forme (A.6), avec $H = I$ et $K = {}^tA^{-1}$, dans la pratique cette méthode est surtout utilisée sous la forme (A.7).

## TROISIÈME PARTIE : CLASSEMENT DES ALGORITHMES

En guise de conclusion à l'étude générale qui précède, on peut classer les algorithmes en trois catégories, suivant leur vitesse de convergence et les propriétés de la matrice du système $A$ :

1) Les algorithmes dont la vitesse de convergence est liée à cond ($A$), et qui convergent si $A$ est symétrique définie positive.

2) Les algorithmes qui convergent pour toute matrice $A$ régulière, avec une vitesse liée à cond $(A)^2$

3) Enfin une classe intermédiaire d'algorithmes qui convergent dès que la partie symétrique de $A$ est définie positive

Cette dernière catégorie est intéressante car elle contient les méthodes itératives de type ORTHOMIN $(m)$ et MIN $(m)$ qui offrent de nombreuses possibilités choix des directions conservées, contrôle de la convergence avec redémarrages éventuels En particulier, on peut noter l'importance fondamentale de la technique de préconditionnement, qui influence, non seulement la vitesse de convergence de ces algorithmes, mais qui conditionne aussi leur utilisation à partir d'une matrice $A$ régulière, on peut chercher $L$ et $U$ telles que la matrice $\hat{A} = L^{-1} A U^{-1}$ ait sa partie symétrique définie positive (ou se rapprocher de cette situation) le préconditionnement peut ainsi permettre de changer de classe d'algorithme

*Remarque* 16 La méthode du double gradient conjugué correspond au choix

$$H = \begin{bmatrix} 0 & A \\ {}^tA & 0 \end{bmatrix} \quad \text{et} \quad K = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

Les conditions suffisantes de convergence de l'algorithme (A 2) ne sont pas satisfaites et cet algorithme, qui n'utilise plus le principe de minimisation de la fonction $E$ peut dégénérer (voir Fletcher [7], et Jacobs [12])

*Remarque* 17 Toutes ces méthodes peuvent être aussi envisagées sous la forme des méthodes de projection oblique (voir par exemple Saad [16])

## BIBLIOGRAPHIE

1 M AMARA, J C NEDELEC, *Resolution du systeme matriciel indefini par une decomposition sur une double suite orthogonale* C R A S , Paris, 1982

2 O AXELSSON, *Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations*, Linear Algebra App 29 (1980)

3 P CONCUS, G H GOLUB, *A generalized conjugate gradient method for nonsymmetric systems of linear equations*, Lecture Notes in Economics and Mathematical systems, 134 , R Glowinski, J L Lions eds Springer Verlag, Berlin, 1976

4 J W DANIEL, *The conjugate gradient methol for linear and non linear operator equations* SIAM, J Num Anal , 4 (1967)

5 S C EISENSTAT, *A note on the generalized conjugate gradient method* SIAM, J Num Anal , 20 (1983)

6 S C EISENSTAT, H C ELMAN, M H SCHULTZ, *Variational Iteration methods for nonsymmetric systems of linear equations*, SIAM, J Num Anal , 20 (1983)

7 R Fletcher, *Conjugate gradient Methods for indefinite systems Proceedings of Dundee conference* in Num Anal (1979)

8 G H Golub, G Meurant, *Résolution numerique des grands systemes linéaires,* Eyrolles, 1984

9 L A Hageman, D M Young, *Applied iterative methods,* Academic Press (1981)

10 B P Il'In, *Some estimates for conjugate gradient methods,* USSR computational Mathematics and Math Physics (1978)

11 K Ito, *An iterative method for indefinite systems of linear equations,* ICASE Report n° 84-13

12 D A H Jacobs, *Generalization of the conjugate method for solving nonsymmetric and complex systems of algebraic equations*

13 P Joly, *Resolutions de systemes linéaires non symetriques par des methodes de gradient conjugue,* Publications du Laboratoire d'Analyse Numerique Universite Paris 6 (1982)

14 P Joly, *Methodes de gradient conjugue,* Publications du Laboratoire d'Analyse Numerique Universite Paris 6 (1984)

15 D S Kershaw, *The incomplete Cholesky-conjugate gradient method for the iterative solution of systems of linear equations,* J of Computational Physics V, 26 (1978)

16 Y Saad, *The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems* SIAM, J Num Anal, 19 (1982)

17 P K W Vinsome, *Orthomin, an iterative method for solving sparse sets of simultaneous linear equations* Proceedings 4th Symposium on Reservoir Simultation Society of Petroleum Engineers or AIME (1976)

18 O Widlung, *A Lanczos method for a class of nonsymmetric systems of linear equations,* SIAM, J Num Anal, 15 (1976)

19 D M Young, K C Jea, *Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods,* Linear Algebra Appl, 34 (1980)