RAIRO. ANALYSE NUMÉRIQUE

JEAN BEUNEU

Méthodes de projection-minimisation pour les problèmes linéaires

RAIRO. Analyse numérique, tome 17, n° 3 (1983), p. 221-248

<http://www.numdam.org/item?id=M2AN_1983__17_3_221_0>

R.A.I.R.O. Analyse numérique/Numerical Analysis
(vol. 17, n° 3, 1983, p. 221 à 248)

MÉTHODES DE PROJECTION-MINIMISATION POUR LES PROBLÈMES LINÉAIRES (*)

par Jean BEUNEU ([1])

Communiqué par F. ROBERT

Résumé. — *On montre que les méthodes de projection pour résoudre les systèmes d'équations linéaires se rattachent à un même formalisme général. On donne une condition suffisante de convergence pour la classe particulière des méthodes de minimisation, qui permet de prouver, entre autres, la convergence des méthodes basées sur une itération linéaire et celle des méthodes de partitionnement.*

Abstract. — *Projection methods for solving linear systems are shown to belong to the same general formalism. A sufficient convergence condition for the particular class of minimization methods is given, which proves, among other things, the convergence of methods based on a linear iteration, and of partitioning methods.*

1. INTRODUCTION

L'idée d'utiliser une projection pour résoudre les systèmes d'équations linéaires n'est pas nouvelle [9], [11]; les méthodes de projection sont cependant à l'heure actuelle de plus en plus utilisées pour les systèmes de grande taille [1], [6], [13], [17], [20], [22].

On présente un cadre théorique général pour ces méthodes, utilisant la projection oblique au sens de Galerkin-Petrov; chaque étape de ces méthodes consiste à résoudre un système linéaire de taille plus petite que le système donné. A ce formalisme se rattachent de nombreuses méthodes [4], [5]; on peut citer en particulier les travaux de Saad [18], [19] sur les méthodes de type Arnoldi ou Lanczos.

(*) Reçu en juillet 1982.
([1]) Université de Lille-I, I.E.E.A., Informatique, 59655 Villeneuve d'Ascq, Cedex.

R.A.I.R.O. Analyse numérique/Numerical Analysis, 0399-0516/1983/221/$ 5.00

On présente ensuite une classe de méthodes de projection, dites méthodes de minimisation, dont le principe consiste à minimiser en norme le résidu; on retrouve ainsi les méthodes de projection au sens de Householder et Bauer [11], [12]. On propose pour ces méthodes une condition suffisante générale de convergence, simple et fournissant dans les algorithmes un moyen de contrôle de la rapidité de convergence. Cette condition suffisante permet de prouver la convergence de nombreuses méthodes connues [4].

On propose enfin deux méthodes particulières, dont on montre la convergence à l'aide de la condition suffisante précédente :

Les méthodes basées sur une itération linéaire

Le sous-espace sur lequel on projette est engendré par certains vecteurs du tableau des différences finies construit sur les itérés successifs d'une itération linéaire. Ce type d'algorithmes s'inspire de méthodes d'accélération de convergence proposées par Kaniel-Stein [13] et Germain-Bonne [10].

Les méthodes de partitionnement

L'idée du rééquilibrage par partitionnement a été introduite par Wachspress [22] et étudiée par Froehlich [8] et Nakamura [15], [16] comme procédé d'accélération de convergence de la méthode de sur-relaxation (SOR) pour des systèmes linéaires obtenus à partir de problèmes d'équations aux dérivées partielles avec conditions aux limites (équations de diffusion des neutrons dans un réacteur). On l'utilise ici de manière indépendante.

2. LES MÉTHODES DE PROJECTIONS

2.1. Principe général des méthodes

Soit à résoudre l'équation linéaire

$$Ax = b \qquad x, b \in \mathbb{R}^n$$

où A est une application linéaire quelconque inversible de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On définit une méthode itérative de résolution :

A l'étape k, on dispose d'une solution approchée $\overset{k}{x}$; soient $\overset{k}{e} = \overset{k}{x} - x$ et $\overset{k}{r} = A\overset{k}{x} - b$ l'erreur et le résidu associés. On a $A\overset{k}{e} = \overset{k}{r}$.

Une méthode de projections consiste à utiliser à l'itération k une projection du type Galerkin-Petrov pour déterminer une approximation $\overset{k+1}{z}$ de la solution $-\overset{k}{e}$ de l'équation $Az + \overset{k}{r} = 0$.

Soit $p_k \in \mathbb{N}^+$ tel que $p_k \leqslant n$. Soient V_{p_k} et W_{p_k} deux sous-espaces de dimension p_k, et Π_{p_k} l'opérateur de projection oblique sur V_{p_k} orthogonalement à W_{p_k}. L'itération s'écrit :

Trouver $\overset{k+1}{z}$ tel que

$$\begin{cases} \overset{k+1}{z} \in V_{p_k} \\ \Pi_{p_k}(A\overset{k+1}{z} + \overset{k}{r}) = 0 . \end{cases}$$

Comme $\Pi_{p_k}\overset{k+1}{z} = \overset{k+1}{z}$, l'itération revient à résoudre dans V_{p_k}

$$\Pi_{p_k} A \Pi_{p_k} \overset{k+1}{z} + \Pi_{p_k} \overset{k}{r} = 0 .$$

On prend pour itéré suivant

$$\begin{aligned} &\overset{k+1}{x} = \overset{k}{x} + \overset{k+1}{z} \\ &\Rightarrow \overset{k+1}{e} = \overset{k+1}{x} - x = \overset{k}{x} - x + \overset{k+1}{z} = \overset{k}{e} + \overset{k+1}{z} . \end{aligned}$$

$\overset{k+1}{x}$ est donc une approximation de x d'autant meilleure que $\overset{k+1}{z}$ est une bonne approximation de $-\overset{k}{e}$,

$$\overset{k+1}{r} = A\overset{k+1}{e} = A\overset{k}{e} + A\overset{k+1}{z} = A\overset{k+1}{z} + \overset{k}{r} .$$

L'itération peut s'écrire :

Trouver $\overset{k+1}{x} = \overset{k}{x} + \overset{k+1}{z}$ tel que :

$$\begin{cases} \overset{k+1}{z} \in V_{p_k} \\ A\overset{k+1}{z} + \overset{k}{r} \perp W_{p_k} \quad \text{ou} \quad \overset{k+1}{r} \perp W_{p_k} . \end{cases}$$

Soit H_{p_k} l'image par A de V_{p_k}.

PROPOSITION 2.1 : *$\overset{k+1}{r}$ est la projection oblique de $\overset{k}{r}$ sur $W_{p_k}^{\perp}$ parallèlement à H_{p_k}.*

Démonstration : Soit E_{p_k} l'opérateur de projection oblique sur H_{p_k} orthogonalement à W_{p_k}

$$\begin{aligned} &\Pi_{p_k}(A\overset{k+1}{z} + \overset{k}{r}) = 0 \Rightarrow E_{p_k}\Pi_{p_k}(A\overset{k+1}{z} + \overset{k}{r}) = 0 : \\ &E_{p_k}\Pi_{p_k} = E_{p_k} \Rightarrow E_{p_k}(A\overset{k+1}{z} + \overset{k}{r}) = 0 . \\ &E_{p_k}A\overset{k+1}{z} = A\overset{k+1}{z} \quad \text{puisque} \quad A\overset{k+1}{z} \in H_{p_k} \Rightarrow A\overset{k+1}{z} = -E_{p_k}\overset{k}{r} . \\ &\overset{k+1}{r} = A\overset{k+1}{z} + \overset{k}{r} = \overset{k}{r} - E_{p_k}\overset{k}{r} = (I - E_{p_k})\overset{k}{r} . \quad \square \end{aligned}$$

2.2. Interprétation matricielle

Pour simplifier les notations on désigne par A la représentation matricielle, par rapport à la base canonique de $\mathbb{R}^n$, de l'opérateur linéaire A. A est donc une matrice régulière $n \times n$.

Soient $\{ \overset{k}{y}_1, \overset{k}{y}_2, \ldots, \overset{k}{y}_{p_k} \}$ une base de V_{p_k} et $\{ \overset{k}{u}_1, \overset{k}{u}_2, \ldots, \overset{k}{u}_{p_k} \}$ une base de W_{p_k}. On pose $\overset{k}{Y} = (\overset{k}{y}_1, \overset{k}{y}_2, \ldots, \overset{k}{y}_{p_k})$ et $\overset{k}{U} = (\overset{k}{u}_1, \overset{k}{u}_2, \ldots, \overset{k}{u}_{p_k})$ les matrices $n \times p_k$ dont les colonnes sont respectivement les $\overset{k}{y}_i$ et les $\overset{k}{u}_i$

$$\overset{k+1}{z} \in V_{p_k} \Rightarrow \overset{k+1}{z} = \overset{k}{Y}\overset{k}{\mu} \quad \text{avec} \quad \overset{k}{\mu} \in \mathbb{R}^{p_k}.$$

L'itération s'écrit :

Trouver $\overset{k}{\mu}$ tel que : $A\overset{k}{Y}\overset{k}{\mu} + \overset{k}{r} \perp W_{p_k}$.

La projection orthogonale est définie par rapport à un produit scalaire dans $\mathbb{R}^n$. Notons $(.,.)$ ce produit scalaire, et $(.,.)_2$ le produit scalaire euclidien dans $\mathbb{R}^{p_k}$.

Soit $\overset{k}{U}{}^*$ la matrice $p_k \times n$ telle que :

$$(x, \overset{k}{U}y) = (\overset{k}{U}{}^* x, y)_2 \qquad \forall x \in \mathbb{R}^n, y \in \mathbb{R}^{p_k}.$$

L'itération k consiste à résoudre le système linéaire $p_k \times p_k$

$$\overset{k}{U}{}^* A \overset{k}{Y}\overset{k}{\mu} + \overset{k}{U}{}^* \overset{k}{r} = 0$$

qui s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{u}_j, A\overset{k}{y}_i) + (\overset{k}{u}_j, \overset{k}{r}) = 0 \qquad j = 1, \ldots, p_k$$

et on prend pour itéré suivant :

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \overset{k}{y}_i .$$

Si $\overset{k}{U}{}^* A \overset{k}{Y}$ est inversible on obtient :

$$\overset{k}{\mu} = -(\overset{k}{U}{}^* A\overset{k}{Y})^{-1} \overset{k}{U}{}^* \overset{k}{r} \Rightarrow \overset{k+1}{x} = \overset{k}{x} - \overset{k}{Y}(\overset{k}{U}{}^* A\overset{k}{Y})^{-1} \overset{k}{U}{}^* \overset{k}{r}.$$

Remarque :

$$\begin{aligned} \overset{k+1}{r} &= A\overset{k+1}{z} + \overset{k}{r} = A\overset{k}{Y}\overset{k}{\mu} + \overset{k}{r} \\ &= [I - A\overset{k}{Y}(\overset{k}{U}{}^* A\overset{k}{Y})^{-1} \overset{k}{U}{}^*]\,\overset{k}{r}. \end{aligned}$$

$\overset{k}{A}\overset{k}{Y}(\overset{k}{U}^* \overset{k}{A}\overset{k}{Y})^{-1} \overset{k}{U}^*$ étant la représentation matricielle de E_{p_k}, on retrouve le résultat de la proposition 2.1.

3. LES MÉTHODES DE MINIMISATION

3.1. Principe des méthodes

Il s'agit d'un cas particulier de méthode de projection pour lequel on choisit $W_{p_k} = H_{p_k}$.

E_{p_k} est alors l'opérateur de projection orthogonale sur H_{p_k} et $\overset{k+1}{r}$ est la projection orthogonale de $\overset{k}{r}$ sur $H_{p_k}^{\perp}$ le complémentaire orthogonal de H_{p_k}.

Notons $\| \, . \, \|$ la norme par rapport à laquelle est définie la projection orthogonale

$$\| \overset{k+1}{r} \| = \| (I - E_{p_k}) \overset{k}{r} \| = \min_{\rho \in H_{p_k}} \| \overset{k}{r} + \rho \|$$
$$= \min_{z \in V_{p_k}} \| \overset{k}{r} + Az \| .$$

L'étape k consiste donc à minimiser $\| \overset{k+1}{r} \| = \| A \overset{k+1}{z} + \overset{k}{r} \|$ en fonction de $\overset{k+1}{z}$ pour $\overset{k+1}{z} \in V_{p_k}$.

Matriciellement, on a :

$$\| \overset{k+1}{r} \| = \min_{\mu \in \mathbb{R}^{p_k}} \| \overset{k}{r} + A\overset{k}{Y}\mu \| .$$

H_{p_k} est engendré par les vecteurs $\{ A\overset{k}{y}_1, ..., A\overset{k}{y}_{p_k} \}$; on peut donc choisir cette base dans W_{p_k} ; on a alors $\overset{k}{u}_i = A\overset{k}{y}_i$, $i = 1, ..., p_k$, et $\overset{k}{U} = A\overset{k}{Y}$. On retrouve ainsi la méthode « des moindres carrés » appliquée aux matrices [14]

$$\| \overset{k+1}{r} \| = \min_{\mu \in \mathbb{R}^{p_k}} \| \overset{k}{r} + \overset{k}{U}\mu \| .$$

Le système $p_k \times p_k$ de 2.2 devient :

$$\overset{k}{U}^*(\overset{k}{U}\overset{k}{\mu} + \overset{k}{r}) = 0$$

ou

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A\overset{k}{y}_j, A\overset{k}{y}_i) + (A\overset{k}{y}_j, \overset{k}{r}) = 0 \qquad j = 1, ..., p_k$$

système « aux moindres carrés » correspondant.

$\overset{k}{U}{}^{*} \overset{k}{U}$ étant toujours inversible, on obtient :

$$\overset{k}{\mu} = -(\overset{k}{U}{}^{*} \overset{k}{U})^{-1} \overset{k}{U}{}^{*} \overset{k}{r}.$$

3.2. Relations avec la méthode des moments [21]

Le sous-espace H_{p_k} est engendré par les vecteurs $\overset{k}{u}_i = A\overset{k}{y}_i$, $i = 1, \ldots, p_k$. On considère $\overset{k}{u}_{p_k+1}$ choisi de sorte que les $\overset{k}{u}_i$, $i = 1, \ldots, p_k + 1$, soient indépendants. La méthode des moments consiste à déterminer l'opérateur linéaire A_{p_k} défini sur H_{p_k} tel que :

$$\left| \begin{array}{ll} \overset{k}{u}_{i+1} = A_{p_k} \overset{k}{u}_i & i = 1, \ldots, p_k - 1 \\ E_{p_k} \overset{k}{u}_{p_k+1} = A_{p_k} \overset{k}{u}_{p_k}. & \end{array} \right.$$

Puisque $E_{p_k} \overset{k}{u}_{p_k+1} \in H_{p_k}$, on a

$$E_{p_k} \overset{k}{u}_{p_k+1} = -\sum_{i=1}^{p_k} \overset{k}{\mu}_i \overset{k}{u}_i \qquad \overset{k}{\mu}_i \in \mathbb{R}, \qquad i = 1, \ldots, p_k.$$

$E_{p_k} \overset{k}{u}_{p_k+1}$ étant la projection de $\overset{k}{u}_{p_k+1}$ sur H_{p_k}, $\overset{k}{u}_{p_k+1} - E_{p_k} \overset{k}{u}_{p_k+1}$ est orthogonal à H_{p_k} et donc à $\overset{k}{u}_1, \overset{k}{u}_2, \ldots, \overset{k}{u}_{p_k}$. Les $\overset{k}{\mu}_i$ sont donc solution du système linéaire d'ordre p_k

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{u}_j, \overset{k}{u}_i) + (\overset{k}{u}_j, \overset{k}{u}_{p_k+1}) = 0 \qquad j = 1, \ldots, p_k.$$

Si on prend $\overset{k}{u}_{p_k+1} = \overset{k}{r}$, on retrouve le système $p_k \times p_k$ de 3.1.

Dans la méthode des moments, A est un opérateur compact représenté par une matrice infinie, et on construit une suite d'opérateurs A_{p_k} convergeant vers l'opérateur A en faisant tendre p_k vers l'infini avec k. Par contre, dans les méthodes de projection, la valeur de p_k est totalement arbitraire (elle peut par exemple rester constante d'une itération à l'autre) et la suite des A_{p_k} n'est en général pas convergente. Donc, bien qu'on utilise implicitement, à chaque étape d'une méthode de minimisation, le même opérateur A_{p_k} que pour une itération de la méthode des moments, les deux méthodes sont fondamentalement différentes.

3.3. Quelques classes de méthodes de minimisation

Soient $(.,.)_2$ le produit scalaire euclidien et $\| . \|_2$ la norme euclidienne associée dans $\mathbb{R}^m$ $\forall m \in \mathbb{N}^+$. Étant donnée une matrice Q quelconque $n \times m$, on note Q^* la matrice telle que $(Q^* x, y)_2 = (x, Qy)$ $\forall x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.

LEMME 3.3 : *Toute norme associée à un produit scalaire est elliptique.*

Démonstration : Soit Q une matrice $n \times n$ quelconque inversible.
Posons $M = (Q^T)^{-1} Q^*$.
M est symétrique définie positive. En effet

$$(Q^{T-1} Q^* x, y)_2 = (Q^* x, Q^{-1} y)_2 = (x, y)$$
$$= (x, Q^{T-1} Q^* y)_2$$
$$(Q^{T-1} Q^* x, x)_2 = (x, x) = \| x \|^2 > 0 .$$

Donc $(x, y) = (Mx, y)_2 = (x, My)_2$ $\forall x, y \in \mathbb{R}^n$ avec M symétrique définie positive. De plus M est unique et indépendante de Q. En effet, s'il existe 2 matrices M_1 et M_2 telles que :

$$(M_1 x, y)_2 = (M_2 x, y)_2 = (x, y)$$
$$\Rightarrow ((M_1 - M_2) x, y)_2 = 0 \qquad \forall x, y \in \mathbb{R}^n$$
$$\Rightarrow (M_1 - M_2) x = 0 \qquad \forall x \in \mathbb{R}^n \Rightarrow M_1 - M_2 = 0 . \quad \square$$

$$(x, \overset{k}{U} y) = (\overset{k}{U}{}^* x, y)_2 \qquad \forall x \in \mathbb{R}^n, y \in \mathbb{R}^{p_k}$$
$$= (Mx, \overset{k}{U} y)_2 = (\overset{k}{U}{}^T Mx, y)_2$$
$$\Rightarrow \overset{k}{U}{}^* = \overset{k}{U}{}^T M .$$

Le système $p_k \times p_k$ de 3.1 devient :

$$\overset{k}{U}{}^T M(\overset{k}{U} \overset{k}{\mu} + \overset{k}{r}) = 0$$

ce qui s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (A\overset{k}{y}_j, MA\overset{k}{y}_i)_2 + (A\overset{k}{y}_j, M\overset{k}{r})_2 = 0 \qquad j = 1, ..., p_k .$$

La solution est $\overset{k}{\mu} = -(\overset{k}{U}{}^T M\overset{k}{U})^{-1} \overset{k}{U}{}^T M\overset{k}{r}$.

Sous cette forme on retrouve les méthodes de projection au sens de Householder-Bauer [11], [12].

D'un point de vue algorithmique, les matrices M les plus intéressantes sont celles pour lesquelles les calculs à effectuer sont les plus simples possibles. M sera donc en général liée à A.

On peut prendre par exemple :

$$M = A^{Tm} A^m \qquad m \in \mathbb{N}\,.$$

Le système à résoudre à l'étape k s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A\overset{k}{y}_j, A^{Tm} A^{m+1} \overset{k}{y}_i)_2 + (A\overset{k}{y}_j, A^{Tm} A^m \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k$$

$$\Rightarrow \sum_{i=1}^{p_k} \overset{k}{\mu}_i(A^{m+1} \overset{k}{y}_j, A^{m+1} \overset{k}{y}_i)_2 + (A^{m+1} \overset{k}{y}_j, A^m \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k\,.$$

$\overset{k+1}{x}$ est ensuite calculé par la relation

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \overset{k}{y}_i$$

et on itère avec $\overset{k+1}{x}$ au lieu de $\overset{k}{x}$.

Cas particuliers :

i) $m = 0$.

Alors $M = I$ opérateur identité. Le système devient

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A\overset{k}{y}_j, A\overset{k}{y}_i)_2 + (A\overset{k}{y}_j, \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k\,.$$

ii) $m = -1$.

Le système devient

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{y}_j, \overset{k}{y}_i)_2 + (A^{T-1} \overset{k}{y}_j, \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k\,.$$

A^{T-1} n'étant pas connue, ce système ne peut être résolu que si les $\overset{k}{y}_i$ sont de la forme $\overset{k}{y}_i = A^T \overset{k}{v}_i$, $\overset{k}{v}_i \in \mathbb{R}^n$.

Si A est symétrique et définie positive, on peut prendre $M = A^m$, $m \in \mathbb{N}$.

Le système s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A\overset{k}{y}_j, A^{m+1} \overset{k}{y}_i)_2 + (A\overset{k}{y}_j, A^m \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k$$

ou

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A^{m+2} \overset{k}{y}_j, \overset{k}{y}_i)_2 + (A^{m+1} \overset{k}{y}_j, \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k .$$

Cas particulier : $m = -1$.

Le système devient :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{y}_j, A\overset{k}{y}_i)_2 + (\overset{k}{y}_j, \overset{k}{r})_2 = 0 \qquad j = 1, \dots, p_k .$$

4. CONVERGENCE DES MÉTHODES DE MINIMISATION [2], [3]

4.1. Une condition suffisante de convergence

Si $\overset{k}{r} = 0$ on a nécessairement $\overset{k}{x} = x$ et dans ce cas la convergence est triviale. Dans le cas contraire :

Soit θ_k l'angle aigu entre $\overset{k}{r}$ et H_{p_k}, et soit $\overset{k}{\alpha} \in \mathbb{R}^{p_k}$ le vecteur dont les composantes sont :

$$\overset{k}{\alpha}_i = \left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}, \frac{\overset{k}{u}_i}{\|\overset{k}{u}_i\|}\right), \qquad \forall i = 1, \dots, p_k .$$

On a $|\overset{k}{\alpha}_i| \leqslant 1 \quad \forall i$.

Lemme 4.1 : *Soit* $\delta \in \mathbb{R}$ *tel que* $0 < \delta \leqslant 1$. *Si* $\exists t^k \in H_{p_k}$ *tel que*

$$\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}, \frac{\overset{k}{t}}{\|\overset{k}{t}\|}\right) \geqslant \delta ,$$

alors $1 \geqslant \cos\theta_k \geqslant \delta$.

Démonstration :

$$\cos\theta_k = \left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}, \frac{E_{p_k}\overset{k}{r}}{\|E_{p_k}\overset{k}{r}\|}\right) = \max_{\rho \in H_{p_k}} \left|\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}, \frac{\rho}{\|\rho\|}\right)\right| \geqslant$$

$$\geqslant \left|\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}, \frac{\overset{k}{t}}{\|\overset{k}{t}\|}\right)\right| \geqslant \delta . \quad \square$$

Proposition 4.1.1 : *Soit* $\delta \in \mathbb{R}$ *tel que* $0 < \delta \leqslant 1$. *Si* $\exists$ *une norme vectorielle* ϕ *sur* $\mathbb{R}^{p_k}$ *telle que* $\phi(\overset{k}{\alpha}) \geqslant \delta$, *alors* $\exists a > 0$, $1 \geqslant \cos\theta_k \geqslant a\delta$.

Démonstration :

$$\overset{k}{u}_i \in H_{p_k} \quad \forall i \Rightarrow \cos\theta_k \geqslant \max_i \left| \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|}, \frac{\overset{k}{u}_i}{\| \overset{k}{u}_i \|} \right) \right| = \max_i | \overset{k}{\alpha}_i | = \phi_\infty(\overset{k}{\alpha}) .$$

D'après l'équivalence des normes vectorielles dans $\mathbb{R}^{p_k}$, $\exists a > 0$ tel que

$$a\phi(\overset{k}{\alpha}) \leqslant \phi_\infty(\overset{k}{\alpha}) \Rightarrow \cos\theta_k \geqslant a\phi(\overset{k}{\alpha}) \geqslant a\delta > 0 . \quad \square$$

COROLLAIRE 4.1 : *Soit $\delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$ avec δ indépendant de k. Si $\exists$ une norme de Holder ϕ_q sur $\mathbb{R}^{p_k}$ telle que $\phi_q(\overset{k}{\alpha}) \geqslant \delta$, alors $\exists \delta' \in \mathbb{R}$ tel que $0 < \delta' \leqslant 1$ avec δ' indépendant de k, tel que $1 \geqslant \cos\theta_k \geqslant \delta'$.*

Démonstration :

$$\frac{\phi_q(\overset{k}{\alpha})}{p_k^{1/q}} \leqslant \phi_\infty(\overset{k}{\alpha}) \Rightarrow a = \frac{1}{p_k^{1/q}} \geqslant \frac{1}{n^{1/q}} \Rightarrow \cos\theta_k \geqslant a\phi_q(\overset{k}{\alpha}) \geqslant \alpha\delta \geqslant \frac{\delta}{n^{1/q}} = \delta' . \quad \square$$

Remarque : $\forall \delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$, on peut toujours trouver des vecteurs $\overset{k}{y}_i$ de façon qu'$\exists$ une norme ϕ_q telle que $\phi(\overset{k}{\alpha}) \geqslant \delta$. En effet, il suffit de prendre

$$\overset{k}{y}_1 = \overset{k}{e} \Rightarrow \overset{k}{\alpha}_1 = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|}, \frac{A\overset{k}{e}}{\| A\overset{k}{e} \|} \right) = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|}, \frac{\overset{k}{r}}{\| \overset{k}{r} \|} \right) = 1 .$$

Comme $| \overset{k}{\alpha}_i | \leqslant 1 \quad \forall i$ et $\overset{k}{\alpha}_1 = 1 \Rightarrow \phi_\infty(\overset{k}{\alpha}) = 1 \geqslant \delta$.

PROPOSITION 4.1.2 : *Soit $\delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$ avec δ indépendant de k. Si $\forall k$ on choisit les vecteurs $\overset{k}{y}_i$ de façon qu'$\exists$ une norme de Holder ϕ_q sur $\mathbb{R}^{p_k}$ telle que $\phi_q(\overset{k}{\alpha}) \geqslant \delta$, alors la suite des itérés $\overset{k}{x}$ converge linéairement vers x.*

Démonstration :

$$\| \overset{k+1}{r} \| = \| (I - E_{p_k}) \overset{k}{r} \| = \| \overset{k}{r} \| \sin\theta_k \Rightarrow$$

$$\Rightarrow \| \overset{k+1}{r} \|^2 = \| \overset{k}{r} \|^2 (1 - \cos^2\theta_k) \leqslant \| \overset{k}{r} \|^2 [1 - a^2 \phi_q^2(\overset{k}{\alpha})]$$

$$\leqslant \| \overset{k}{r} \|^2 (1 - a^2\delta^2) \leqslant \| \overset{k}{r} \|^2 (1 - \delta'^2)$$

avec $1 - \delta'^2 < 1$ et δ' indépendant de k, ce qui implique la convergence. $\square$

PROPOSITION 4.1.3 : *Soit $\delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$ avec δ indépendant de k. Si on choisit les vecteurs $\overset{k}{y}_i$ de façon qu'$\exists$ une norme de Holder ϕ_q sur $\mathbb{R}^{p_k}$ telle que $\forall K$, $\exists k > K \Rightarrow \phi_q(\overset{k}{\alpha}) \geqslant \delta$ alors la suite des itérés $\overset{k}{x}$ converge vers x.*

Démonstration : La suite des $\| \overset{k}{r} \|$ est non croissante et minorée par 0 donc converge. D'après l'hypothèse, il existe une sous-suite infinie d'indices k_j tels que $\phi_q(\overset{k_j}{\alpha}) \geqslant \delta$. Un raisonnement analogue à celui de la démonstration de la proposition 4.1.2 montre que :

$$
\begin{aligned}
&\| \overset{k_2}{r} \| \leqslant \| \overset{k_2-1}{r} \| \leqslant \dots \leqslant \| \overset{k_1+1}{r} \| \leqslant \sqrt{1-\delta'^2}\, \| \overset{k_1}{r} \| \\
&\| \overset{k_3}{r} \| \leqslant \| \overset{k_3-1}{r} \| \leqslant \dots \leqslant \| \overset{k_2+1}{r} \| \leqslant \sqrt{1-\delta'^2}\, \| \overset{k_2}{r} \| \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\| \overset{k_{j+1}}{r} \| \leqslant \| \overset{k_{j+1}-1}{r} \| \leqslant \dots \leqslant \| \overset{k_j+1}{r} \| \leqslant \sqrt{1-\delta'^2}\, \| \overset{k_j}{r} \| \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}
$$

La sous-suite des $\| \overset{k_j}{r} \|$ converge donc vers 0. La suite des $\| \overset{k}{r} \|$, admettant 0 pour point d'accumulation, a donc pour limite 0. □

4.2. Estimation du taux de convergence

Les propositions précédentes montrent que

$$
\| \overset{k+1}{r} \|^2 \leqslant \| \overset{k}{r} \|^2 \left[1 - \left(\frac{1}{p_k^{1/q}} \right)^2 \overline{\phi_q^2(\overset{k}{\alpha})} \right].
$$

Pour les trois normes de Holder classiques on obtient :

i) $$\phi_q = \phi_\infty \Rightarrow \| \overset{k+1}{r} \|^2 \leqslant \| \overset{k}{r} \|^2 \left[1 - \max_i | \overset{k}{\alpha}_i |^2 \right]$$

ii) $$\phi_q = \phi_1 \Rightarrow \| \overset{k+1}{r} \|^2 \leqslant \| \overset{k}{r} \|^2 \left[1 - \frac{1}{p_k^2} \left(\sum_{i=1}^{p_k} | \overset{k}{\alpha}_i | \right)^2 \right]$$

iii) $$\phi_q = \phi_2 \Rightarrow \| \overset{k+1}{r} \|^2 \leqslant \| \overset{k}{r} \| \left[1 - \frac{1}{p_k} \left(\sum_{i=1}^{p_k} \overset{k}{\alpha}{}_i^2 \right) \right].$$

Ces inégalités fournissent des bornes du taux de convergence ; d'un point de vue pratique, pour obtenir à l'itération k une diminution en norme de $\overset{k}{r}$ la plus élevée possible, on pourra choisir, pour un p_k donné, les vecteurs $\overset{k}{y}_i$ de façon que $\max | \overset{k}{\alpha}_i |$, $\sum_{i=1}^{p_k} \overset{k}{\alpha}_i$, ou $\sum_{i=1}^{p_k} \overset{k}{\alpha}{}_i^2$ aient la plus grande valeur possible.

On peut cependant, à chaque itération, donner une autre expression de la diminution en norme de $\overset{k}{r}$:

Soit H_i le sous-espace de dimension i engendré par $\overset{k}{u}_1, \overset{k}{u}_2, \dots, \overset{k}{u}_i$ pour $i = 1, \dots, p_k$, et soit ρ_i la projection de $\overset{k}{r}$ sur $H_i^\perp$

$$
\overset{k}{u}_{i+1} = E_i\, \overset{k}{u}_{i+1} + v_i \quad \text{avec} \quad v_i \in H_i^\perp \quad i = 1, \dots, p_k - 1 .
$$

H_{i+1} est donc engendré par les vecteurs $\overset{k}{u}_1, \dots, \overset{k}{u}_i, v_i$.

Soit ϕ_0 l'angle entre $\overset{k}{u}_1$ et $\overset{k}{r}$, et ϕ_i l'angle entre v_i et ρ_i, $i = 1, ..., p_k - 1$.

PROPOSITION 4.2

$$\| \overset{k+1}{r} \| = \| \overset{k}{r} \| \prod_{i=0}^{p_k-1} \sin \phi_i .$$

Démonstration : $v_i \in H_i^\perp$ et $\in H_{i+1}$; donc $H_i^\perp$ est la somme de $H_{i+1}^\perp$ et v_i ; $\exists a \in \mathbb{R}$ et $w_i \in H_{i+1}^\perp$ tels que $\rho_i = w_i + av_i$.

Par définition, $\overset{k}{r} = \rho_i + E_i \overset{k}{r}$; $E_i \overset{k}{r} \in H_i \Rightarrow E_i \overset{k}{r} \perp w_i$.

Donc

$$\overset{k}{r} = w_i + E_i \overset{k}{r} + av_i \quad \text{avec} \quad w_i \in H_{i+1}^\perp \quad \text{et} \quad E_i \overset{k}{r} + av_i \in H_{i+1} .$$

Comme $\overset{k}{r} = \rho_{i+1} + E_{i+1} \overset{k}{r}$, on en déduit $w_i = \rho_{i+1}$ et $E_i \overset{k}{r} + av_i = E_{i+1} \overset{k}{r}$. Donc $\rho_i = \rho_{i+1} + av_i$ avec $\rho_{i+1} \in H_{i+1}^\perp$ et $av_i \in H_{i+1}$ ce qui montre que ρ_{i+1} est la projection de ρ_i sur $H_{i+1}^\perp$ et $av_i = E_{i+1}\, \rho_i$.

Donc

$$\| \rho_{i+1} \| = \| \rho_i \| \sin \phi_i \qquad i = 1, ..., p_k - 1 .$$

D'autre part

$$\| \rho_1 \| = \| \overset{k}{r} \| \sin \phi_0 \quad \text{et} \quad \overset{k+1}{r} = \rho_{pk} .$$

Donc

$$\| \overset{k+1}{r} \| = \| \overset{k}{r} \| \prod_{i=0}^{p_k-1} \sin \phi_i . \quad \square$$

Si $v_i \perp \rho_i$, c'est-à-dire $v_i \perp \overset{k}{r}$, alors $\rho_{i+1} = \rho_i$. Par contre, si v_i est colinéaire à ρ_i, alors $\rho_{i+1} = 0$. Il est donc toujours possible de choisir v_i tel que

$$\| \rho_{i+1} \| < \| \rho_i \| ,$$

sauf si $\overset{k}{r} \in H_i$; dans ce dernier cas $E_i \overset{k}{r} = \overset{k}{r} \Rightarrow \rho_i = 0$.

On déduit de la proposition 4.2 une nouvelle borne du taux de convergence :

Soit

$$\overset{k}{\gamma}_i = \left(\frac{\overset{k}{u}_{i+1}}{\| \overset{k}{u}_{i+1} \|}, \frac{\rho_i}{\| \rho_i \|} \right) \qquad i = 1, ..., p_k - 1$$

et

$$\overset{k}{\gamma}_0 = \overset{k}{\alpha}_1 = \left(\frac{\overset{k}{u}_1}{\| \overset{k}{u}_1 \|}, \frac{\overset{k}{r}}{\| \overset{k}{r} \|}\right)$$

v_i étant la projection de $\overset{k}{u}_{i+1}$ sur $H_i^{\perp}$, on a $\cos \phi_i \geqslant | \overset{k}{\gamma}_i |$ $i = 1, \dots, p_k - 1$. De plus

$$\cos \phi_0 = | \overset{k}{\gamma}_0 | .$$

Donc

$$\| \overset{k+1}{r} \| \leqslant \| \overset{k}{r} \| \prod_{i=0}^{p_k - 1} \sqrt{1 - \overset{k}{\gamma}{}_i^2} .$$

D'un point de vue pratique, pour avoir à l'itération k une bonne amélioration sur $\| \overset{k}{r} \|$, on pourra choisir les vecteurs $\overset{k}{y}_i$ de façon que les $| \overset{k}{\gamma}_i |$ aient une ~~valeur élevée.~~

La proposition 4.2 montre de plus que si on augmente p_k en ajoutant de nouveaux vecteurs $\overset{k}{y}_i$ aux précédents on obtient une plus forte diminution en norme de $\overset{k}{r}$.

4.3. Autre forme de la condition suffisante de convergence

La norme utilisée est une norme elliptique associée à une matrice symétrique définie positive M : soit R la matrice racine carrée de M

$$\| \overset{k}{r} \| = \| R\overset{k}{r} \|_2 \qquad \| \overset{k}{u}_i \| = \| R\overset{k}{u}_i \|_2$$

$$\Rightarrow \overset{k}{\alpha}_i = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|}, \frac{\overset{k}{u}_i}{\| \overset{k}{u}_i \|}\right) = \frac{(\overset{k}{r}, M\overset{k}{u}_i)_2}{\| R\overset{k}{r} \|_2 \, \| R\overset{k}{u}_i \|_2} .$$

Soit $\overset{k}{\beta} \in \mathbb{R}^{p_k}$ le vecteur dont les composantes sont :

$$\overset{k}{\beta}_i = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|_2}, \frac{M\overset{k}{u}_i}{\| M\overset{k}{u}_i \|_2}\right)_2 \qquad \forall i = 1, \dots, p_k .$$

PROPOSITION 4.3 : *Soit $\delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$ avec δ indépendant de k. Si $\forall k$ on choisit les vecteurs $\overset{k}{y}_i$ de façon qu'$\exists$ une norme de Holder ϕ_q sur $\mathbb{R}^{p_k}$ telle que $\phi_q(\overset{k}{\beta}) \geqslant \delta$, alors la suite des itérés $\overset{k}{x}$ converge linéairement vers x.*

Démonstration :

$$|\overset{k}{\alpha}_i| = \left|\left(\frac{\overset{k}{r}}{\| R\overset{k}{r} \|_2}, \frac{M\overset{k}{u}_i}{\| R^{-1} M\overset{k}{u}_i \|_2}\right)_2\right|$$

$$\geqslant \frac{1}{\| R \|_2 \, \| R^{-1} \|_2}\left|\left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|_2}, \frac{M\overset{k}{u}_i}{\| M\overset{k}{u}_i \|_2}\right)_2\right|$$

$$\Rightarrow |\overset{k}{\alpha}_i| \geqslant \frac{1}{\gamma_2(R)} |\overset{k}{\beta}_i| \qquad i = 1, \ldots, p_k$$

$\gamma_2(R)$ étant le conditionnement de R

$$\Rightarrow \phi_q(\overset{k}{\alpha}) \geqslant \frac{1}{\gamma^2(R)} \phi_q(\overset{k}{\beta}) \geqslant \frac{\delta}{\gamma_2(R)} = \delta^*, 0 < \delta^* \leqslant \delta \leqslant 1 .$$

La proposition 4.1.2 implique donc la convergence. □

On pourrait également donner une proposition analogue à la proposition 4.1.3 en remplaçant $\overset{k}{\alpha}$ par $\overset{k}{\beta}$.

Applications

i) $M = I$:

$$\overset{k}{\beta}_i = \overset{k}{\alpha}_i = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|_2}, \frac{A\overset{k}{y}_i}{\| A\overset{k}{y}_i \|_2}\right)_2$$

ii) $M = A^{T-1} A^{-1}$:

$$\overset{k}{\beta}_i = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|_2}, \frac{A^{T-1}\overset{k}{y}_i}{\| A^{T-1}\overset{k}{y}_i \|_2}\right)_2$$

iii) A symétrique définie positive et $M = A^{-1}$:

$$\overset{k}{\beta}_i = \left(\frac{\overset{k}{r}}{\| \overset{k}{r} \|_2}, \frac{\overset{k}{y}_i}{\| \overset{k}{y}_i \|_2}\right)_2 .$$

5. CAS DES SYSTÈMES SUR-DÉTERMINÉS

On considère le système d'équations linéaires : $Ax = b$, $x \in \mathbb{R}^m$, $b \in \mathbb{R}^n$, où A est une matrice $n \times m$ quelconque, de rang maximum, avec $m < n$. Ce système n'admet pas en général de solution. On peut cependant lui appliquer la méthode de projections précédente :

Soit $p_k \in \mathbb{N}^+$ tel que $p_k \leqslant m$, et $\{\overset{k}{y}_1, \overset{k}{y}_2, \ldots, \overset{k}{y}_{p_k}\}$ une base de V_{p_k}. L'itération k consiste à résoudre le système linéaire $p_k \times p_k$:

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(A\overset{k}{y}_j, A\overset{k}{y}_i) + (A\overset{k}{y}_j, \overset{k}{r}) = 0 \qquad j = 1, \ldots, p_k$$

et on prend pour itéré suivant :

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \overset{k}{y}_i .$$

On note $(.,.)_2$ le produit scalaire euclidien dans $\mathbb{R}^m$.
Soit A^* la matrice $m \times n$ telle que :

$$(Ax, y) = (x, A^* y)_2 \qquad \forall x \in \mathbb{R}^m, y \in \mathbb{R}^n .$$

Si la norme utilisée dans la méthode de projections est une norme elliptique associée à une matrice $n \times n$ symétrique définie positive M, on a $A^* = A^T M$. Le système s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{y}_j, A^* A\overset{k}{y}_i)_2 + (\overset{k}{y}_j, A^* \overset{k}{r})_2 = 0 \qquad j = 1, \ldots, p_k .$$

Posons

$$C = A^* A \quad \text{et} \quad \overset{k}{\rho} = A^* \overset{k}{r} = A^*(A\overset{k}{x} - b) = C\overset{k}{x} - A^* b .$$

C est symétrique et définie positive. En effet :

$$(x, Cy)_2 = (x, A^* Ay)_2 = (Ax, Ay) = (Ay, Ax) = (y, Cx)_2 = (Cx, y)_2$$

$$(x, Cx)_2 = (x, A^* Ax)_2 = (Ax, Ax) = \| Ax \|^2 > 0 .$$

Le système à résoudre est donc :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i(\overset{k}{y}_j, C\overset{k}{y}_i)_2 + (\overset{k}{y}_j, \overset{k}{\rho})_2 = 0 \qquad j = 1, \ldots, p_k .$$

D'après 3.3, on voit que la méthode revient en fait à résoudre le système :

$$A^* Ax = A^* b$$

par une méthode de projections utilisant la norme elliptique associée à la matrice C^{-1}.

Le vecteur $\overset{k}{\alpha}'$ associé a pour composantes

$$\overset{k}{\alpha}'_i = \left(\frac{\overset{k}{\rho}}{\| \overset{k}{\rho} \|_{C^{-1}}}, \frac{C\overset{k}{y}_i}{\| C\overset{k}{y}_i \|_{C^{-1}}}\right)_{C^{-1}}$$

$$(\overset{k}{\rho}, C\overset{k}{y}_i)_{C^{-1}} = (\overset{k}{\rho}, \overset{k}{y}_i)_2 = (A^* \overset{k}{r}, \overset{k}{y}_i)_2 = (\overset{k}{r}, A\overset{k}{y}_i)$$

$$(\overset{k}{\rho}, \overset{k}{\rho})_{C^{-1}} = (A^* \overset{k}{r}, (A^* A)^{-1} A^* \overset{k}{r})_2 = (\overset{k}{r}, A(A^* A)^{-1} A^* \overset{k}{r})$$

$$(C\overset{k}{y}_i, C\overset{k}{y}_i)_{C^{-1}} = (C\overset{k}{y}_i, \overset{k}{y}_i)_2 = (A^* A\overset{k}{y}_i, \overset{k}{y}_i)_2 = (A\overset{k}{y}_i, A\overset{k}{y}_i)$$

$$\Rightarrow (\overset{k}{\alpha}'_i)^2 = \frac{(\overset{k}{r}, A\overset{k}{y}_i)^2}{(\overset{k}{r}, A(A^* A)^{-1} A^* \overset{k}{r}) \, \| A\overset{k}{y}_i \|^2}.$$

Le vecteur $\overset{k}{\beta}'$ associé a pour composantes :

$$\overset{k}{\beta}'_i = \left(\frac{\overset{k}{\rho}}{\| \overset{k}{\rho} \|_2}, \frac{\overset{k}{y}_i}{\| \overset{k}{y}_i \|_2}\right)_2$$

$$(\overset{k}{\rho}, \overset{k}{y}_i)_2 = (A^* \overset{k}{r}, \overset{k}{y}_i)_2 = (\overset{k}{r}, A\overset{k}{y}_i)$$

$$(\overset{k}{\rho}, \overset{k}{\rho})_2 = (A^* \overset{k}{r}, A^* \overset{k}{r})_2$$

$$\Rightarrow \overset{k}{\beta}'_i = \frac{(\overset{k}{r}, A\overset{k}{y}_i)}{\| A^* \overset{k}{r} \|_2 \, \| \overset{k}{y}_i \|_2}.$$

PROPOSITION 5 : *Soit $\delta \in \mathbb{R}$ tel que $0 < \delta \leqslant 1$ avec δ indépendant de k. Si $\forall k$ on choisit les vecteurs $\overset{k}{y}_i$ de façon qu'$\exists$ une norme de Holder ϕ_q sur $\mathbb{R}^{p_k}$ telle que $\phi_q(\overset{k}{\alpha}') \geqslant \delta$ (respectivement $\phi_q(\overset{k}{\beta}') \geqslant \delta$), alors la suite des itérés $\overset{k}{x}$ converge linéairement vers le x qui minimise $\| Ax - b \|$.*

Démonstration : D'après la proposition 4.1.2 (respectivement 4.3), la suite des $\overset{k}{x}$ converge vers le x qui vérifie

$$A^* Ax = A^* b.$$

Soit $r = Ax - b$

$$A^* r = 0 \Rightarrow (A^* r, z)_2 = 0 \qquad \forall z \in \mathbb{R}^m$$

$$\Rightarrow (r, Az) = 0 \qquad \forall z \in \mathbb{R}^m. \quad \square$$

Remarque : Dans le cas où $m = n$, on a :

$$(\overset{k}{r}, A(A^* A)^{-1} A^* \overset{k}{r}) = (\overset{k}{r}, \overset{k}{r})$$

$$\Rightarrow \overset{k}{\alpha}{}'_i = \frac{(\overset{k}{r}, A\overset{k}{y}_i)}{\| \overset{k}{r} \| \; \| A\overset{k}{y}_i \|} = \overset{k}{\alpha}_i .$$

Donc $\overset{k}{\alpha}{}' = \overset{k}{\alpha}$; dans ce cas on retrouve la méthode générale.

6. DEUX MÉTHODES PARTICULIÈRES

6.1. Utilisation d'une itération linéaire [3], [4]

Soit $B = I - A$. L'équation matricielle $Ax = b$ s'écrit sous la forme de l'équation de point fixe $x = Bx + b$.

Pour un entier $q > 0$, on considère, à partir de la solution approchée $\overset{k}{x}$ obtenue à l'étape k, l'itération linéaire suivante :

$$\left| \begin{aligned} & s_1 = \overset{k}{x} \\ & s_{i+1} = Bs_i + b \qquad\qquad i = 1, \ldots, p_k + q . \end{aligned} \right.$$

Soit

$$\Delta s_i = s_{i+1} - s_i \qquad\qquad i = 1, \ldots, p_k + q - 1 .$$

On a

$$\Delta s_i = B \, \Delta s_{i-1} \qquad\qquad i = 2, \ldots, p_k + q - 1 .$$

Soit, pour $j = 2, \ldots, q$

$$\Delta^j s_i = \Delta^{j-1} s_{i+1} - \Delta^{j-1} s_i \qquad i = 1, \ldots, p_k + q - j .$$

On montre facilement par récurrence que

$$\Delta^j s_{i+1} = B \, \Delta^j s_i \qquad\qquad i = 1, \ldots, p_k + q - j - 1$$

$$\Delta^j s_i = - A \, \Delta^{j-1} s_i \qquad\qquad i = 1, \ldots, p_k + q - j .$$

Supposons $p_k + q - 1 \leqslant n$. On fait l'hypothèse que B a au minimum $p_k + q - j$ valeurs propres distinctes et donc au minimum $p_k + q - j$ vecteurs propres indépendants. Soient respectivement $\lambda_1, \lambda_2, \ldots, \lambda_{p_k+q-j}$ ces valeurs propres et $v_1, v_2, \ldots, v_{p_k+q-j}$ ces vecteurs propres. On sait que $\lambda_i \neq 1$ puisque A est inversible.

Soit $\Delta s_1 = s_2 - s_1 = a_1 v_1 + a_2 v_2 + \cdots + a_{p_k+q-j} v_{p_k+q-j} + \cdots$.

LEMME 6.1 : *Si $a_1 \neq 0, a_2 \neq 0, \ldots, a_{p_k+q-j} \neq 0$, les vecteurs $\Delta^j s_1, \ldots, \Delta^j s_{p_k+q-j}$ sont linéairement indépendants $\forall j = 1, \ldots, q$.*

Démonstration :

$$\Delta^j s_i = (-1)^{j-1} A^{j-1} \Delta s_i = (-1)^{j-1} A^{j-1} B^{i-1} \Delta s_1$$
$$= a_1 \lambda_1^{i-1}(\lambda_1 - 1)^{j-1} v_1 + a_2 \lambda_2^{i-1}(\lambda_2 - 1)^{j-1} v_2 + \cdots$$

Supposons qu'il existe $b_1, b_2, \ldots, b_{p_k+q-j}$ tels que $\sum_{i=1}^{p_k+q-j} b_i \Delta^j s_i = 0$.

En écrivant que les coefficients des $p_k + q - j$ vecteurs propres $v_1, v_2, \ldots, v_{p_k+q-j}$ sont nuls, on obtient :

$$a_1(\lambda_1 - 1)^{j-1} \sum_{i=1}^{p_k+q-j} \lambda_1^{i-1} b_i = 0$$
$$a_2(\lambda_2 - 1)^{j-1} \sum_{i=1}^{p_k+q-j} \lambda_2^{i-1} b_i = 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{p_k+q-j}(\lambda_{p_k+q-j} - 1)^{j-1} \sum_{i=1}^{p_k+q-j} \lambda_{p_k+q-j}^{i-1} b_i = 0 .$$

Les λ_i étant distincts, $\lambda_i \neq 1$, et $a_i \neq 0\ \forall i = 1, \ldots, p_k + q - j$, le déterminant (Van der Monde) de ce système est non nul et on en déduit :

$$b_1 = b_2 = \cdots = b_{p_k+q-j} = 0 . \quad \square$$

Dans la suite, on suppose que A a au moins p_k valeurs propres distinctes (et donc B aussi).

On se donne un entier l tel que $1 \leqslant l \leqslant q - 1$, et on choisit pour vecteurs $\overset{k}{y}_i$ les vecteurs $\Delta^l s_i$ construits à partir des itérés $s_1, s_2, \ldots, s_{p_k+l}$

$$\overset{k}{y}_i = \Delta^l s_i \qquad \forall i = 1, \ldots, p_k .$$

On a :

$$\overset{k}{r} = A s_1 - b = s_1 - B s_1 - b = s_1 - s_2 = -\Delta s_1 .$$

Le système linéaire de l'étape k s'écrit donc :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (A \Delta^l s_j, A \Delta^l s_i) - (A \Delta^l s_j, \Delta s_1) = 0 \qquad j = 1, \ldots, p_k$$

ou

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (\Delta^{l+1} s_j, \Delta^{l+1} s_i) + (\Delta^{l+1} s_j, \Delta s_1) = 0 \qquad j = 1, \ldots, p_k$$

et

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \Delta^l s_i .$$

$A^* A$ est symétrique et définie positive. En effet :

$$(x, A^* Ay)_2 = (Ax, Ay) = (A^* Ax, y)_2 \qquad x, y \in \mathbb{R}^n$$
$$(x, A^* Ax)_2 = (Ax, Ax) = \| Ax \|^2 > 0 \qquad x \in \mathbb{R}^n .$$

Soit Q la matrice racine carrée de $A^* A$.

Proposition 6.1.1 : *Si la matrice A^* est symétrique définie positive, la suite des $\overset{k}{x}$ converge vers x $\forall l$.*

Démonstration :

$$|\overset{k}{\alpha}_1| = \left| \left(\frac{\Delta s_1, A\, \Delta^l s_1}{\| \Delta s_1 \| \; \| A\, \Delta^l s_1 \|} \right) \right| = \left| \left(\frac{\Delta s_1, A\, \Delta^l s_1}{\| \Delta s_1 \| \; \| A^l\, \Delta s_1 \|} \right) \right|$$

$$(x, y) = (A^{-1} x, A^* AA^{-1} y)_2 = (QA^{-1} x, QA^{-1} y)_2 \quad \forall x, y \in \mathbb{R}^n .$$

Posons $A^{-1} \Delta s_1 = -A^{-1} \overset{k}{r} = -\overset{k}{e}$

$$|\overset{k}{\alpha}_1| = \left| \left(\frac{Q\overset{k}{e}, QA^l \overset{k}{e}}{\| Q\overset{k}{e} \|_2 \; \| QA^l \overset{k}{e} \|_2} \right)_2 \right| = \left| \left(\frac{Q\overset{k}{e}, (QA^l Q^{-1})\, Q\overset{k}{e}}{\| Q\overset{k}{e} \|_2 \; \| (QA^l Q^{-1})\, Q\overset{k}{e} \|_2} \right)_2 \right|$$

$$QA^l Q^{-1} = (QAQ^{-1})^l$$

QAQ^{-1} est symétrique définie positive. En effet

$$A^T Q^2 = A^T A^* A = (A^T A^* A)^T = (A^* A)^T A = A^* A^2 = Q^2 A$$

$$\Rightarrow (QAQ^{-1})^T = Q^{-1} A^T Q = Q^{-1} A^T Q^2 Q^{-1} = Q^{-1} Q^2 AQ^{-1} = QAQ^{-1}$$

$$(x, QAQ^{-1} x)_2 = (Q^{-1} x, Q^2 AQ^{-1} x)_2 = (AQ^{-1} x, A^* AQ^{-1} x)_2 > 0$$

$$|\overset{k}{\alpha}_1| \geqslant \frac{|(Q\overset{k}{e}, (QAQ^{-1})^l\, Q\overset{k}{e})_2|}{\| Q\overset{k}{e} \|_2^2} \; \frac{1}{\| (QAQ^{-1})^l \|_2} .$$

Soient ν_1 et ν_n la plus grande et la plus petite valeur propre de A ; QAQ^{-1} étant semblable à A, ce sont aussi la plus grande et la plus petite valeur propre de QAQ^{-1}

$$\frac{|(Q\overset{k}{e}, (QAQ^{-1})^l\, Q\overset{k}{e})_2|}{\| Q\overset{k}{e} \|_2^2} \geqslant \nu_n^l$$

$$\| (QAQ^{-1})^l \|_2 = \nu_1^l$$

$$\Rightarrow |\overset{k}{\alpha}_1| \geqslant \left(\frac{\nu_n}{\nu_1} \right)^l .$$

Les hypothèses de la proposition 4.1.2 sont donc satisfaites avec $\delta = (\nu_n/\nu_1)^l$. □

De la proposition 6.1.1 on déduit une borne sur le taux de convergence

$$\| \overset{k+1}{r} \| \leqslant \sqrt{1 - \left(\frac{\nu_n}{\nu_1}\right)^{2l}} \, \| \overset{k}{r} \| .$$

Si $l = 1$, on peut donner une borne plus précise :

PROPOSITION 6.1.2 (voir [14]) : *Sous les hypothèses de la proposition* 6.1.1, *on a pour* $l = 1$:

$$\| \overset{k+1}{r} \| \leqslant \frac{1}{\left| T_{p_k}\left(- \dfrac{\nu_1 + \nu_n}{\nu_1 - \nu_n} \right) \right|} \| \overset{k}{r} \| .$$

T_{p_k} *étant le polynôme de Tchebycheff de degré* p_k.

Démonstration :

$$\begin{aligned} \overset{k+1}{r} &= \overset{k}{r} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \, A \, \Delta s_i \\ &= \overset{k}{r} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \, A B^{i-1} \, \Delta s_1 \\ &= \overset{k}{r} - \sum_{i=1}^{p_k} \overset{k}{\mu}_i \, A B^{i-1} \, \overset{k}{r} \\ &= \overset{k}{r} - \sum_{i=1}^{p_k} \overset{k}{\mu}_i A (I - A)^{i-1} \, \overset{k}{r} \\ &= P(A) \, \overset{k}{r} \text{ avec } P(0) = 1, \ P \text{ de degré } p_k . \end{aligned}$$

Le polynôme P est celui qui minimise $\| P(A) \overset{k}{r} \|$ parmi tous les polynômes de degré p_k.

Donc

$$\| \overset{k+1}{r} \| \leqslant \min_{P \text{ de d}^\circ p_k} \| P(A) \| \ \| \overset{k}{r} \|$$

$$\begin{aligned} \| P(A) \| &= \max_x \frac{\| P(A) \, x \|}{\| x \|} = \max_x \frac{\| Q P(A) \, A^{-1} \, x \|_2}{\| Q A^{-1} \, x \|_2} \\ &= \max_y \frac{\| Q P(A) \, Q^{-1} \, y \|_2}{\| y \|_2} = \max_y \frac{\| P(Q A Q^{-1}) \, y \|_2}{\| y \|_2} \\ &= \| P(Q A Q^{-1}) \|_2 \\ &= \rho[P(Q A Q^{-1})] \text{ car } Q A Q^{-1} \text{ est symétrique} \end{aligned}$$

$$
\begin{aligned}
&= \rho[P(A)] \\
&= \max_{\nu \text{ valeur propre de } A} | P(\nu) |
\end{aligned}
$$

$$
\begin{aligned}
\min_{P \text{ de d}^\circ p_k} \| P(A) \| &= \min_{P \text{ de d}^\circ p_k} \ \max_{\nu \text{ valeur propre de } A} | P(\nu) | \\
&= \max_{\nu \text{ valeur propre de } A} \ \min_{P \text{ de d}^\circ p_k} | P(\nu) | \\
&\leqslant \max_{\nu_n \leqslant \nu \leqslant \nu_1} \ \min_{P \text{ de d}^\circ p_k} | P(\nu) | \\
&= \frac{1}{\left| T_{p_k}\left(- \dfrac{\nu_1 + \nu_n}{\nu_1 - \nu_n} \right) \right|} < 1 . \quad \square
\end{aligned}
$$

COROLLAIRE 6.1 (voir [1]) : *Sous les hypothèses de la proposition* 6.1.1, *dans le cas* $l = 1$, *la valeur de* p_k *à choisir pour obtenir une réduction en norme de* $\overset{k}{r}$ ~~*inférieure à un* ε *fixé est proportionnelle à la racine carrée du conditionnement*~~ *de* A.

Démonstration : Le conditionnement euclidien de A est :

$$
\gamma_2(A) = \| A \|_2 \| A^{-1} \|_2 = \frac{\nu_1}{\nu_n} .
$$

Supposons

$$
p_k \geqslant \frac{1}{2} \sqrt{\frac{\nu_1}{\nu_n}} \operatorname{Log} \frac{2}{\varepsilon}
$$

$$
\frac{1 + \sqrt{\nu_n/\nu_1}}{1 - \sqrt{\nu_n/\nu_1}} \geqslant e^{2\sqrt{\nu_n/\nu_1}}
$$

$$
\Rightarrow p_k \operatorname{Log} \left(\frac{1 + \sqrt{\nu_n/\nu_1}}{1 - \sqrt{\nu_n/\nu_1}} \right) \geqslant 2\, p_k \sqrt{\frac{\nu_n}{\nu_1}} \geqslant \operatorname{Log} \frac{2}{\varepsilon}
$$

$$
\Rightarrow 2 \left(\frac{1 - \sqrt{\nu_n/\nu_1}}{1 + \sqrt{\nu_n/\nu_1}} \right)^{p_k} \leqslant \varepsilon
$$

$$
\Rightarrow \frac{1}{T_{p_k}\left(\dfrac{\nu_1 + \nu_n}{\nu_1 - \nu_n} \right)} \leqslant 2 \left(\frac{1 - \sqrt{\nu_n/\nu_1}}{1 + \sqrt{\nu_n/\nu_1}} \right)^{p_k} \leqslant \varepsilon
$$

$$
\Rightarrow \frac{\| \overset{k+1}{r} \|}{\| \overset{k}{r} \|} \leqslant \varepsilon . \quad \square
$$

Dans le cas général (l quelconque) le fait que $\delta = (\nu_n/\nu_1)^l$ montre que la vitesse de convergence est d'autant plus élevée que le conditionnement de A est petit. Dans le cas $l = 1$, la propriété précédente montre l'influence de la valeur de p_k sur le taux de convergence.

1er *cas particulier*

A est symétrique définie positive et on utilise la norme elliptique associée à $M = A^{m-1}$ avec m tel que $m + l \leqslant q$.

Le système à résoudre s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (\Delta^{l+1} s_j, A^{m-1} \Delta^{l+1} s_i)_2 + (\Delta s_1, A^{m-1} \Delta^{l+1} s_j)_2 = 0 \qquad j = 1, \dots, p_k$$

ou

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (\Delta^{l+1} s_j, \Delta^{m+l} s_i)_2 + (\Delta s_1, \Delta^{m+l} s_j)_2 = 0 \qquad j = 1, \dots, p_k$$

avec

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \Delta^l s_i .$$

La proposition 6.1.1 montre la convergence $\forall l$ et $\forall m$. Si on choisit m et l tels que $m + l = 1$, le système devient :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (\Delta^{l+1} s_j, \Delta s_i)_2 + (\Delta s_1, \Delta s_j)_2 = 0 \qquad j = 1, \dots, p_k .$$

2e *cas particulier*

On utilise la norme elliptique associée à $M = A^{T^{m-1}} A^{m-1}$ avec m tel que $m + l \leqslant q$. Le système à résoudre s'écrit :

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (A^{m-1} \Delta^{l+1} s_j, A^{m-1} \Delta^{l+1} s_i)_2 + (A^{m-1} \Delta s_1, A^{m-1} \Delta^{l+1} s_j)_2 = 0$$

$$j = 1, \dots, p_k$$

ou

$$\sum_{i=1}^{p_k} \overset{k}{\mu}_i (\Delta^{m+l} s_j, \Delta^{m+l} s_i)_2 + (\Delta^m s_1, \Delta^{m+l} s_j)_2 = 0 \qquad j = 1, \dots, p_k$$

avec

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \Delta^l s_i .$$

Si A est symétrique définie positive, $A^* = A^{2m-1}$ l'est aussi ; la proposition 6.1.1 montre alors la convergence. Sinon, on peut montrer une propriété plus faible :

PROPOSITION 6.1.3 : $\forall \delta$ *tel que* $0 < \delta < 1, \exists m^*, \forall m$ *tel que*

$$m > m^* \Rightarrow | \overset{k}{\alpha}_1 | > \delta .$$

Démonstration :

$$| \overset{k}{\alpha}_1 | = \left| \left(\frac{\Delta s_1, A\, \Delta^l s_1}{\| \Delta s_1 \| \;\; \| A\, \Delta^l s_1 \|} \right) \right|$$

$$\Rightarrow \overset{k}{\alpha}{}_1^2 = \frac{(A^{m-1}\, \Delta s_1, A^{m+l-1}\, \Delta s_1)_2^2}{(A^{m-1}\, \Delta s_1, A^{m-1}\, \Delta s_1)_2\, (A^{m+l-1}\, \Delta s_1, A^{m+l-1}\, \Delta s_1)_2} .$$

Soient $\nu_1, \nu_2, \ldots, \nu_n$ les valeurs propres de $A^T A$ supposées telles que $\nu_1 > \nu_2 \geqslant \ldots \geqslant \nu_n > 0$, et $w_1, w_2, \ldots, w_n$ les vecteurs propres correspondants.

Soit

$$\Delta s_1 = \sum_{i=1}^{n} a_i\, w_i \quad \text{et} \quad A^l\, \Delta s_1 = \sum_{i=1}^{n} b_i\, w_i$$

et supposons $a_1 \neq 0$ et $b_1 \neq 0$.

$$(A^T A)^{m-1}\, \Delta s_1 = \sum_{i=1}^{n} a_i\, \nu_i^{m-1}\, w_i$$

$$(A^T A)^{m-1}\, A^l\, \Delta s_1 = \sum_{i=1}^{n} b_i\, \nu_i^{m-1}\, w_i$$

$$(A^{m-1}\, \Delta s_1, A^{m+l-1}\, \Delta s_1)_2 = (\Delta s_1, (A^T A)^{m-1}\, A^l\, \Delta s_1)_2 = \sum_{i=1}^{n} a_i\, b_i\, \nu_i^{m-1}$$

$$(A^{m-1}\, \Delta s_1, A^{m-1}\, \Delta s_1)_2 = (\Delta s_1, (A^T A)^{m-1}\, \Delta s_1)_2 = \sum_{i=1}^{n} a_i^2\, \nu_i^{m-1}$$

$$(A^{m+l-1}\, \Delta s_1, A^{m+l-1}\, \Delta s_1)_2 = (A^l\, \Delta s_1, (A^T A)^{m-1}\, A^l\, \Delta s_1)_2 = \sum_{i=1}^{n} b_i^2\, \nu_i^{m-1} \Rightarrow$$

$$\Rightarrow \overset{k}{\alpha}{}_1^2 = \frac{\left[1 + \sum_{i=2}^{n} \frac{a_i\, b_i}{a_1\, b_1} \left(\frac{\nu_i}{\nu_1} \right)^{m-1} \right]^2}{\left[1 + \sum_{i=2}^{n} \left(\frac{b_i}{b_1} \right)^2 \left(\frac{\nu_i}{\nu_1} \right)^{m-1} \right] \left[1 + \sum_{i=2}^{n} \left(\frac{a_i}{a_1} \right)^2 \left(\frac{\nu_i}{\nu_1} \right)^{m-1} \right]}$$

On voit que

$$\lim_{m \to \infty} \overset{k}{\alpha}{}_1^2 = 1 . \quad \square$$

Le rang m^* dépend de k et, bien que fini, est non borné. Or la norme utilisée, et donc la valeur de m, ne doit pas changer au cours des itérations. Même en choisissant m assez grand, on ne peut donc conclure de façon certaine à la convergence.

6.2. Utilisation d'un procédé de partitionnement [2]

Soit $\overset{k}{y} \in \mathbb{R}^n$ quelconque. On a $\overset{k}{y} = \sum_{j=1}^{n} \overset{k}{a}_j e_j$ (e_j vecteurs de la base canonique).

On suppose que $\overset{k}{a}_j \neq 0 \quad \forall j = 1, \dots, n$.

On partitionne l'ensemble d'indices $\{ 1, \dots, n \}$ en p_k ensembles d'indices $I_1, I_2, \dots, I_{p_k}$

$$\{ 1, \dots, n \} = I_1 \cup I_2 \cup \dots \cup I_{p_k}$$

avec

$$I_j \cap I_l = \varnothing \qquad \forall j, l \in \{ 1, \dots, p_k \}, j \neq l$$

et on prend

$$\overset{k}{y}_i = \sum_{j \in I_i} \overset{k}{a}_j e_j \qquad i = 1, \dots, p_k .$$

Les vecteurs $\overset{k}{y}_i$ sont donc orthogonaux ; d'autre part, $\overset{k}{y} \in V_{p_k} \quad \forall k$ puisque

$$\overset{k}{y} = \sum_{i=1}^{p_k} \overset{k}{y}_i .$$

PROPOSITION 6.2.1 : *A chaque itération k, $\forall \overset{k}{y} \in \mathbb{R}^n$, $\forall p_k$ tel que $1 < p_k \leqslant n$, on peut déterminer un partitionnement de $\overset{k}{y}$ de façon à assurer la convergence de la suite des itérés $\overset{k}{x}$ vers x.*

Démonstration :

$$\overset{k}{y} = \overset{k}{a}_i e_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \overset{k}{a}_j e_j \qquad \forall i \in \{ 1, \dots, n \} .$$

Pour un i donné, on pose $\overset{k}{y}_1 = \overset{k}{a}_i e_i$ et on choisit les $\overset{k}{y}_j$, $j = 2, \dots, p_k$, tels que

$$\sum_{j=2}^{p_k} \overset{k}{y}_j = \sum_{\substack{j=1 \\ j \neq i}}^{n} \overset{k}{a}_j e_j .$$

$$\begin{aligned}
|\overset{k}{\alpha}_1| &= \left|\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}\,\frac{A\overset{k}{y}_1}{\|A\overset{k}{y}_1\|}\right)\right| = \left|\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|}\,\frac{\overset{k}{a}_i\,Ae_i}{|\overset{k}{a}_i|\,\|Ae_i\|}\right)\right| \\
&= \left|\left(\frac{\overset{k}{r}}{\|\overset{k}{r}\|},\frac{Ae_i}{\|Ae_i\|}\right)\right| \\
&= \frac{|(A^*\overset{k}{r},e_i)|}{\|\overset{k}{r}\|\;\|Ae_i\|} = \frac{|(A^*\overset{k}{r})_i|}{\|\overset{k}{r}\|\;\|Ae_i\|}\,.
\end{aligned}$$

On choisit le i qui maximise $|(A^*\overset{k}{r})_i|$, c'est-à-dire le i tel que

$$\begin{aligned}
\phi_\infty(A^*\overset{k}{r}) = |(A^*\overset{k}{r})_i| \Rightarrow |\overset{k}{\alpha}_1| &= \frac{\phi_\infty(A^*\overset{k}{r})}{\|\overset{k}{r}\|}\,\frac{1}{\|Ae_i\|} \\
&= \frac{\phi_\infty(A^*\overset{k}{r})}{\|\overset{k}{r}\|}\,\frac{\phi_\infty(e_i)}{\|Ae_i\|} \\
&\geqslant \frac{1}{S_{N_\infty}(A^{*-1})}\,\frac{1}{S_{N_\infty}(A)}\,.
\end{aligned}$$

Les hypothèses de la proposition 4.1.2 sont donc vérifiées avec

$$\delta = \frac{1}{S_{N_\infty}(A^{*-1})\,S_{N_\infty}(A)}\,. \quad \square$$

De la proposition 6.2.1 on déduit une borne sur le taux de convergence

$$\|\overset{k+1}{r}\| \leqslant \sqrt{1-\frac{1}{S^2_{N_\infty}(A^{*-1})\,S^2_{N_\infty}(A)}}\;\|\overset{k}{r}\|\,.$$

On peut encore donner une autre borne :

PROPOSITION 6.2.2 : *En utilisant le même partitionnement que dans la proposition* 6.2.1, *on a* :

$$\|\overset{k+1}{r}\| \leqslant \sqrt{1-\frac{1}{n\gamma_2(M)\,\gamma_2^2(A)}}\;\|\overset{k}{r}\|\,.$$

Démonstration : A^*A est symétrique et définie positive; soit Q la matrice telle que $Q^2 = A^*A$

$$(x,y) = (A^{-1}x, A^*AA^{-1}y)_2 = (QA^{-1}x, QA^{-1}y)_2 \quad \forall x, y \in \mathbb{R}^n\,.$$

$$\phi_\infty(A^* \overset{k}{r}) \geqslant \frac{\phi_2(A^* \overset{k}{r})}{\sqrt{n}}$$

$$\overset{k}{\alpha}{}_1^2 = \frac{\phi_\infty^2(A^* \overset{k}{r})}{\| \overset{k}{r} \|^2} \frac{1}{\| Ae_i \|^2} \geqslant \frac{\| A^* \overset{k}{r} \|_2^2}{n \| \overset{k}{r} \|^2} \frac{1}{\| Ae_i \|^2}$$

$$\| A^* \overset{k}{r} \|_2^2 = (A^* \overset{k}{r}, A^* \overset{k}{r})_2 = (A^* A\overset{k}{e}, A^* A\overset{k}{e})_2$$

$$= (Q^2 \overset{k}{e}, Q^2 \overset{k}{e})_2 = (Q\overset{k}{e}, Q^2 Q\overset{k}{e})_2$$

$$\| \overset{k}{r} \|^2 = (\overset{k}{r}, \overset{k}{r}) = (QA^{-1} \overset{k}{r}, QA^{-1} \overset{k}{r})_2 = (Q\overset{k}{e}, Q\overset{k}{e})_2 = \| Q\overset{k}{e} \|_2^2$$

$$\| Ae_i \|^2 = (Ae_i, Ae_i) = (Qe_i, Qe_i)_2 = \| Qe_i \|_2^2$$

$$\Rightarrow \overset{k}{\alpha}{}_1^2 \geqslant \frac{1}{n} \frac{(Q\overset{k}{e}, Q^2 Q\overset{k}{e})_2}{\| Q\overset{k}{e} \|_2^2} \frac{1}{\| Qe_i \|_2^2}.$$

Soient λ_1 et λ_n la plus grande et la plus petite valeur propre de Q

$$\frac{(Q\overset{k}{e}, Q^2 Q\overset{k}{e})_2}{\| Q\overset{k}{e} \|_2^2} \geqslant \lambda_n^2$$

$$\| Qe_i \|_2 \leqslant \| Q \|_2 \| e_i \|_2 = \| Q \|_2 = \lambda_1$$

$$\Rightarrow \overset{k}{\alpha}{}_1^2 \geqslant \frac{1}{n} \frac{\lambda_n^2}{\lambda_1^2} = \frac{1}{n\gamma_2^2(Q)} = \frac{1}{n\gamma_2(A^* A)}$$

$$A^* = A^T M \Rightarrow \gamma_2(A^* A) = \gamma_2(A^T MA) \leqslant \gamma_2(M)\, \gamma_2(A)\, \gamma_2(A^T)$$

$$\| A \|_2 = \| A^T \|_2 \Rightarrow \gamma_2(A) = \gamma_2(A^T)$$

$$\Rightarrow \gamma_2(A^* A) \leqslant \gamma_2(M)\, \gamma_2^2(A) = \gamma_2^2(R)\, \gamma_2^2(A)$$

$$\Rightarrow \overset{k}{\alpha}{}_1^2 \geqslant \frac{1}{n\gamma_2^2(R)\, \gamma_2^2(A)}.$$

Les hypothèses de la proposition 4.1.2 sont donc satisfaites avec

$$\delta = \frac{1}{\sqrt{n}\, \gamma_2(R)\, \gamma_2(A)}. \quad \square$$

Ce résultat montre que la vitesse de convergence est d'autant plus élevée que le conditionnement de A est petit.

Le vecteur $\overset{k}{y}$ est arbitraire et de nombreux choix simples sont possibles. On propose les deux cas particuliers suivants :

i) Corrections multiplicatives :

$$\overset{k}{y} = \overset{k}{x} \qquad \forall k .$$

On a alors

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \sum_{j \in I_i} \overset{k}{x}_j e_j \Rightarrow \overset{k+1}{x} = \sum_{i=1}^{p_k} (1 + \overset{k}{\mu}_i) \sum_{j \in I_i} \overset{k}{x}_j e_j .$$

Pour obtenir $\overset{k+1}{x}$ on multiplie donc des groupes de composantes de $\overset{k}{x}$ par les coefficients $1 + \overset{k}{\mu}_i$.

Dans ce cas le partitionnement est indispensable.

En effet, pour $p_k = 1$, $\overset{k+1}{r}$ est la projection de $\overset{k}{r}$ sur $(A\overset{k}{x})^\perp \Rightarrow (\overset{k+1}{r}, A\overset{k}{x}) = 0$

$$\overset{k+1}{x} = (1 + \overset{k}{\mu})\, \overset{k}{x} \Rightarrow (\overset{k+1}{r}, A\overset{k+1}{x}) = (1 + \overset{k}{\mu})\, (\overset{k+1}{r}, A\overset{k}{x}) = 0$$
$$\Rightarrow \overset{k+2}{r} = \overset{k+1}{r}$$

ce qui montre que $\overset{k+i}{x} = \overset{k+1}{x} \quad \forall i \geqslant 2$.

ii) Corrections additives :

$$\overset{k}{y} = \sum_{j=1}^{n} e_j \quad \forall k \qquad (\text{donc } \overset{k}{a}_j = 1 \quad \forall j = 1, \ldots, n) .$$

On a alors

$$\overset{k+1}{x} = \overset{k}{x} + \sum_{i=1}^{p_k} \overset{k}{\mu}_i \sum_{j \in I_i} e_j .$$

Pour obtenir $\overset{k+1}{x}$ on ajoute donc à des groupes de composantes de $\overset{k}{x}$ les coefficients $\overset{k}{\mu}_i$.

Ici encore le partitionnement est indispensable.

En effet, pour $p_k = 1$, y_k reste constant et donc aussi H_{p_k} au cours des itérations ; donc $\overset{k+2}{r} = \overset{k+1}{r}$ et $\overset{k+i}{x} = \overset{k+1}{x}$ $\forall i \geqslant 2$.

BIBLIOGRAPHIE

[1] O. AXELSSON, *Solution of linear systems of equations : iterative methods,* Sparse Matrix Techniques, Springer-Verlag, 1977, 1-51.

[2] J. BEUNEU, *Résolution des systèmes d'équations linéaires par la méthode des compensations,* Publication n° 69 du Laboratoire de Calcul de l'Université de Lille I, 1976.

[3] J. BEUNEU, *Un procédé d'accélération par projection d'itérations linéaires,* Publication ANO n° 47, Université de Lille I, 1981.

[4] J. BEUNEU, *Les méthodes de projection-minimisation pour les systèmes linéaires,* Publication ANO n° 58, Université de Lille I, 1981.

[5] J. BEUNEU, *Méthodes de projection à convergence finie ; remarques sur leur forme incomplète,* Publication ANO n° 80. Université de Lille I, 1982.

[6] F. CHATELIN, W. L. MIRANKER, *Acceleration by aggregation of successive approximation methods,* Linear Algebra and its applications, *43,* 1982, 17-47.

[7] C. ESPINOZA, *Contribution à la résolution numérique de certains systèmes d'équations,* Thèse de 3e cycle, Université de Grenoble, 1977.

[8] R. FROEHLICH, *A theoretical foundation for coarse mesh variational techniques,* General Atomic, Rapport n° GA-7870, 1967.

[9] N. GASTINEL, *Analyse numérique linéaire,* Hermann, Paris, 1966.

[10] B. GERMAIN-BONNE, *Etude de quelques problèmes d'accélération de convergence,* Publication n° 65 du Laboratoire de Calcul de l'Université de Lille I, 1976.

[11] A. S. HOUSEHOLDER, F. L. BAUER, *On certain iterative methods for solving linear systems,* Numerische Mathematik, *2,* 1960, 55-59.

[12] A. S. HOUSEHOLDER, *The theory of matrices in numerical analysis,* Blaisdell Publishing Company, 1964.

[13] S. KANIEL, J. STEIN, *Least-square acceleration of iterative methods for linear equations,* Journal of optimization theory and applications, *14,* 4 (1974), 431-437.

[14] M. A. KRASNOSELSKII, *Approximate solution of operator equations,* Wolters-Nordhoff, Gröningen, 1972.

[15] S. NAKAMURA, *A variational rebalancing method for linear iterative convergence schemes of neutron diffusion and transport equations,* Nuclear Sci. and Engineering, *39,* 1970, 278-283.

[16] S. NAKAMURA, *Coarse mesh acceleration of iterative solution of neutron diffusion equation,* Nuclear Sci. and Engineering, *43,* 1971, 116-120.

[17] D. RYAN, G. TRAPP, *Partitioning methods for accelerating Gauss-Seidel iterations,* Proceedings of West Virginia Academy of Science, *45,* 1973, 323-326.

[18] Y. SAAD, *Krylov subspace methods for solving large unsymmetric linear systems,* Math. Comp. *37,* 1981, 105-126.

[19] Y. SAAD, *The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems,* SIAM Journ. Numer. Anal. *19,* 1982, 485-506.

[20] A. SETTARI, K. AZIZ, *A generalization of the additive correction methods for the iterative solution of matrix equations,* SIAM J. Numer. Anal., *10,* 3 (1973) 506-521.

[21] Y. VOROBYEV, *Method of moments in applied mathematics,* Gordon and Breach, New York, 1965.

[22] E. L. WACHSPRESS, *Iterative solution of elliptic systems and applications to the neutron diffusion equations of reactor physics,* Prentice-Hall, Englewood Cliffs, N. J., 1966.