REVUE FRANÇAISE D'AUTOMATIQUE, INFORMATIQUE, RECHERCHE OPÉRATIONNELLE. MATHÉMATIQUE

N. XUAN LUONG

Sur la méthode de sur-relaxation, dans le cas des problèmes avec contrainte : un résultat de convergence asymptotique

Revue française d'automatique, informatique, recherche opérationnelle. Mathématique, tome 7, nº 2 (1973), p. 107-113.

<http://www.numdam.org/item?id=M2AN_1973__7_2_107_0>

R.A.I.R.O.
(7e année, R-2, 1973, p. 107-113)

SUR LA METHODE DE SUR-RELAXATION, DANS LE CAS DES PROBLEMES AVEC CONTRAINTE : UN RESULTAT DE CONVERGENCE ASYMPTOTIQUE

par N. Xuan LUONG [1]

Résumé. — *On considère des méthodes de type relaxation appliquées à des systèmes non linéaires de R^n provenant de la perturbation d'un opérateur diagonal maximal monotone, par un opérateur linéaire ; méthodes dont on étudie le comportement asymptotique.*

I. POSITION DU PROBLEME

Soit à résoudre le système non linéaire :

$$(1) \qquad \left| \begin{array}{l} \text{chercher } x \text{ tel que} \\ f \in Ax + M(x) \\ x \in R^n \cap D(M) \end{array} \right.$$

où f est un élément fixé de R^n, A une matrice $n \times n$ de diagonale strictement positive, M un opérateur diagonal, $M(x) = \{ M_1(x_1), ..., M_n(x_n) \}$, dont chaque composante M_i est un opérateur maximal monotone de R dans R et $D(M)$ est le domaine de définition de M.

II. ALGORITHMES DE RESOLUTION

Miellou dans [3], [4], [5] propose des algorithmes de résolution pour le problème (1).

(1) Laboratoire de Calcul, Université de Besançon.

Soit $I = \{ 1, 2, \ldots, N \}$, avec $N \leqslant n$. Soit un partitionnement de A en bloc matrice, de blocs A_{ij}, $i, j \in I$, tel que A_{ii}, $i \in I$, soient des matrices carrées de type $\alpha_i \times \alpha_i$.

On note par $x = (X_i)_{i \in I}$, $f = (F_i)_{i \in I}$ le partitionnement correspondant des vecteurs x et f.

Pour tout $i \in I$ on associe $S_i = \{ 1, 2, \ldots, i - 1, i + 1, \ldots, N \}$, On considère deux applications τ et $\bar{\tau}$ de I dans $\mathcal{P}(I)$, l'ensemble des parties de I, tels que $\tau(i) \subset S_i$ et $\bar{\tau}(i)$ soit le complémentaire de $\tau(i)$ dans S_i. Une telle application τ est appelée application de découpe.

Soit $\omega \in]0,2]$. Soit x^0 donné de R^n. On considère les itérés successifs $x^{p+1} = T_{\omega,\tau} x^p$, $p = 0, 1, 2, \ldots$, $T_{\omega,\tau}$ définie par :

$$(2) \quad \begin{cases} \omega F_i - \omega \sum\limits_{j \in \tau(i)} A_{ij} X_j^{p+1} - \omega \sum\limits_{\in \bar{\tau}(j)} A_{ij} X_j^p \in A_{ii}(X_i^{p+1} + (\omega - 1) X_i^p) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \omega M_i(X_i^{p+1}) \\ \text{pour} \quad i = 1, 2, \ldots, N. \end{cases}$$

Remarques. — Si l'on choisi $\tau(i) = \{ 1, \ldots, i - 1 \}$ $\forall i = 2, \ldots, N$, alors $\omega = 1$ correspond à une méthode de résolution de type Gauss-Seidel par bloc, $\omega \neq 1$ correspond à une méthode de type sur(sous)-relaxation. Dans [5] cité précédemment, on trouve d'autres exemples de fonctions τ qui correspondent à des méthodes adaptées au calcul en parallèle. Dans la partie IV nous citerons des théorèmes de convergence des algorithmes, dans des situations particulières.

III. COMPORTEMENT ASYMPTOTIQUE DE LA METHODE DE RELAXATION

Nous considérons le problème (1) avec les hypothèses :

(3) A est une matrice par bloc, $A_{i,j}$, $i, j \in I = [1, 2, \ldots, N]$. On suppose que

$$\forall i \in I, (A_{ii} u_i, u_i) > 0 \quad \forall u_i \in R^{\alpha_i}, u_i \neq 0.$$

(4) La solution z du problème (1) et le graphe de l'opérateur M sont tels que : $\forall i \in [1, 2, \ldots n]$, il existe un voisinage V_i de $(z_i, f_i - [Az]_i)$ tels que si G_i est le graphe de M_i dans $R \times R$, $G_i \cap V_i$ soit prolongeable par une droite D_i de pente nulle ou infinie.

(5) On note par H l'ensemble des indices tels que la pente de D_i, $i \in [1, 2, \ldots, n]$, soit nulle et $\bar{H}$ le complémentaire de H.

On considère la matrice B, sous-matrice de A obtenue en supprimant les lignes et les colonnes dont l'indice appartient à $\bar{H}$. La structure de bloc-matrice considérée de A induit sur B une structure de bloc-matrice, chaque bloc B_{ij} s'obtient de A_{ij} de la même manière que B de A. On peut écrire : $B = (B_{ij})$, i, $j \in J = [1, 2, ..., M]$ avec $M \leqslant N$.

En utilisant l'hypothèse (3), on vérifie que $\forall i \in J$, les blocs B_{ii} sont inversibles.

On note par D la matrice diagonale par blocs, de blocs B_{ii}, $i \in J$. τ étant une application de découpe de ligne, $\forall j \in J$, on note par $\tau'(j) = \tau(j) \cap J$ et $\bar{\tau}'(j) = \bar{\tau}(j) \cap J$.

Soient L et U, les matrices par blocs définies par :

$$L_{ij} = \begin{cases} -B_{ii}^{-1}B_{ij} \ \forall j \in \tau'(i), & i \in J \\ \quad 0 \qquad \forall j \notin \tau'(i), & i \in J \end{cases}$$

$$U_{ij} = \begin{cases} -U_{ii}^{-1}U_{ij} \ \forall j \in \bar{\tau}'(i) & i \in J \\ \quad 0 \qquad \forall j \notin \bar{\tau}'(i) & i \in J \end{cases}$$

Soit $\mathfrak{L}_{\omega,\tau} = (I - \omega L)^{-1}((1 - \omega)I + \omega U)$
$\omega \in]0,2]$, soit $\rho(\mathfrak{L}_{\omega,\tau})$ le rayon spectral de $\mathfrak{L}_{\omega,\tau}$.

Proposition 1

Sous les hypothèses (3), (4), la suite $\{x^p\}$ étant définie par $x^{p+1} = T_{\omega,\tau}x^p, p = 0, 1, 2, ...,$ et en supposant de plus que $\{x^p\}$ converge vers z, solution du problème (1), alors on a :

$$\limsup_{\substack{p\to\infty \\ x^0 \in R^n}} \|x^p - z\|^{1/p} \leqslant \rho(\mathfrak{L}_{\omega,\tau})$$

Démonstration : On peut réécrire (2) sous la forme :

(6) $$\omega\chi_i(X_i^{p+1}) + R_i^p = 0, \text{ pour } i = 1, 2, ..., N,$$

où $$\chi_i(X_i^{p+1}) \in M_i(X_i^{p+1})$$
et

(7) $$R_i^p = \omega \sum_{j \in \tau(i)} A_{ij}X_j^{p+1} + A_{ii}(X_i^{p+1} + (\omega - 1)X_i^p) + \omega \sum_{j \in \bar{\tau}(i)} A_{ij}X_j^p - \omega F_i$$

Par hypothèse x^p, x^{p+1} convergent vers z solution de (1). R_i^p est une fonction continue de x^p, x^{p+1}, les relations (7), pour $i = 1, 2, ..., N$, pour $p \to \infty$, déterminent un vecteur $\mathfrak{R}$ de R^n, de composantes r_i. La relation (6) montre que $\chi_j(x_j^{p+1})$ converge vers $-\frac{1}{\omega} r_j$, $\forall j = 1, 2, \dots, n$. Considérons les voisinages V_i de l'hypothèse (4). Lorsque $p \to \infty$, $(x_i^{p+1}, \chi_i(x_i^{p+1}))$ converge vers $(z_i, \chi_i(z_i))$. On en déduit qu'il existe un entier p_0 tel que $\forall p \geqslant p_0$ on ait $(x_i^{p+1}, \chi_i(x_i^{p+1})) \in V_i$, $i = 1, 2, \dots, n$. Avec l'hypothèse (5) on a :

$$\forall p \geqslant p_0, \forall i \in \bar{H}, x_i^{p+1} = z_i \tag{8}$$

$$\forall p \geqslant p_0, \forall i \in H, \chi_i(x_i^{p+1}) = M_i(x_i^{p+1}) = C_i, C_i \tag{9}$$

est une constante, pour $x_i^{p+1} \in V_i$.

Notons par G la matrice $n \times n$ de coefficients g_{ij} avec :

$$g_{ij} = \begin{cases} a_{ij} & \text{pour } i \in H \text{ et } j \in \bar{H} \\ 0 & \text{autrement} \end{cases}$$

Notons par $\bar{X}_i$, le sous-vecteur de dimension l_i de X_i obtenu en enlevant les éléments de X_i dont l'indice appartient à $\bar{H}$. On déduit de (2), (8), (9),

$$\forall p \geqslant p_0 \text{ et } i = 1, 2, \dots, M$$

$$\sum_{j \in \bar{T}'(i)} \omega B_{ij} \bar{X}_j^{p+1} + \sum_{j \in \bar{T}'(i)} B_{ij} \bar{X}_j^p + B_{ii}(\bar{X}_i^{p+1} + (\omega - 1)\bar{X}_i^p = \omega Q_i \tag{10}$$

où Q_i est un sous-vecteur de même dimension que $\bar{X}_i$, en notant Q_{ik}, $k = 1, 2, \dots, l_i$, les composantes de Q_i,

$$Q_{ik} = f_{ik} - c_{ik} - \sum_{j=1}^{n} g_{i_k j} z_j$$

Soit Q le bloc vecteur de composantes Q_i, $i = 1, 2, \dots, M$.

Les relations (10) peuvent se réécrire :

$$\begin{cases} \forall p \geqslant p_0 \\ \bar{x}^{p+1} = (I - \omega L)^{-1} ((1 - \omega) I + \omega U) \bar{x}^p + \omega (I - \omega L)^{-1} Q \end{cases}$$

d'où le résultat.

IV. SITUATIONS PARTICULIERES SUR- (SOUS) RELAXATION PAR POINTS

1. Cas des H-matrices

Une matrice A, dont les termes diagonaux sont strictement positifs est une H-matrice si la matrice $\tilde{A}$, avec $\tilde{a}_{ij} = a_{ij}$, pour $i = j$ et $\tilde{a}_{ij} = - |a_{ij}|$ pour $i \neq j$ est inversible et $\tilde{A}^{-1}$ a tous ses termes non négatifs.

Considérons le problème 1 avec l'hypothèse : (H_1) A est une H-matrice.

L'algorithme provenant de (2), dans le cas de la sur-(sous)relaxation s'écrit :

$$\left| \begin{array}{l} x^0 \in D(M) \\ x^{p+1} = T_\omega x^p,\ T_\omega \text{ définie par} \\ \omega f_i + (1-\omega) a_{ii} x_i^p - \omega \sum_{j=1}^{i-1} a_{ij} x_j^{p+1} - \omega \sum_{j=i+1}^{n} a_{ij} x_j^p \in a_{ii} x_i^{p+1} + \omega M(x_i^{p+1}) \\ \text{pour } i = 1, 2, \dots, n. \end{array} \right.$$

On note par J la matrice de Jacobi associée à la M-matrice $\tilde{A}$, et $J = L + U$, où L est une matrice strictement triangulaire inférieure et U une matrice strictement triangulaire supérieure. Nous énonçons le résultat suivant, qui est un corollaire des résultats de Miellou [3], [5].

Proposition 2

Le problème 1, avec l'hypothèse H_1 a une solution unique x. Pour tout ω tel que $0 < \omega < \dfrac{2}{1 + \rho(J)}$, le processus itératif associé à T_ω converge vers x; on a la relation :

(11) $\limsup \|x - x^p\|^{1/p} \leqslant \rho(t_\omega)$ où $\rho(t_\omega)$ est le rayon spectral inférieur à 1 de la matrice : $t_\omega = (I - \omega L)^{-1}((\omega - 1)I + \omega U)$.

2. Cas des matrices symétriques, définies positives

Soit K_i un convexe fermé de R. On note par $\partial\Psi_{K_i}$ l'opérateur sous-différentiel de la fonction indicatrice de K_i, on sait que $\partial\Psi_{K_i}$ est maximal monotone.

Soit $K = \prod_{i=1}^{n} K_i$.

On note par $\partial\Psi_K = (\partial\Psi_{K_1}, \dots, \partial\Psi_{K_n})$.

On considère le problème 1 avec l'hypothèse :
(H_2) A est une matrice symétrique, définie positive et $M = \partial\Psi_K$.

Si $K_i = [a_i, b_i]$, a_i, b_i peuvent prendre des valeurs infinies, et si $y \in R$ on note par :

$$\operatorname{Proj} y|_{K_i} = \begin{cases} y & \text{si } y \in K_i \\ a_i & \text{si } y < z \quad \forall z \in K_i \\ b_i & \text{si } y > z \quad \forall z \in K_i \end{cases}$$

La sur(sous)-relaxation, pour le problème 1 avec l'hypothèse H_2 est définie par l'algorithme :

$$\left|\begin{array}{l} x^0 \text{ donné de } K\,;\ x^{p+1} = T_\omega x^p,\ T_\omega \text{ définie par} \\ a_{ii}x_i^{p+1/2} = \omega f_i + (1-\omega)a_{ii}x_i^p - \omega \sum_{j=1}^{i-1} a_{ij}x_j^{p+1} - \omega \sum_{j=i+1}^{n} a_{ij}x_j^p \\ x_i^{p+1} = \operatorname{proj} x_i^{p+1/2} \mid K_i \\ \text{pour } i = 1, 2, \ldots, n \end{array}\right.$$

On a le résultat de convergence suivant, qui est un corollaire de résultats de Glowinski [2].

Proposition 3

Avec l'hypothèse H_2, pour tout ω, $0 < \omega < 2$, x^p converge vers x, solution unique du problème 1.

3. Vitesse de convergence

Dans le cas linéaire, rappelons que la vitesse de convergence asymptotique pour la sur-(sous)relaxation est :

$$(12) \qquad R_\infty(\mathfrak{L}_\omega(A)) = -\log\left[\rho(\mathfrak{L}_\omega(A))\right] = -\log\left[\limsup_{\substack{k\to\infty \\ x^0\in R^n}} \|x^k - x\|^{1/k}\right]$$

$\rho(\mathfrak{L}_\omega(A))$ étant le rayon spectral de la matrice de relaxation associée à A, cf. [7]. Dans le cas général d'un problème non linéaire, la vitesse de convergence asymptotique d'un processus itératif T peut être définie par, cf. [6] :

$$(13) \qquad R(T) = -\log\left[\limsup_{\substack{k\to\infty \\ x^0\in R^n}} \|x^k - x\|^{1/k}\right]$$

avec les notations de III, on déduit de la proposition 1, de (12) et (13).

Corollaire 1

Si la solution du problème 1 avec l'hypothèse H_j, $j = 1$ ou 2, est telle que l'hypothèse (4) soit vérifiée, alors :

$$R(T_\omega) = \begin{cases} \infty & \text{si } H = \varnothing \\ R_\infty[\mathfrak{L}_\omega(B)] & \text{si } H \neq \varnothing \end{cases}$$

REMARQUES

a) La majoration de la vitesse de convergence asymptotique en (11) ne nécessite pas l'hypothèse (4). Mais l'utilisation de la théorie de Young-Frankel ne permet pas de rendre compte dans ce cas de l'amélioration apportée par l'introduction du paramètre de relaxation. La proposition 1 et le corollaire 1 permettent au contraire d'apporter une justification théorique à l'introduction de paramètre de sur-relaxation.

b) On trouve un résultat de convergence de type énoncé dans le corollaire 1, dans [1], dans le cadre de la sur-(sous)relaxation ponctuelle appliquée à un problème d'optimisation avec des contraintes particulières $K_i = R^+$.

c) Les résultats de cette partie sont encore valables pour la sur(sous)-relaxation par blocs. En effet les propositions 1 et 2 sont énoncées à partir de résultats plus généraux de [3], [2].

BIBLIOGRAPHIE

[1] C. W. CRYER, *The solution of a quadratic programming problem using systematic over-relaxation*, SIAM J. Control, vol. 9, n° 3, 1971.

[2] R. GLOWINSKI, *La méthode de relaxation*, Quaderni dei rendiconti di Matematica, 14, Rome, 1971.

[3] J. C. MIELLOU, *Opérateurs paramonotones*, Thèse, Grenoble, 1970.

[4] J. C. MIELLOU, *Méthodes de Jacobi, Gauss-Seidel, Sur(sous)-relaxation*, C.R.A.S., t. 273, 1971.

[5] J. C. MIELLOU, Colloque d'Analyse Numérique d'Épinal, 1972.

[6] J. M. ORTEGA, W. C. RHEINBOLDT, *Iterative solution of nonlinear equations in several variables*, Academic Press, 1970.

[7] R. VARGA, *Matrix Iterative Analysis*, Prentice Hall, 1962.