

F. ROBERT

Algorithmes tronques de découpe linéaire

Revue française d'automatique informatique recherche opérationnelle. Mathématiques, tome 6, n° R2 (1972), p. 45-64

http://www.numdam.org/item?id=M2AN_1972__6_2_45_0

© AFCET, 1972, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique informatique recherche opérationnelle. Mathématiques » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ALGORITHMES TRONQUES DE DECOUPE LINEAIRE

par F. ROBERT ⁽¹⁾

Résumé. — *Un théorème dit « de Stein-Rosenberg tronqué » est établi, qui permet d'étudier la convergence d'un algorithme tronqué de découpe linéaire. Applicables mais onéreux, pour des systèmes linéaires, ces résultats constituent en fait une base théorique pour des algorithmes de résolution d'équations de point fixe dans R^n , non linéaires mais contractantes en norme vectorielle.*

INTRODUCTION

A partir des notions de *découpe* de matrice et de *norme compatible* avec une découpe (*) est défini, dans [2], un algorithme dit « de découpe linéaire » opérant de la façon suivante :

Partant d'une équation de point fixe linéaire dans R^n , l'algorithme produit une suite d'équations de point fixe équivalentes, et qui « convergent » vers une équation limite

$$x = Ux + h$$

équivalente à celle de départ, et pouvant être beaucoup plus simple à résoudre : le choix de la découpe particulière utilisée permet d'imposer à U d'avoir des éléments nuls « où l'on veut ».

La caractéristique de cet algorithme est de transformer, de façon itérative, l'opérateur linéaire dont on part. Il ne rentre donc :

— ni dans la classe des « méthodes directes » (Gauss, Jordan, Choleski) qui transforment bien l'opérateur de départ, mais en un nombre fini d'opérations;

(1) Analyse Numérique, Département de Mathématiques, Université de Lyon I-Villeurbanne.

(*) Notions rappelées brièvement à la fin de cette introduction.

— ni dans la classe des « méthodes itératives » d'itération linéaire (Jacobi, Gauss-Seidel, relaxation...) qui procèdent bien de façon itérative, mais dans l'espace R^n où l'on cherche la solution, et non pas sur l'opérateur.

L'intérêt principal de cet algorithme est de pouvoir s'étendre, quant aux résultats de base, et de façon très naturelle, au traitement (formel) d'équations de point fixe sur R^n , *non linéaires mais contractantes en norme vectorielle* [1]. L'algorithme obtenu a alors été baptisé « algorithme de découpe non linéaire ». Précisément, dans l'algorithme non linéaire, c'est sur la matrice de contraction de l'équation donnée que s'applique l'algorithme de découpe linéaire, pour produire la suite des matrices de contraction des équations de point fixe successivement construites.

Le but du présent travail est alors le suivant : définir et étudier la convergence, pour le cas linéaire, d'algorithmes de découpe *tronqués*, basés sur la remarque suivante :

L'algorithme de découpe linéaire consiste, à chaque pas, à calculer $(I - L)^{-1}U$ à partir de deux matrices L et U obtenues au pas précédent [2].

Dans le contexte de l'algorithme, $(I - L)^{-1}U$ coïncide avec $\left[\sum_{k=0}^{\infty} L^k \right] U$.

L'algorithme *tronqué* de découpe linéaire, va consister à remplacer, dans un but d'économie de calcul, $(I - L)^{-1}U$ par une approximation, égale à

$$[I + L + \dots L^k]U + L^{k+1}$$

l'entier k , supérieur ou égal à 1, pouvant varier arbitrairement à chaque pas. La *troncature maximum*, qui consiste à prendre systématiquement $k = 1$, revient donc à remplacer $(I - L)^{-1}U$ par l'approximation suivante

$$(I + L)U + L^2$$

Nous montrons alors, dans ce travail, que, sous la même condition de convergence établie pour l'algorithme de découpe linéaire [à savoir : $N(G_0) < 1$, G_0 étant la matrice de départ, et N une norme compatible avec la découpe utilisée] on obtient encore les mêmes résultats pour l'algorithme tronqué, à savoir :

1) que cet algorithme produit encore une suite d'équations linéaires de point fixe équivalentes à celle de départ;

2) que cette suite d'équations « converge » vers une équation limite, encore équivalente à celle de départ, et dont la matrice U a des éléments nuls placés à volonté, selon le choix de la découpe utilisée : l'équation limite peut donc être très simple à résoudre, selon la forme imposée a priori à U : triangulaire par exemple.

Cette convergence de l'algorithme tronqué sous la même condition que pour l'algorithme non tronqué prouve donc que dans le calcul de $(I - L)^{-1}U$ à

chaque pas, on peut s'en tenir à une approximation même très grossière (par exemple $L^2 + LU + U$) : on rejoint par là un résultat valable pour d'autres problèmes résolus de façon itérative : on peut souvent sans dommage s'en tenir à une résolution très grossière du problème posé à chaque pas, l'essentiel étant de faire un calcul — même très approché — et d'itérer le processus.

Enfin, l'étude de la convergence de l'algorithme tronqué nous a amenés à élaborer un « théorème de Stein-Rosenberg tronqué » (théorème 2 dans le texte) dont on pense qu'il présente un intérêt intrinsèque : nous l'avons établi de façon autonome en début de papier.

Là encore, l'intérêt réel de l'algorithme (tronqué) de découpe linéaire n'est pas tant dans la résolution numérique d'équations linéaires de point fixe (le coût du calcul reste élevé) que dans son extension à des équations non linéaires de point fixe sur R^n , contractantes en norme vectorielle, particulièrement pour le cas, numériquement intéressant, de découpes implicites [1].

Pour l'immédiat, le présent travail se réfère uniquement aux notions élaborées dans [2], et que nous rappelons très brièvement :

M_n désignant l'algèbre des matrices réelles (n, n) , on appelle *découpe* de M_n en $\mathcal{L} + \mathcal{U}$, toute décomposition de M_n en la somme directe des sous-espaces \mathcal{L} et \mathcal{U} choisis ainsi : \mathcal{L} et \mathcal{U} sont engendrés par une partition de la base canonique de M_n en 2 sous-ensembles. Autrement dit, selon une découpe s donnée, toute matrice G de M_n se décompose (de façon unique) en $L + U$, où L est formée de certains éléments de G (au choix de la découpe s), les autres étant nuls. U complète L pour donner $G = L + U$.

Une découpe s étant donnée sur M_n , une norme N sur M_n est alors dite *compatible* avec s si elle vérifie les deux axiomes suivants :

- 1) N est sous-multiplicative : $\forall A, B \in M_n, N(AB) \leq N(A)N(B)$
- 2) pour tout G de M_n décomposé selon s en $L + U$, on a

$$N(G) = N(L) + N(U)$$

On renvoie, pour l'existence et l'étude de telles normes, ainsi que pour les résultats rappelés ici, à [2].

En ce qui concerne les notations, un vecteur de R^n est dit ≥ 0 si toutes ses composantes sont ≥ 0 ; ≥ 0 s'il est ≥ 0 et $\neq 0$; > 0 si toutes ses composantes sont > 0 . Un vecteur ≥ 0 est dit non négatif; un vecteur > 0 est dit positif.

De façon analogue, une matrice S est dite ≥ 0 (non négative) si tous ses éléments sont ≥ 0 . On ne considérera que des matrices carrées. Si S appartient à M_n , $\rho(S)$ désigne le rayon spectral de S , à savoir le module maximum des valeurs propres de S . On rappelle que si N est une norme sous multiplicative sur M_n (et en particulier si elle est compatible avec une découpe donnée sur M_n), on a, pour toute matrice $A \in M_n$:

$$\rho(A) \leq N(A)$$

I. GENERALITES

§ 1. Rappels

Pour toute matrice carrée $S \geq 0$, et tout vecteur $u \neq 0$, on pose [3]

$$\eta_S(u) = \sup_{\alpha \geq 0} (\alpha u \leq Su)$$

$$\tau_S(u) = \inf_{\beta \geq 0} (Su \leq \beta u) \quad (\text{par convention, } \inf(\emptyset) = +\infty)$$

d'où

$$0 \leq \eta_S(u) \leq \tau_S(u)$$

Pour $u > 0$, il vient

$$\eta_S(u) = \min_i \frac{[Su]_i}{u_i} \quad \text{et} \quad \tau_S(u) = \max_i \frac{[Su]_i}{u_i} = S_{\infty}(U^{-1}SU)$$

avec $U = \text{diag}(u)$, S_{∞} désignant la norme de matrice engendrée par la norme du max.

On rappelle alors la formulation suivante du théorème de Perron-Frobenius.

Théorème 1 [3]

Soit S une matrice ≥ 0 et carrée. Son rayon spectral $\rho(S)$ est valeur propre et il lui correspond un vecteur propre $\omega \geq 0$. On a alors

$$\sup_{\substack{u \geq 0 \\ u \neq 0}} \eta_S(u) = \rho(S) = \inf_{u > 0} \tau_S(u) \quad (1)$$

avec évidemment : $\eta_S(\omega) = \rho(S)$: la borne supérieure est atteinte ⁽¹⁾.

REMARQUES

1) Nous n'énonçons pas de résultat plus précis concernant le cas où S est irréductible : seul le résultat le plus général (1) interviendra dans la suite.

2) On retrouve, grâce à (1) le résultat (classique) suivant : si S , matrice carrée ≥ 0 , admet un vecteur propre $\omega > 0$, ω est nécessairement attaché à une valeur propre λ égale au rayon spectral $\rho(S)$.

En effet, il vient $\tau_S(\omega) = \lambda$

or a priori

$$\tau_S(\omega) \geq \rho(S) \geq |\lambda|$$

donc

$$\lambda = \rho(S)$$

(1) On peut montrer que la borne inférieure est atteinte si et seulement si il existe $u > 0$ tel que $Su \leq \rho(S)u$; c'est vrai en particulier si S est irréductible.

Corollaire 1.

Pour toute matrice carrée $S \geq 0$ et tout $\varepsilon > 0$, il existe $u_\varepsilon > 0$ tel que

$$Su_\varepsilon \leq [\rho(S) + \varepsilon]u_\varepsilon \quad (2)$$

En effet, puisque $\rho(S) = \inf_{u > 0} \tau_S(u)$, on a, pour tout $\varepsilon > 0$, l'existence de $u_\varepsilon > 0$ tel que

$$\rho(S) \leq \tau_S(u_\varepsilon) \leq \rho(S) + \varepsilon/2$$

Or

$$\tau_S(u_\varepsilon) = \inf (\beta \geq 0 : Su_\varepsilon \leq \beta u_\varepsilon)$$

il existe donc $\beta_\varepsilon \geq 0$ tel que

$$\tau_S(u_\varepsilon) \leq \beta_\varepsilon \leq \tau_S(u_\varepsilon) + \varepsilon/2$$

et

$$Su_\varepsilon \leq \beta_\varepsilon u_\varepsilon$$

d'où

$$Su_\varepsilon \leq [\rho(S) + \varepsilon]u_\varepsilon$$

REMARQUE

On pourrait, de façon analogue, démontrer, pour toute matrice carrée $S \geq 0$ et tout $\varepsilon > 0$, l'existence de $v_\varepsilon \neq 0$ tel que :

$$[\rho(S) - \varepsilon]v_\varepsilon \leq Sv_\varepsilon \quad (3)$$

Mais en fait il suffit de prendre ω , vecteur propre $\neq 0$ de S , attaché à la valeur propre $\rho(S)$ (théorème 1) pour avoir :

$$\rho(S)\omega = S\omega \quad (4)$$

ce qui redonne (3) avec en plus le fait que ω ne dépend pas de ε (et l'égalité au lieu de l'inégalité).

§ 2. Un théorème du type « Stein-Rosenberg-tronqué »

Soient L et U deux matrices carrées ≥ 0 . On pose

$$S_0 = L + U \geq 0$$

et

$$S_{k+1} = LS_k + U \quad (k = 0, 1, \dots)$$

Il vient

$$S_k = [I + L + \dots L^k]U + L^{k+1} \geq 0.$$

On pose $\rho_k = \rho(S_k)$.

On a alors le

Théorème 2.

- 1) Si $\rho_0 = 0$ alors, pour tout k , $\rho_k = 0$, et toutes les matrices S_k admettent un vecteur propre commun $\omega \geq 0$ attaché à la valeur propre 0 ⁽¹⁾.
- 2) Si $0 < \rho_0 < 1$ alors, pour tout k , $0 < \rho_0^{k+1} \leq \rho_k \leq \rho_0 < 1$.
- 3) Si $\rho_0 = 1$ alors, pour tout k , $\rho_k = 1$ et toutes les matrices S_k admettent un vecteur propre commun $\omega \geq 0$ attaché à la valeur propre 1.
- 4) Si $\rho_0 > 1$, alors, pour tout k , $1 < \rho_0 \leq \rho_k \leq \rho_0^{k+1}$.

Preuve

Pour tout $\varepsilon > 0$, il existe (Corollaire 1) $u_\varepsilon > 0$ tel que

$$S_0 u_\varepsilon \leq [\rho_0 + \varepsilon] u_\varepsilon$$

d'autre part, si $\omega \neq 0$ est un vecteur propre de S_0 attaché à la valeur propre ρ_0 , on a

$$\rho_0 \omega = S_0 \omega$$

Alors

- 1) si $\rho_0 = 0$ limitons ε à varier dans $]0, 1[$ et montrons que

$$\forall k, \forall \varepsilon \in]0, 1[\quad S_k u_\varepsilon \leq \varepsilon u_\varepsilon \quad (5)$$

Ce résultat est vrai pour $k = 0$. Le supposant vrai pour k , il vient

$$S_{k+1} u_\varepsilon = L S_k u_\varepsilon + U u_\varepsilon \leq \varepsilon L u_\varepsilon + U u_\varepsilon \leq [L + U] u_\varepsilon = S_0 u_\varepsilon \leq \varepsilon u_\varepsilon$$

par récurrence, le résultat (5) est établi. D'où il résulte que

$$0 \leq \rho_k \leq \tau_{S_k}(u_\varepsilon) \leq \varepsilon$$

Ainsi pour tout k , et tout $\varepsilon \in]0, 1[$ $0 \leq \rho_k \leq \varepsilon$ d'où il résulte que, pour tout k , $\rho_k = 0$.

D'autre part, montrons que pour tout k , $S_k \omega = 0$: Vrai pour $k = 0$, ce résultat se démontre par récurrence grâce aux inégalités

$$0 \leq S_{k+1} \omega = L S_k \omega + U \omega = U \omega \leq [L + U] \omega = S_0 \omega = 0$$

et le point 1) est établi.

- 2) Si $0 < \rho_0 < 1$. Limitons $\varepsilon > 0$ à être tel que $\rho_0 + \varepsilon \leq 1$. Montrons alors que pour tout k

$$S_k u_\varepsilon \leq [\rho_0 + \varepsilon] u_\varepsilon \quad (6)$$

(1) Et tous les S_k sont réductibles, une matrice ≥ 0 irréductible ne pouvant avoir son rayon spectral nul.

Ce résultat, vrai pour $k = 0$, se démontre par récurrence de la même façon que précédemment, grâce aux inégalités :

$$S_{k+1}u_\varepsilon = LS_k u_\varepsilon + Uu_\varepsilon \leq (\rho_0 + \varepsilon)Lu_\varepsilon + Uu_\varepsilon \leq [L + U]u_\varepsilon = S_0 u_\varepsilon \leq [\rho_0 + \varepsilon]u_\varepsilon$$

Montrons également que pour tout k

$$\rho_0^{k+1}\omega \leq S_k \omega \quad (7)$$

Ce résultat est vrai pour $k = 0$. Le supposant vrai pour k , il vient

$$S_{k+1}\omega = LS_k \omega + U\omega \geq \rho_0^{k+1}L\omega + U\omega \geq \rho_0^{k+1}[L + U]\omega = \rho_0^{k+1}S_0\omega = \rho_0^{k+2}\omega$$

Il en résulte que pour tout k , et tout $\varepsilon > 0$ assez petit :

$$\rho_0^{k+1} \leq \eta_{S_k}(\omega) \leq \rho_k \leq \tau_{S_k}(u_\varepsilon) \leq \rho_0 + \varepsilon$$

d'où

$$\rho_0^{k+1} \leq \rho_k \leq \rho_0$$

et le point 2) est établi.

3) Si $\rho_0 = 1$.

Montrons que pour tout k , et tout $\varepsilon > 0$

$$S_k u_\varepsilon \leq (1 + \varepsilon)^{k+1} u_\varepsilon \quad (8)$$

Vrai pour $k = 0$, ce résultat se démontre par récurrence grâce aux inégalités :

$$\begin{aligned} S_{k+1}u_\varepsilon &= LS_k u_\varepsilon + Uu_\varepsilon \leq (1 + \varepsilon)^{k+1}Lu_\varepsilon + Uu_\varepsilon \\ &\leq (1 + \varepsilon)^{k+1}[L + U]u_\varepsilon = (1 + \varepsilon)^{k+1}S_0 u_\varepsilon \leq (1 + \varepsilon)^{k+2}u_\varepsilon \end{aligned}$$

Montrons également que, pour tout k

$$S_k \omega = \omega \quad (9)$$

Vrai pour $k = 0$, ce résultat se démontre par récurrence grâce aux égalités $S_{k+1}\omega = LS_k \omega + U\omega = L\omega + U\omega = S_0\omega = \omega$. ω est donc vecteur propre commun à tous les S_k , et attaché à la valeur propre commune 1 (1).

Il en résulte que, pour tout k , et tout $\varepsilon > 0$

$$1 = \eta_{S_k}(\omega) \leq \rho_k \leq \tau_{S_k}(u_\varepsilon) \leq (1 + \varepsilon)^{k+1}$$

(1) On ne peut en déduire par ce simple argument que 1 est rayon spectral des S_k car $\omega \neq 0$, n'est pas nécessairement > 0 .

d'où, pour tout k

$$\rho_k = 1$$

4) Si $\rho_0 > 1$.

Montrons que, pour tout k , et tout $\varepsilon > 0$

$$S_k u_\varepsilon \leq (\rho_0 + \varepsilon)^{k+1} u_\varepsilon \quad (10)$$

Ce résultat, vrai pour $k = 0$, se démontre par récurrence grâce aux inégalités

$$\begin{aligned} S_{k+1} u_\varepsilon &= L S_k u_\varepsilon + U u_\varepsilon \leq (\rho_0 + \varepsilon)^{k+1} L u_\varepsilon + U u_\varepsilon \leq (\rho_0 + \varepsilon)^{k+1} [L + U] u_\varepsilon \\ &= (\rho_0 + \varepsilon)^{k+1} S_0 u_\varepsilon \leq (\rho_0 + \varepsilon)^{k+2} u_\varepsilon \end{aligned}$$

Montrons de plus que, pour tout k

$$\rho_0 \omega \leq S_k \omega \quad (11)$$

Vrai pour $k = 0$, ce résultat se démontre par récurrence grâce aux inégalités :

$$S_{k+1} \omega = L S_k \omega + U \omega \geq \rho_0 L \omega + U \omega \geq [L + U] \omega = S_0 \omega \geq \rho_0 \omega$$

Ainsi, pour tout k et tout $\varepsilon > 0$, il vient

$$\rho_0 \leq \eta_{S_k}(\omega) \leq \rho_k \leq \tau_{S_k}(u_\varepsilon) \leq (\rho_0 + \varepsilon)^{k+1}$$

d'où, pour tout k

$$\rho_0 \leq \rho_k \leq \rho_0^{k+1}$$

et le théorème 2 est démontré :

REMARQUE. Si $L = 0$, ce résultat devient trivial, puisqu'alors tous les S_k coïncident avec $S_0 = U$.

Autre cas élémentaire : si $U = 0$. On a alors $S_k = L^{k+1}$.

Notons que dans ces deux cas triviaux, le résultat est vrai même si U , (resp. L) ne sont pas ≥ 0 : on peut les prendre quelconques.

Le théorème précédent a pour conséquence le théorème classique de comparaison des rayons spectraux de $L + U$ et $(I - L)^{-1}U$ pour L et $U \geq 0$ (Stein-Rosenberg) [4].

En effet :

a) Si $\rho_0 = 0$

alors d'une part

$$0 \leq \rho(L) \leq \rho(L + U) = \rho_0 = 0 < 1$$

donc $(I - L)^{-1}U$ existe et coïncide avec

$$\sum_{i=0}^{\infty} L^i U$$

autrement dit

$$\lim_{k \rightarrow \infty} S_k = (I - L)^{-1}U$$

Or $\rho_k = 0$ d'où $\rho[(I - L)^{-1}U] = 0$.

b) Si $0 < \rho_0 < 1$

on a encore l'existence de $(I - L)^{-1}U$ qui est limite des S_k .

Or

$$\rho_0^{k+1} \leq \rho_k \leq \rho_0 < 1$$

d'où

$$\rho[(I - L)^{-1}U] \leq \rho(L + U) < 1$$

c) Si $\rho_0 = 1$

et si $\rho(L) < 1$, $(I - L)^{-1}U$ existe et est limite des S_k

or

$$\rho_k = 1 \quad \text{d'où} \quad \rho[(I - L)^{-1}U] = 1$$

d) Si $\rho_0 > 1$

et si $\rho(L) < 1$, $(I - L)^{-1}U$ existe et est limite des S_k .

or

$$1 < \rho_0 \leq \rho_k \leq \rho_0^{k+1}$$

d'où $1 < \rho(L + U) \leq \rho[(I - L)^{-1}U]$

D'autre part, nous utiliserons le théorème 2 dans l'étude des algorithmes tronqués de découpe linéaire :

II. ALGORITHME TRONQUE DE DECOUPE LINEAIRE

Soit s une *découpe* de M_n en $\mathcal{L} \oplus \mathcal{U}$ et $G_0 \in M_n$.

Rappelons les résultats concernant 1'.

§ 1. Algorithme de découpe linéaire associé à s et partant de G_0 [2]

On envisage la suite suivante dans M_n :

pour G_r décomposé selon s en $L_r + U_r$, on pose, si elle existe,

$$G_{r+1} = (I - L_r)^{-1}U_r \quad (r = 0, 1, 2 \dots)$$

On a alors le résultat suivant d'existence et de convergence :

Théorème 3 [2]

Soit N une norme compatible avec la découpe s

si $N(G_0) < 1$ alors :

1) G_r existe pour tout r ($I - L_r$ inversible pour tout r)

2) G_r converge vers une limite $U \in \mathcal{U}$, et l'on a :

$$N(U) \leq \dots \leq N(G_{r+1}) \leq N(G_r) \dots \leq N(G_0) < 1$$

3) Si $G_0 \geq 0$, G_r est ≥ 0 pour tout r et

$$\rho(U) \leq \dots \rho(G_{r+1}) \leq \rho(G_r) \dots \leq \rho(G_0) \leq N(G_0) < 1$$

$$U_0 \leq \dots U_r \leq U_{r+1} \leq \dots U$$

d'où

$$\rho(U_0) \leq \dots \rho(U_r) \leq \rho(U_{r+1}) \leq \dots \rho(U)$$

$\rho(U)$ est donc limite des deux suites (monotones) $\rho(G_r)$ et $\rho(U_r)$.

REMARQUE

Le calcul de $(I - L_r)^{-1}U_r$ est onéreux.

On se propose de modifier l'algorithme précédent en prenant comme approximation de $(I - L)^{-1}U$ (pour $\rho(L) < 1$)

$$A = [I + L + \dots L^k]U + L^{k+1} \quad k \text{ entier } \geq 1 \text{ donné}$$

Cette approximation est calculable de la façon suivante :

$$S_0 = L + U$$

$$S_0 = LS_0 + U$$

$$S_{p+1} = LS_p + U \quad (p = 0, 1, \dots, k-1)$$

alors

$$A = S_k$$

On définit 1'.

§ 2. Opérateur de découpe, tronqué au rang k , associé à s

La découpe s étant donnée sur M_n , tout H de M_n est décomposé selon s en $H = L + U$.

k étant un entier ≥ 0 , on pose

$$F_k(H) = [I + L + \dots L^k]U + L^{k+1} \quad (12)$$

Cet opérateur F_k est défini sur tout M_n ; pour $k = 0$, il redonne l'application identique.

Pour étudier les propriétés de F_k , nous avons besoin de résultats préliminaires :

Lemme 1

Soient l , et $u \geq 0$ dans R .

Posons

$$\theta_k = (1 + l + \dots l^k)u + l^{k+1} \quad [\theta_0 = l + u]$$

alors

- 1) ou bien $l = 0$ et alors $\forall k \quad \theta_k = \theta_0 = u$
- 2) ou bien $l > 0$. Alors, pour tout k , θ_k est > 0 et l'on a
 - $\alpha)$ si $0 < \theta_0 < 1$ alors $\forall k \quad 0 < \theta_{k+1} < \theta_k (< \theta_0 < 1)$
 - $\beta)$ si $\theta_0 = 1$ alors $\forall k \quad \theta_k = 1$
 - $\gamma)$ si $\theta_0 > 1$ alors $\forall k \quad \theta_{k+1} > \theta_k (> \theta_0 > 1)$

Preuve

Le point 1) est évident; pour 2) remarquons que

$$\theta_k - \theta_{k+1} = -l^{k+1}u + l^{k+1} - l^{k+2} = l^{k+1}[1 - (l + u)]$$

d'où $\alpha)$ $\beta)$ $\gamma)$.

REMARQUES

1) dans le cas 2 — $\alpha)$ les θ_k décroissent strictement vers leur limite

$$\theta = \frac{u}{1-l} < \theta_0 \quad (\text{car } \theta_0 < 1 \Rightarrow l < 1)$$

dans le cas 2 — $\gamma)$ les θ_k croissent strictement vers

$$\theta = \frac{u}{1-l} > \theta_0 \quad \text{si } l < 1$$

$$+ \infty \quad \text{si } l \geq 1$$

2) Ce lemme est très analogue au théorème 2, mais n'en est pas cas particulier. En appliquant le théorème 2) aux matrices (1.1) l et $u \geq 0$, on peut compléter le lemme par les indications suivantes :

- $\alpha)$ si $0 < \theta_0 < 1$ alors $\forall k \quad \theta_0^{k+1} \leq \theta_k \leq \theta_0 < 1$
- $\gamma)$ si $1 < \theta_0$ alors $\forall k \quad 1 < \theta_0 \leq \theta_k \leq \theta_0^{k+1}$

Lemme 2

Soit N une norme compatible avec la découpe s .

Alors si $N(H) < 1$, il vient pour tout entier $k \geq 1$

$$N(F_k(H)) \leq N(H) < 1 \quad (13)$$

plus précisément :

ou bien $L = 0$ et $\forall k \quad F_k(H) = H = U \in \mathcal{U}$

ou bien $L \neq 0$ et $\forall k \geq 1 \quad N(F_k(H)) < N(H) < 1$

Preuve

$$\begin{aligned} N(F_k(H)) &= N[U + LU + \dots L^k U + L^{k+1}] \\ &\leq N(U) + N(L)N(U) + \dots [N(L)]^k N(U) + [N(L)]^{k+1} \\ &= (1 + N(L) + \dots [N(L)]^k) N(U) + (N(L))^{k+1}. \end{aligned}$$

Posons $l = N(L)$; $u = N(U)$. Il vient avec les notations du lemme 1

$$N(F_k(H)) \leq \theta_k$$

Or $N(H) = N(L) + N(U) = l + u = \theta_0 < 1$. Il vient, d'après le lemme 1 :

ou bien $l = 0$ c'est-à-dire $L = 0$ et alors $F_k(H) = U = H \in \mathcal{U}$

ou bien $l \neq 0$ c'est-à-dire $L \neq 0$ et alors

$$N(F_k(H)) \leq \theta_k < \theta_0 = N(H)$$

Définissons alors 1'.

§ 3. Algorithme tronqué de découpe linéaire

On se donne une suite $\{k_0, k_1 \dots k_r \dots\}$ d'entiers ≥ 1 ; alors, partant de H_0 pris dans M_n , décomposé selon s en $L_0 + U_0$, on définit

$$H_{r+1} = F_{k_r}(H_r) \quad (r = 0, 1, 2 \dots)$$

Cet algorithme sera appelé *algorithme tronqué de découpe linéaire, partant de H_0 , et associé à la découpe s et à la suite $\{k_r\}$* . Le résultat suivant règle sa convergence.

Théorème 4

Soit s une découpe de M_n en $\mathcal{L} \oplus \mathcal{U}$, N une norme compatible avec s , et H_0 pris dans M_n .

Si $N(H_0) < 1$, alors, quelle que soit la suite $k_0, k_1, \dots k_r$ d'entiers ≥ 1 , l'algorithme de découpe tronqué associé à s et aux $\{k_r\}$ et partant de H_0 ; vérifie les propriétés suivantes :

1) H_r converge vers une limite U appartenant à \mathcal{U}

$$2) N(U) \leq \dots N(H_{r+1}) \leq N(H_r) \leq \dots N(H_0) < 1 \quad (14)$$

3) si $H_0 \geq 0$, alors tous les H_r sont ≥ 0 et

$$\rho(U) \leq \dots \leq \rho(H_{r+1}) \leq \rho(H_r) \leq \dots \leq \rho(H_0) \leq N(H_0) < 1 \quad (15)$$

$$0 \leq U_0 \leq U_1 \leq \dots U_r \leq U_{r+1} \leq \dots U \quad (16)$$

d'où

$$\rho(U_0) \leq \rho(U_1) \leq \dots \leq \rho(U_r) \leq \rho(U_{r+1}) \leq \dots \rho(U) \quad (17)$$

$\rho(U)$ est donc limite des deux suites (monotones) $\rho(H_r)$ et $\rho(U_r)$

Preuve

1) Puisque $H_{r+1} = F_{k_r}(H_r)$ et que $N(H_0) < 1$, les inégalités

$$N(H_{r+1}) \leq N(H_r) \leq \dots N(H_0) < 1$$

résultent, par récurrence, du lemme 2. Comme on va montrer que H_r converge vers une limite U , les inégalités (14) seront acquises.

2) Démontrons cette convergence, et que la limite U des H_r appartient à \mathcal{U} . Pour cela, formons

$$\begin{aligned} H_{r+1} - U_r &= L_{r+1} + U_{r+1} - U_r \\ &= [I + L_r + \dots L_r^{k_r}]U_r + L_r^{k_r+1} - U_r \\ &= L_r[(I + L_r + \dots L_r^{k_r-1})U_r + L_r^{k_r}] = L_r P_r \end{aligned}$$

passant aux normes, il vient, puisque N est compatible avec s

$$N(L_{r+1}) + N(U_{r+1} - U_r) \leq N(L_r)N(P_r)$$

or

$$P_r = F_{k_r-1}(H_r) \quad \text{et} \quad N(H_r) < 1$$

donc, par application du lemme 2, il vient, si $k_r \geq 2$

$$N(P_r) \leq N(H_r) \leq N(H_0) < 1$$

et si $k_r = 1$ $P_r = H_r$, d'où $N(P_r) = N(H_r) \leq N(H_0) < 1$.

On a donc, dans tous les cas

$$N(L_{r+1}) + N(U_{r+1} - U_r) \leq N(L_r)N(H_0) \quad \text{avec} \quad N(H_0) < 1$$

Donc

$$N(L_{r+1}) \leq N(L_r)N(H_0) \leq [N(H_0)]^{r+1}N(L_0) \quad \text{avec} \quad N(H_0) < 1$$

ce qui montre que $N(L_r)$, donc L_r , converge vers 0.

D'autre part

$$N(U_{r+1} - U_r) \leq N(L_r)N(H_0) \leq [N(H_0)]^{r+1}N(L_0) \quad \text{avec} \quad N(H_0) < 1$$

ce qui prouve que la suite U_r est de Cauchy dans \mathcal{U} , qui est complet : donc U_r converge vers une limite $U \in \mathcal{U}$.

Alors $H_r = L_r + U_r$ converge vers U .

Ainsi les points 1) et 2) sont établis.

Si de plus H_0 est ≥ 0 , L_0 et U_0 de même, d'après les propriétés d'une découpe. Donc H_1 est ≥ 0 et tous les H_r de même.

On a alors

$$H_{r+1} = L_{r+1} + U_{r+1} = U_r + L_r U_r + \dots L_r^{k_r} U_r + L_r^{k_r+1}$$

Ainsi

$$H_{r+1} = L_{r+1} + U_{r+1} \geq U_r$$

d'où résulte, d'après les propriétés de non superposition de L_{r+1} et U_{r+1} , que

$$U_{r+1} \geq U_r \geq 0$$

et la limite U est \geq à tous les U_r , d'où (16).

On déduit les mêmes inégalités sur les rayons spectraux de ces matrices non négatives, d'où (17).

Enfin, puisque L_r et U_r sont ≥ 0 et que $\rho(H_r) \leq N(H_r) < 1$, on déduit du théorème 2 que

$$\rho[H_{r+1}] = \rho[(I + L_r + \dots L_r^{k_r})U_r + L_r^{k_r+1}] \leq \rho(H_r) < 1$$

ce qui établit les inégalités (15).

D'ailleurs, le théorème 2 permet de préciser que, pour tout r

$$[\rho(H_r)]^{k_r+1} \leq \rho(H_{r+1}) \leq \rho(H_r) < 1 \quad (18)$$

Le théorème 4 est donc démontré.

REMARQUES

1) pour $k_r = 0$, l'algorithme stationne en H_0 ;

2) Soit $|H_0|$ la matrice obtenue en remplaçant dans H_0 tous les éléments par leur valeur absolue. Dans l'algorithme ci-dessus, soit H'_r la suite de matrices engendrée à partir de $H'_0 = |H_0|$, et H_r la suite de matrices engendrées à partir de H_0 .

Alors pour tout r $0 \leq |H_r| \leq H'_r$

ce qui s'écrit, pour H_r décomposé en $L_r + U_r$, et H'_r en $L'_r + U'_r$ selon s :

$$0 \leq |L_r| \leq L'_r \quad ; \quad 0 \leq |U_r| \leq U'_r$$

En effet, cette relation est vraie pour $r = 0$: $|L_0| = L'_0$; $|U_0| = U'_0$.

Supposons la vraie au rang r , il vient :

$$\begin{aligned} 0 \leq |H_{r+1}| &= |[I + L_r + \dots L_r^{k_r}]U_r + L_r^{k_r+1}| \\ &\leq |[I + L_r + \dots L_r^{k_r}]| |U_r| + |L_r|^{k_r+1} \\ &\leq [I + |L_r| + \dots |L_r|^{k_r}] |U_r| + |L_r|^{k_r+1} \\ &\leq [I + L'_r + \dots L_r'^{k_r}]U'_r + L_r'^{k_r+1} = H'_r \end{aligned}$$

c'est-à-dire

$$0 \leq |L_{r+1}| \leq L'_{r+1} \quad ; \quad 0 \leq |U_{r+1}| \leq U'_{r+1}.$$

En particulier, si, relativement à une norme N compatible avec la découpe s , on a à la fois $N(H_0) < 1$ et $N(H'_0) < 1$ (par exemple parce que $N(H_0) < 1$ et que N est absolue; alors $N(|H_0|) = N(H_0) < 1$), il y aura convergence de H'_r vers $U \in \mathcal{U}$, de H' vers $U' \in \mathcal{U}$, et l'on aura

$$0 \leq |U| \leq U'$$

d'où, d'ailleurs, la même inégalité sur les rayons spectraux de ces matrices non-négatives; on a donc, finalement :

$$\rho(U) \leq \rho(|U|) \leq \rho(U') < 1$$

EXEMPLES

1) Reprenons la matrice H_0 donnée dans l'introduction de [2].

$$H_0 = \frac{1}{100} \begin{vmatrix} 4 & -2 & 4 & 8 & -3 & 1 \\ -1 & 2 & 6 & 3 & 1 & 2 \\ 4 & 6 & 3 & 6 & -4 & -2 \\ -2 & 1 & 2 & 5 & -5 & 1 \\ 5 & 2 & 1 & -6 & 7 & -3 \\ -2 & -1 & 1 & 3 & -4 & 2 \end{vmatrix}$$

Sur M_6 , on sait [2] que la norme N définie par

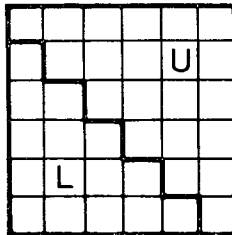
$$N(H) = \max_i |h_{ii}| + \sum_{\substack{i,j \\ i \neq j}} |h_{ij}|$$

est compatible avec toute découpe intègre, c'est-à-dire toute découpe ne scindant pas la diagonale (ce qui revient à dire que la matrice unité s'y décompose soit en $I + 0$, soit en $0 + I$);

et on vérifie que $N(H_0) < 1$: les algorithmes tronqués de découpe seront donc convergents pour toutes les découpes intègres. Voici quelques résultats numériques, pour de telles découpes.

Dans tous ces exemples, la suite $\{k_r\}$ a été prise stationnaire en la valeur (entière, ≥ 1) k . Différentes valeurs ont été essayées pour k : 1, 5, 10, 20, $+\infty$ (ce qui correspond alors à l'algorithme non tronqué de découpe linéaire).

A) Découpe de Gauss Seidel



Pour toutes les valeurs de k testées, la matrice H_r est stabilisée sur 6 chiffres après 6 ou 7 itérations selon les valeurs de k , sur la même matrice :

$\frac{1}{100}$	4	- 2	4	8	- 3	1
	0	2.02093	5.95833	2.91666	1.03125	1.98958
	0	0	3.52647	6.50946	- 4.06272	- 1.83818
	0	0	0	4.99794	- 5.00991	0.96259
	0	0	0	0	7.09471	- 2.98973
	0	0	0	0	0	2.10541

Pour $k = 1$, il a fallu 7 itérations

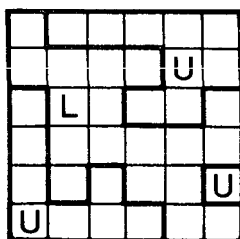
$k = 5, 10, 20, + \infty$ 6 itérations

Dans cet exemple, où L_r est triangulaire inférieure (6,6), on sait que L_r^6 est nul. Autrement dit pour $k \geq 5$

$$F_k(L_r + U_r) = ([I + L_r \dots + L_r^k]U_r + L_r^{k+1})$$

coïncide avec $(I - L_r)^{-1}U_r$, d'où un calcul identique pour $k = 5, 10, 20$ et $+\infty$.

B) Reprenons la découpe (utilisée dans [2])



Les matrices sont stabilisées sur 6 chiffres :

a) pour $k = 1$, en 5 itérations; on obtient, pour matrice limite :

$\frac{1}{100}$	0	- 2.08333	4.16666	8.33332	- 3.12500	1.04166
	0	0	0	0	0.64259	1.95107
	4.11754	0	0	6.32972	- 3.99518	0
	- 2.35491	0	0	0	0	0
	5.51035	0	1.07799	0	0	- 3.25552
	2.26367	- 1.10232	0.97779	3.30805	0	0

b) pour $k = 5, 10, 20, +\infty$, en 5 itérations aussi; on obtient la limite :

	0	- 2.08333	4.16666	8.33332	- 3.12500	1.04166
	0	0	0	0	0.79665	2.04590
$\frac{1}{100}$	4.17814	0	0	6.14326	- 4.07313	0
	- 2.34478	0	0	0	0	0
	5.53198	0	1.07759	0	0	- 3.19503
	- 2.26956	- 1.02178	0.97769	3.06539	0	0

2) Pour une matrice H_0 (24, 24) et pour la découpe de Gauss-Seidel, le calcul a été effectué, en double précision, en prenant la suite des k , stationnaire en la valeur k , pour différentes valeurs de k . Il est intéressant de comparer le nombre d'itérations nécessaires et les temps de calcul correspondants pour une même précision machine (stabilisation sur 10 chiffres décimaux).

k	Nombre d'itérations	Temps de calcul
1	8	1 m 30 s
5	7	3 m 52 s
10	7	7 m 27 s
20	7	14 m 03 s

On constate, sur cet exemple, que pour obtenir une précision donnée, le nombre d'itérations nécessaires est pratiquement indépendant de l'entier de troncature k . Il en résulte que le temps de calcul total pour un même résultat, croît, pratiquement linéairement, avec k . Cette constatation numérique montre que le calcul le plus économique correspond à la troncature maximum : $k = 1$.

Si l'on fait le calcul pour k infini (algorithme de découpe non tronqué) en calculant

$$\left[\sum_{j=0}^{k=\infty} L^j \right] U = (I - L)^{-1} U$$

par inversion directe de $(I - L)$ grâce à la méthode de Gauss, et non plus par une suite de schémas de Horner ($S_0 = L + U$; $S_{k+1} = LS_k + U$), on obtient, sur ce même exemple :

∞	6	1 m 51 s
----------	---	----------

Il ne faut donc pas trop compter sur un grand gain de temps : dans cet exemple, il n'est effectif que pour la troncature maximum ($k = 1$) (et inférieur à 25 %).

III. APPLICATION A DES EQUATIONS LINEAIRES DE POINT FIXE

Pour H_0 donné dans M_n et h_0 donné dans R^n , considérons l'équation de point fixe

$$(E_0) \quad x = H_0 x + h_0$$

Soit s une découpe donnée sur M_n et $\{k_r\}$ une suite d'entiers ≥ 1 .

Définissons alors les deux suites :

$$H_{r+1} = F_{k_r}(H_r); h_{r+1} = [I + L_r + \dots L_r^{k_r}]h_r \quad (r = 0, 1 \dots)$$

pour H_r décomposé en $L_r + U_r$ selon s .

Cet algorithme sera encore appelé algorithme tronqué de découpe linéaire, associé à s et à la suite $\{k_r\}$, et partant de (H_0, h_0) .

Alors, pour tout r , on associe au couple (H_r, h_r) l'équation de point fixe (E_r) suivante dans R^n :

$$(E_r) \quad x = H_r x + h_r$$

On a alors le

Théorème 5

Si $N(H_0) < 1$, où N est une norme compatible avec la découpe s , alors, quelle que soit la suite $\{k_r\}$ ($k_r \geq 1$).

- 1) L'équation de point fixe (E_0) admet une solution unique x_0 .
- 2) Toutes les équations (E_r) sont équivalentes à (E_0) : elles admettent la même solution unique x_0 .
- 3) H_r converge vers une limite $U \in \mathcal{U}$; h_r converge vers une limite h dans R^n .
- 4) Alors l'équation limite :

$$(E) \quad x = Ux + h$$

admet pour unique solution la solution x_0 de (E_0) .

- 5) Si H_0 est ≥ 0 , tous les H_r le sont, et les inégalités (15) (16) et (17) sont vérifiées.

Preuve

a) Puisque $\rho(H_0) \leq N(H_0) < 1$, (E_0) admet une solution unique, que nous notons x_0 .

b) D'après le théorème 4, on sait que $N(H_r) \leq N(H_0) < 1$
d'où

$$\rho(H_r) \leq N(H_r) < 1, \quad (E_r) \text{ admet une solution unique.}$$

Montrons alors par récurrence que x_0 est solution de (E_r) . C'est vrai pour $r = 0$. Supposons que

$$x_0 = H_r x_0 + h_r \quad (19)$$

Alors

$$\sum_{p=0}^{k_r} L_r^p x_0 = \sum_{p=0}^{k_r} L_r^p (L_r + U_r) x_0 + \sum_{p=0}^{k_r} L_r^p h_r$$

soit

$$\sum_{p=0}^{k_r} (L_r^p - L_r^{p+1}) x_0 = (H_{r+1} - L_r^{k_r+1}) x_0 + h_{r+1}$$

c'est-à-dire

$$x_0 = H_{r+1} x_0 + h_{r+1}$$

et (19) est établie pour tout r .

Toutes les équations (E_r) ont donc même solution unique x_0 que (E_0) .

c) d'après le théorème 4, on sait, puisque $N(H_0) < 1$, que H_r converge vers une limite $U \in \mathcal{U}$. Or on a

$$h_r = x_0 - H_r x_0$$

ce qui montre que h_r converge; de plus il converge vers h tel que

$$h = x_0 - U x_0$$

ce qui montre que x_0 est encore solution de l'équation limite (E) ; d'autre part, puisque $\rho(U) \leq N(U) \leq N(H_0) < 1$, l'équation limite admet une solution unique, qui est donc x_0 .

d) Enfin le point 5) reproduit les résultats du théorème 4 dans le cas où $H_0 \geq 0$.

Le théorème 5 est démontré.

REMARQUE : Dans le cas de la découpe triviale de M_n en $\mathcal{L} \oplus \mathcal{U}$ où $\mathcal{U} = \{0\}$ et $\mathcal{L} = M_n$, il vient

$$(E_r) \quad x = H_r x + h_r \quad (r = 0, 1, \dots)$$

avec

$$H_{r+1} = H_r^{k_r+1} \quad \text{et} \quad h_{r+1} = [I + H_r + \dots + H_r^{k_r}] h_r$$

Toute norme N sous-multiplicative sur M_n est alors compatible avec la découpe envisagée. La condition $N(H_0) < 1$ entraîne que $\rho(H_0) < 1$ (et d'ailleurs, $\rho(H_0)$ est l'inf des $N(H_0)$, pour N sous multiplicative). Il est alors simple de démontrer directement dans ce cas les résultats du théorème 5 : puisque $\rho(H_0) < 1$, H_r tend vers 0, et h_r tend vers $(I - H_0)^{-1} h_0$. L'équation limite s'écrit

$$x = (I - H_0)^{-1} h_0$$

En particulier, pour la troncature maximum (k_r constant et égal à 1) il vient

$$H_r = H_0^{2^r} \quad \text{et} \quad h_r = [I + H_0 + \dots H_0^{2^r-1}]h_0$$

EXEMPLE. On a traité ainsi, pour la découpe de Gauss-Seidel, un système linéaire (24, 24)

$$x = H_0 x + h_0$$

où la matrice H_0 est celle utilisée dans l'exemple (2) cité plus haut. En double précision, et avec la troncature maximum ($k_r = 1$) on obtient la stabilisation de H_r et h_r sur 10 chiffres pour $r = 8$. Le traitement du second membre fait passer le temps de calcul de 1 mn 30 s [cf. exemple (2)] à 1 mn 38 s. H_8 est alors (à la précision indiquée) triangulaire supérieure.

A noter que la triangularisation du système linéaire par la méthode de Gauss ne prend que 6 s! (1).

IV. CONCLUSION

Une première conclusion est négative : l'expérimentation numérique prouve que l'algorithme tronqué de découpe linéaire n'est pas compétitif sous le rapport du temps de calcul nécessaire, avec les méthodes classiques de résolution de systèmes linéaires (le contraire eut été étonnant).

Par contre, le résultat principal de convergence (théorème 4), ainsi que le théorème de Stein-Rosenberg « tronqué » (théorème 2) sont des outils théoriques que nous pensons intéressants : ils constituent le matériel de base pour l'extension de l'algorithme présenté ici à des équations de point fixe dans R^n , non linéaires mais contractantes en norme vectorielle, particulièrement pour le cas de découpes implicites [1].

REFERENCES

- [1] M. CHAMBAT, *Algorithmes non linéaires de découpe*. Thèse de troisième cycle Université de Lyon 1 (Juin 1972).
- [2] F. ROBERT et J. F. MAITRE, *Normes et algorithmes associés à une découpe de matrices*, Num. Math. 1972.
- [3] M. RASCLE, *Irréductibilité et primitivité de certains opérateurs non linéaires* Thèse de troisième cycle Université de Lyon 1 (Juin 1972).
- [4] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, 1962.

(1) Ceci s'explique : pour une matrice (n, n), la triangularisation par Gauss nécessite un peu plus de $n^3/3$ multiplications ou divisions.

D'autre part, avec la troncature maximum, chaque pas de l'algorithme tronqué coûte une multiplication matricielle, soit n^3 multiplications.

Si r pas sont nécessaires, le rapport des temps est alors quelque peu inférieur à $3r$, ce que confirme notre exemple.