Revue française d'informatique et de recherche opérationnelle

Jean Cea

Les méthodes de « descente » dans la théorie de l'optimisation

Revue française d'informatique et de recherche opérationnelle, tome 2, nº R3 (1968), p. 79-101
<http://www.numdam.org/item?id=M2AN_1968__2_3_79_0>

R.I.R.O.
(2e année, N° 13, 1968, p. 79-102)

LES METHODES DE "DESCENTE" DANS LA THEORIE DE L'OPTIMISATION

par Jean CEA [1]

Résumé. — *Il est question d'approcher la solution d'un problème de minimisation d'une forme quadratique à l'aide d'une méthode itérative. Dans chaque itération, on résoud un problème de minimisation à une variable. Dans les nos 1 et 2 on traite le cas des problèmes sans contrainte et avec contraintes ; dans le no 3, il y a une application à un problème de contrôle dans les équations aux dérivées partielles.*

INTRODUCTION

Pour trouver (ou approcher) un point en lequel une fonction de plusieurs variables est minimum, il est classique de remplacer le problème initial par une suite (infinie, en général) de problèmes de minimisation où la fonction économique ne dépend plus que d'une variable : c'est le principe des méthodes de descente. Dans cet article, nous nous proposons d'étudier systématiquement, dans un cadre simple, la méthode de descente, dans le cas sans contraintes d'abord, avec contraintes ensuite. Nous retrouvons ainsi plusieurs méthodes de la théorie de l'optimisation ainsi que les méthodes classiques dans la résolution des grands systèmes linéaires. Pour finir, nous appliquons la méthode de descente à la théorie du contrôle (dans les équations aux dérivées partielles)

1. LE CAS SANS CONTRAINTE

1.1. Position du problème

Tous les éléments introduits sont réels.

V est un espace de Hilbert, de produit scalaire (...) et de norme $\| \; \|$.

A est un opérateur *linéaire* de V sur V.

f est un élément donné dans V.

(1) Faculté des Sciences de Rennes.

Les hypothèses :

*H*1, l'opérateur A est *hermitien.*

*H*2, l'opérateur A est *coercif :* il existe $\alpha > 0$ tel que

$$(Av, v) \geqslant \alpha \|v\|^2 \qquad \forall v \in V$$

*H*3, l'opérateur A est continu; il existe donc $M > 0$ tel que

$$(Au, v) \leqslant M \cdot \|u\| \cdot \|v\| \qquad \forall u, v \in V.$$

Nous savons que les hypothèses précédentes entraînent ceci : l'opérateur A est inversible; nous désignons par u l'élément unique qui vérifie

$$Au = f \tag{1.1}$$

Nous nous proposons de déterminer u comme solution d'un problème de minimisation ; pour cela posons

$$J(v) = (A(v - u), v - u)$$

grâce à l'hypothèse de coercivité, nous avons

$$J(u) = \underset{v \in V}{\operatorname{Min}} J(v)$$

Remarquons qu'il est équivalent de minimiser $(Av, v) - 2(f, v)$.

1.2. Le principe des méthodes de descente

A partir d'un élément u_0 donné, on construit une suite $u_1, \ldots, u_n, \ldots$ destinée à converger vers la solution u du problème. Le passage de u_n à u_{n+1} se fait de la façon suivante : on choisit une direction (de descente) p_{n+1}, u_{n+1} est alors le point de la droite $u_n + \rho \cdot p_{n+1}$, $\rho \in \mathbf{R}$, pour lequel $J(v)$ est minimum. De façon plus précise, posons

$$u_{n+1}(\rho) = u_n + \rho p_{n+1}$$

On a

$$J(u_{n+1}(\rho)) = J(u_n) + 2\rho(A(u_n - u), p_{n+1}) + \rho^2 (Ap_{n+1}, p_{n+1})$$

ou encore, en posant

$$g_n = -(A(u_n - u)) = -(Au_n - f) \tag{1.3}$$

$$J(u_{n+1}(\rho)) = J(u_n) - 2\rho(g_n, p_{n+1}) + \rho^2(Ap_{n+1}, p_{n+1}) \tag{1.4}$$

La fonction $\rho \to J(u_{n+1}(\rho))$ atteint son minimum pour

$$\rho = \rho_{n+1} = \frac{(g_n, p_{n+1})}{(Ap_{n+1}, p_{n+1})} \tag{1.5}$$

et on a

$$(1.6) \qquad \begin{cases} J(u_{n+1}) = J(u_n) - \dfrac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1})} \\ u_{n+1} = u_n + \rho_{n+1} p_{n+1} \end{cases}$$

(Remarquons que grâce à l'hypothèse $H2$, $p_{n+1} \neq 0 \Rightarrow (Ap_{n+1}, p_{n+1}) > 0$.)

Si $u_n \neq u$, alors $J(u_n) \neq 0$ et nous pouvons écrire

$$J(u_{n+1}) = J(u_n) - J(u_n) \cdot \frac{1}{J(u_n)} \cdot \frac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1})}$$

ou encore, puisque $J(u_n) = (A(u_n - u), u_n - u) = (A^{-1}g_n, g_n)$, on a

$$(1.7) \qquad \begin{cases} J(u_{n+1}) = J(u_n)[1 - E_n] \\ E_n = \dfrac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1}) \cdot (A^{-1}g_n, g_n)} \end{cases}$$

Les relations (1.7) sont fondamentales pour la suite.

Si nous choisissons p_{n+1} de façon que

$$(1.8) \qquad E_n \geqslant \beta$$

où β est une constante > 0 indépendante de n, alors

$$J(u_{n+1}) \leqslant J(u_n)(1 - \beta)$$

et par récurrence

$$(1.9) \qquad J(u_n) \leqslant (1 - \beta)^n J(u_0)$$

Si on utilise l'hypothèse $H2$ il vient :

$$\alpha \|u_n - u\|^2 \leqslant (1 - \beta)^n J(u_0)$$

d'où

$$(1.10) \qquad \|u_n - u\| \leqslant (1 - \beta)^{\frac{n}{2}} \cdot \sqrt{\frac{J(u_0)}{\alpha}}$$

Théorème 1.1. — *Pour tout choix « raisonnable » de p_{n+1} [i.e. tel que l'inégalité* (1.8) *ait lieu], la méthode de descente est convergente ; de façon plus précise on a les inégalités* (1.9) *et* (1.10).

Nous allons maintenant donner une condition suffisante pour que (1.8) ait lieu ; grâce aux hypothèses $H2$ et $H3$, nous avons

$$(Av, v) \leqslant M \|v\|^2 \qquad \forall v \in V$$

$$(A^{-1}v, v) \leqslant \frac{1}{\alpha} \|v\|^2 \qquad \forall v \in V$$

et par suite

$$\frac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1}) \cdot (A^{-1}g_n, g_n)} \geqslant \frac{\alpha(g_n, p_{n+1})^2}{M \|p_{n+1}\|^2 \cdot \|g_n\|^2}$$

d'où

(1.11) $$E_n \geqslant \frac{\alpha}{M} \left(\frac{g_n}{\|g_n\|}, \frac{p_{n+1}}{\|p_{n+1}\|} \right)^2$$

Si

(1.12) $$\left| \left(\frac{g_n}{\|g_n\|}, \frac{p_{n+1}}{\|p_{n+1}\|} \right) \right| \geqslant \gamma > 0 \qquad \forall n.$$

Alors

$$E_n \geqslant \frac{\alpha}{M} \cdot \gamma^2$$

et on peut prendre dans la théorie précédente

$$\beta = \frac{\alpha}{M} \cdot \gamma^2$$

d'où le

Théorème 1.2. — *Sous l'hypothèse : il existe* $\gamma > 0$ *tel que*

$$\left| \left(\frac{g_n}{\|g_n\|}, \frac{p_{n+1}}{\|p_{n+1}\|} \right) \right| \geqslant \gamma > 0 \qquad \forall n.$$

la méthode de descente est convergente et on a les inégalités (1.9) *et* (1.10) *avec* $\beta = \frac{\alpha}{M} \cdot \gamma^2$.

Nous allons étudier maintenant quelques cas particuliers. Auparavant, il est bon de remarquer ceci : le théorème précédent montre que toute direction non « presque » orthogonale au gradient conduit à une méthode convergente. Le choix de la direction p_{n+1} se fera en fonction de l'information connue au point u_n : celle-ci est la suivante :

i) le gradient g_n,

ii) les directions des axes (dans le cas de la dimension finie),

iii) s'il n'y a pas de problème de mémorisation, on peut connaître le point u_{n-1} et donc la direction $u_{n-1} - u_n$; il sera peut-être possible de connaître d'autres points.

$\Rightarrow$ | On choisira p_{n+1} comme une combinaison des directions connues après la $n^{\text{ième}}$ itération.

1.3. La méthode du gradient

On choisit

(1.13) $$p_{n+1} = g_n$$

d'où en introduisant la notation E_n^G

$$E_n = E_n^G = \frac{\|g_n\|^4}{(Ag_n, g_n)(A^{-1}g_n, g_n)}$$

ce qui, avec les hypothèses $H1$, $H2$ et $H3$ et l'inégalité de Kantorovitch entraîne

$$E_n^G \geqslant \frac{4\alpha M}{(M+\alpha)^2}. \tag{1.14}$$

On peut donc appliquer le théorème 1.1 avec $\beta = \dfrac{4\alpha M}{(M+\alpha)^2}$.

1.4. La méthode du gradient conjugué

On choisit

$$p_{n+1} = g_n + \mu_{n+1} \cdot p_n \tag{1.15}$$

où μ_{n+1} est déterminé au mieux, c'est-à-dire de façon que E_n soit maximum : on a en introduisant la notation E_n^{GC}

$$E_n = E_n^{GC} = \frac{(g_n, g_n + \mu_{n+1} \cdot p_n)^2}{(A(g_n + \mu_{n+1}p_n), g_n + \mu_{n+1}p_n) \cdot (A^{-1}g_n, g_n)}$$

Géométriquement parlant, nous pouvons dire ceci : l'ellipsoïde $J(v) = J(u_n)$ est tangent en u_n à la droite $u_{n-1} + \rho p_n$ et par conséquent

$$(g_n, p_n) = 0$$

d'où

$$E_n^{GC} = \frac{\|g_n\|^4}{(A(g_n + \mu_{n+1}p_n), g_n + \mu_{n+1}p_n), (A^{-1}g_n, g_n)}$$

Il est clair maintenant que le meilleur choix possible de μ_{n+1} est la valeur de μ pour laquelle le trinôme $\mu \to (A(g_n + \mu p_n), g_n + \mu p_n)$ atteint son minimum

$$\mu_{n+1} = -\frac{(Ag_n, p_n)}{(Ap_n, p_n)} \tag{1.16}$$

De (1.16) il vient immédiatement :

$$(Ap_{n+1}, p_n) = (A(g_n + \mu_{n+1}p_n), p_n) = 0 \tag{1.17}$$

ce qui montre que les *directions* p_n et p_{n+1} sont conjuguées pour les ellipsoïdes $J(v) = c$.

On a donc

$$E_n^{GC} = \frac{\|g_n\|^4}{\left\{ (Ag_n, g_n) - \dfrac{(Ag_n, p_n)^2}{(Ap_n, p_n)} \right\} \cdot (A^{-1}g_n, g_n)} \tag{1.18}$$

et

$$E_n^{GC} \geqslant E_n^G \geqslant \frac{4\alpha M}{(M+\alpha)^2}$$

Par suite la méthode du gradient conjugué est convergente.

1.5. La méthode du gradient axial

Nous sommes maintenant dans le cas de la dimension finie; on peut identifier V avec $\mathbf{R}^q$, une base de V est désignée par les vecteurs $e_1, \dots, e_q$. Nous allons introduire les points

$$u_1, \dots, u_n, u_{n+\frac{1}{q}}, \dots, u_{n+\frac{i}{q}}, \dots, u_{n+1} = u_{n+\frac{q}{q}}, \dots$$

et les éléments $p_{n+\frac{i}{q}}, g_{n+\frac{i}{q}}, \dots$

Nous choisissons, lors du passage de $u_{n+\frac{i}{q}}$ à $u_{n+\frac{i+1}{q}}$,

$$p_{n+\frac{i+1}{q}} = g_{n+\frac{i}{q}} + \mu_{n+\frac{i+1}{q}}, e_{i+1} \tag{1.19}$$

où $\mu_{n+\frac{i+1}{q}}$ est choisi « au mieux » comme nous allons le voir :

Posons

$$p_{n+\frac{i+1}{q}}(\mu) = g_{n+\frac{i}{q}} + \mu e_{i+1}$$

$$E^{GA}_{n+\frac{i}{q}}(\mu) = \frac{\left(g_{n+\frac{i}{q}}, p_{n+\frac{i+1}{q}}(\mu)\right)^2}{\left(Ap_{n+\frac{i+1}{q}}(\mu), p_{n+\frac{i+1}{q}}(\mu)\right) \cdot \left(A^{-1}g_{n+\frac{i}{q}}, g_{n+\frac{i}{q}}\right)}$$

Nous voyons que

$$E^{GA}_{n+\frac{i}{q}}(0) = E^G_{n+\frac{i}{q}}$$

Par suite, pour tout choix de μ tel que $E^{GA}_{n+\frac{i}{q}}(\mu) \geqslant E^{GA}_{n+\frac{i}{q}}(0)$ *la méthode sera convergente.*

Une étude du choix pratique de $\mu_{n+\frac{i+1}{q}}$ et de l'organisation des calculs reste à faire.

1.6. Méthode de plus grande descente axiale

Cas de la dimension finie : $V = \mathbf{R}^q$.

On introduira les points $u_1, \dots, u_n, u_{n+1}, \dots$ et les éléments $p_n, \dots, g_n, \dots$ Le choix de p_{n+1} se fera de la façon suivante :

Soit i_n un indice entre 1 et q pour lequel on a

$$\frac{(g_n, e_{i_n})^2}{(Ae_{i_n}, e_{i_n})} \geqslant \frac{(g_n, e_i)^2}{(Ae_i, e_i)} \forall i, i = 1, \ldots, q$$

Nous posons

$$p_{n+1} = e_{i_n}$$

On peut vérifier qu'il existe une constante positive β telle que

$$E_n = \frac{(g_n, e_{i_n})^2}{(Ae_{i_n}, e_{i_n})} \geqslant \beta \qquad \forall n$$

et par suite *la méthode est convergente.*

La direction e_{i_n} est celle de plus « grande descente ».

REMARQUE 1.1.

Il est possible de construire plusieurs méthodes, en faisant intervenir des points « partiels » $u_{n+\frac{i}{q}}$ et en prenant pour $p_{n+\frac{i+1}{q}}$ le vecteur e_{i+1}, ou une combinaison linéaire d'un certain nombre de vecteurs (e_j) de la base de V.

1.7. La méthode de l'opérateur auxiliaire

Soit un opérateur B donné dans $\mathcal{L}(V, V)$ et vérifiant :

$$\begin{cases} B \text{ est hermitien} & \\ |(Bu, v)| \leqslant N \cdot \|u\| \cdot \|v\| & \forall u, v \in V \\ (Bv, v) \geqslant \beta \|v\|^2 & \forall v \in V ; \beta > 0 \end{cases}$$

Posons :

$$\begin{cases} p_{n+1} = B g_n \\ u_{n+1} = u_n + \rho_{n+1} \cdot p_{n+1} \end{cases}$$

nous avons

$$\left(\frac{g_n}{\|g_n\|}, \frac{Bg_n}{\|Bg_n\|}\right) \geqslant \frac{\beta \|g_n\|^2}{\|g_n\| \cdot \|Bg_n\|} \geqslant \frac{\beta}{N}$$

et donc d'après le théorème 1.2 la méthode est convergente.

Notons ceci : B^{-1} existe et a les mêmes propriétés que B, donc la méthode :

$$\begin{cases} p_{n+1} = B^{-1} g_n \\ u_{n+1} = u_n + \rho_{n+1} p_{n+1} \end{cases}$$

ou encore

$$Bu_{n+1} = Bu_n - \rho_{n+1}(Au_n - f)$$

est encore convergente.

Cette méthode est utilisée par Fletcher et Powell [8], la matrice B variant d'un itéré à l'autre.

Dans le cas non linéaire, une méthode semblable est utilisée par plusieurs auteurs : Browder [5], Brezis et Sibony [4].

1.8. La méthode des directions alternées

C'est en fait une variante de la méthode précédente : on suppose que A s'écrit :

$$A = A_1 + A_2$$

A_1 et A_2 ayant des propriétés analogues à celles de A, ou à celles de l'opérateur B introduit dans 1.7.

On pose :

$$\begin{cases} p_{n+\frac{1}{2}} = A_1^{-1} g_n \\ u_{n+\frac{1}{2}} = u_n + \rho_{n+\frac{1}{2}} \cdot p_{n+\frac{1}{2}} \end{cases}$$

puis

$$\begin{cases} p_{n+1} = A_2^{-1} g_{n+\frac{1}{2}} \\ u_{n+1} = u_{n+\frac{1}{2}} + \rho_{n+1} \cdot p_{n+1} \end{cases}$$

(généralisation immédiate au cas ou $A = A_1 + \ldots + A_m$).

Il s'agit de la méthode du 1.7 avec $B_n = A_1^{-1}$ et $B_{n+\frac{1}{2}} = A_2^{-1}$; la méthode est évidemment convergente.

1.9. Une variante des méthodes de descente

Dans les méthodes précédentes, les formules essentielles étaient les suivantes :

$$u_{n+1} = u_n + \rho_{n+1} \cdot p_{n+1}$$

$$\rho_{n+1} = \frac{(g, p_{n+1})}{(Ap_{n+1}, p_{n+1})}$$

$$J(u_n + \rho_{n+1} p_{n+1}) = J(u_n)[1 - E_n]$$

$$E_n = \frac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1}); (A^{-1} g_n, g_n)}$$

Notations

La valeur du paramètre ρ qui permet de passer de u_n à u_{n+1} est *toujours* désignée par ρ_{n+1}; dans les cas précédents on obtenait ρ_{n+1} comme la valeur de ρ pour laquelle un trinôme de second degré était minimum. Cette valeur sera désignée par ρ^m_{n+1}

(1.20) $$\rho^m_{n+1} = \frac{(g_n, p_{n+1})}{(Ap_{n+1}, p_{n+1})}$$

Nous allons maintenant définir le nouveau ρ_{n+1}.

Soit γ un nombre donné dans]0, 1]

$$0 < \gamma \leqslant 1$$

et soit un nombre ρ_{n+1} qui vérifie

(1.21) $$\gamma\rho^m_{n+1} \leqslant \rho_{n+1} \leqslant (2 - \gamma) \cdot \rho^m_{n+1}$$

Nous allons poser maintenant

(1.22) $$u_{n+1} = u_n + \rho_{n+1} \cdot p_{n+1}$$

Un calcul élémentaire montre que

$$J(u_n) - J(u_n + \rho_{n+1}p_{n+1}) \geqslant \frac{\rho_{n+1}}{\rho^m_{n+1}} [J(u_n) - J(u_n + \rho^m_{n+1}p_{n+1})]$$

d'où il vient

$$J(u_n) - J(u_n + \rho_{n+1} \cdot p_{n+1}) \geqslant \gamma J(u_n) \cdot E_n$$

et donc

(1.23) $$J(u_{n+1}) \leqslant J(u_n)[1 - \gamma E_n]$$

formule que l'on comparera à (1.7) et aux inégalités suivantes.

⇒ On peut donc reprendre toutes les méthodes précédentes et introduire la modification apportée par les formules (1.21) et (1.22) au choix de u_{n+1} : on obtiendra encore une méthode convergente.

REMARQUE 1.2.

M. L. Stein [16] a utilisé la variante précédente dans le cas de la méthode du gradient. En choisissant ρ_{n+1} vérifiant

$$0{,}7 \cdot \rho^m_{n+1} \leqslant \rho_{n+1} \leqslant 0{,}9 \cdot \rho^m_{n+1}$$

il a remarqué ceci : l'effet des « zig-zags » de la méthode du gradient est atténué et la convergence est plus rapide.

A. D. Both [3] a proposé une variante du même genre.

Nous nous proposons maintenant de donner une variante de la méthode du gradient qui nous conduira à des inégalités analogues à celles de M. L. Stein; rappelons que dans la méthode du gradient on a :

$$\begin{cases} u_{n+1} = u_n + \overset{m}{\rho}_{n+1} \cdot g_n = u_n - \overset{m}{\rho}_{n+1} \cdot (Au_n - f) \\ \overset{m}{\rho}_{n+1} = \dfrac{\|g_n\|^2}{(Ag_n, g_n)} \end{cases}$$

Introduisons la variante suivante :

$$u_{n+1} = u_n - \rho_{n+1}(Au_n - f)$$

et cherchons $\rho_{n+1} = \theta =$ constante pour laquelle la méthode est convergente. Il suffit pour cela que l'application $u \to w = (I - \theta A)u$ soit une contraction stricte. Or compte tenu des propriétés de A, on a :

$$\|w\|^2 = \|u\|^2 - 2\theta(Au, u) + \theta^2 \|Au\|^2$$

$$\|w\|^2 \leqslant \|u\|^2 - 2\theta\alpha \|u\|^2 + \theta^2 M^2 \|u\|^2$$

ou encore :

$$\begin{cases} \|w\|^2 \leqslant T(\theta) \cdot \|u\|^2 \\ T(\theta) = 1 - 2\theta\alpha + \theta^2 M^2 \end{cases}$$

le meilleur choix de θ est

$$\theta = \theta_{\text{opt}} = \frac{\alpha}{M^2}$$

Alors $|T(\theta_{\text{opt}})| < 1$ et la variante de la méthode du gradient est convergente. Nous avons :

$$\frac{\rho_{n+1}}{\overset{m}{\rho}_{n+1}} = \frac{\theta}{\overset{m}{\rho}_{n+1}} = \frac{\alpha}{M^2} \cdot \frac{(Ag_n, g_n)}{\|g_n\|^2}$$

d'où il vient facilement

$$\frac{\alpha^2}{M^2} \leqslant \frac{\rho_{n+1}}{\overset{m}{\rho}_{n+1}} \leqslant \frac{\alpha}{M}$$

relation que l'on comparera à celle proposée par Stein.

1.10. La méthode de descente et des boules

On pose toujours

$$u_{n+1}(\rho) = u_n + \rho \cdot p_{n+1}$$

et on introduit un nombre r arbitraire positif. Dans cette méthode, u_{n+1} sera le point de la forme $u_{n+1}(\rho)$, soumis à la contrainte $\|u_{n+1}(\rho) - u_n\| \leqslant r$, et qui rend $J(u_{n+1}(\rho))$ minimum. Désignons par ρ_{n+1} la valeur optimale de ρ, nous avons donc :

$$u_{n+1} = u_n + \rho_{n+1} \cdot p_{n+1}$$

et si ρ_{n+1}^m désigne toujours la valeur optimale de ρ (dans le cas sans contrainte)

$$\rho_{n+1}^m = \frac{(g_n, p_{n+1})}{(Ap_{n+1}, p_{n+1})}$$

nous avons les deux cas suivants :

1er cas :

$$\|\rho_{n+1}^m \cdot p_{n+1}\| \leqslant r; \qquad \text{alors}$$

(1.24) $$\rho_{n+1} = \rho_{n+1}^m$$

2e cas :

$$\|\rho_{n+1}^m p_{n+1}\| > r; \qquad \text{alors}$$

$$\rho_{n+1} < \rho_{n+1}^m$$

et on a

$$\|\rho_{n+1} p_{n+1}\| = r$$

d'où

(1.25) $$\frac{\rho_{n+1}}{\rho_{n+1}^m} = \frac{r}{\rho_{n+1}^m \, \|p_{n+1}\|}$$

Montrons que $\rho_{n+1}^m \, \|p_{n+1}\|$ est borné supérieurement; en effet : d'après le n° 1.2, on a

$$J(u_n + \rho_{n+1}^m p_{n+1}) \leqslant J(u_n) \leqslant J(u_0)$$

d'où, compte tenu de la coercivité

$$\|u_n + \rho_{n+1}^m p_{n+1}\| \leqslant \sqrt{\frac{J(u_0)}{\alpha}}$$

$$\|u_n\| \leqslant \sqrt{\frac{J(u_0)}{\alpha}}$$

et

$$\|\rho_{n+1}^m p_{n+1}\| \leqslant 2\sqrt{\frac{J(u_0)}{\alpha}}$$

par suite, dans ce 2e cas, la formule (1.25) conduit à

(1.26) $$\frac{\rho_{n+1}}{\rho_{n+1}^m} \geqslant \frac{r}{2}\sqrt{\frac{\alpha}{J(u_0)}}$$

En regroupant les deux cas, les formules (1.24) et (1.25) montrent que

(1.27) $$\frac{r}{2}\sqrt{\frac{\alpha}{J(u_0)}}\,\rho_{n+1}^m \leqslant \rho_{n+1} \leqslant \rho_{n+1}^m$$

ce qui prouve que nous sommes dans un cas particulier de la variante des méthodes de descente (voir 1.9).

REMARQUE 1.3.

Avec une méthode convergente, pour n assez grand on aura

$$\|u_n + \rho_{n+1} p_{n+1} - u_n\| \leqslant r$$

et par suite $\rho_{n+1} = \rho_{n+1}^m$. L'influence de l'introduction de r ne se fera sentir que pour les premiers itérés. C'est ce qu'avait remarqué M. L. Stein. En fait, dans la méthode du gradient, par exemple, pour r assez petit, u_{n+1} est « à peu près » le point de la boule $\|v - u_n\| \leqslant r$, en lequel $J(v)$ est minimum.

REMARQUE 1.4.

On pourrait introduire un rayon r, variable à chaque itération, de façon que $\frac{\rho_{n+1}}{\rho_{n+1}^m} \geqslant \gamma > 0$; et cela conduirait encore à une méthode convergente.

2. LE CAS AVEC CONTRAINTES

Nous proposons une généralisation de l'algorithme de M. Franck et P. Wolfe [10]; rappelons les généralisations de A. Auslender et F. Brodeau [1], V. F. Dem'janov and A. M. Rubinov [7] ; Y. Haugazeau [11], B. T. Poljak [14] et M. Valadier [17].

2.1. Position du problème

L'espace V, l'opérateur A et l'élément f sont donnés, comme dans le cas précédent, ils vérifient les mêmes hypothèses. On donne *de plus, un sous-ensemble convexe fermé* K de V; on pose

$$J(v) = (Av, v) - 2(f, v)$$

(A une constante additive près, la fonction économique est identique à la précédente.)

Il s'agit maintenant de déterminer un élément $u \in K$ vérifiant

$$J(u) \leqslant J(v) \qquad \forall v \in K. \tag{2.1}$$

On sait que l'élément u existe, qu'il est unique et qu'il est caractérisé par

$$(Au - f, v - u) \geqslant 0 \qquad \forall v \in K. \tag{2.1'}$$

2.2. Le principe de l'approximation de u

On va construire une suite d'itérés u_n; le passage de u_n à u_{n+1} se fera de la façon suivante : on choisit une direction de descente p_{n+1}, u_{n+1} sera alors le *point de* K de la forme $u_n + \rho p_{n+1}$, $\rho \in \mathbf{R}$, pour lequel $J(v)$ est minimum.

Notations

Posons

$$g_n = -(Au_n - f)$$

$$u_{n+1}(\rho) = u_n + \rho p_{n+1}$$

(2.2) $$\rho^m_{n+1} = \frac{(g_n, p_{n+1})}{(Ap_{n+1}, p_{n+1})}$$

(ρ^m_{n+1} est déterminé par : $J(u_{n+1}(\rho^m_{n+1})) \leqslant J(u_{n+1}(\rho)) \quad \forall \rho \in R$).

(2.3) $$\rho^1_{n+1} = \sup_{\substack{\rho \geqslant 0 \\ u_{n+1}(\rho) \in K}} \rho$$

Nous avons

(2.4) $$J(u_n + \rho p_{n+1}) = J(u_n) - 2\rho(g_n, p_{n+1}) + \rho^2(Ap_{n+1}, p_{n+1})$$

et par conséquent le meilleur choix possible de ρ [bien entendu lorsque (g_n, p_{n+1}) est positif] est donné par

(2.5) $$\rho_{n+1} = \min(\rho^1_{n+1}, \rho^m_{n+1})$$

On prend

(2.6) $$u_{n+1} = u_n + \rho_{n+1} \cdot p_{n+1}$$

Le choix de p_{n+1} : On donne maintenant un nombre strictement positif r et on définit K_n par

$$v \in K_n \Leftrightarrow v \in K \qquad \text{et} \qquad \|v - u_n\| \leqslant r$$

Dans une grande famille de méthodes on peut définir p_{n+1} de la façon suivante : soit un élément $z_n \in \mathbf{K}$, et qui vérifie

(2.7) $$\begin{cases} (g_n, z_n) \geqslant (g_n, v) & \forall v \in K_n \\ \|z_n\| \leqslant c, & c = \text{constante indépendante de } n. \end{cases}$$

Posons alors

(2.8) $$p_{n+1} = z_n - u_n$$

Si $(g_n, p_{n+1}) = 0$, alors avec (2.7) et (2.8) il vient

$$(Au_n - f, v - u_n) \geqslant 0 \qquad \forall v \in K_n$$

et avec (2.1′) on vérifierait que $u_n = u$.

Le cas où la conclusion n'est pas immédiate est donc celui où $(g_n, p_{n+1}) > 0$, c'est-à-dire le cas où il y a effectivement une direction de descente. Notons ceci : dans ce cas on a $p_{n+1} \neq 0$ et donc

$$z_n \neq u_n$$

et puisque $z_n \in K$, il vient [voir (2.3)]

$$\rho^1_{n+1} \geqslant 1 \tag{2.9}$$

En résumé : On se donne u_0; on choisit z_n vérifiant (2.7), on définit ensuite p_{n+1} par (2.8), ρ_{n+1} par (2.5) et enfin u_{n+1} par (2.6).

Théorème 2.1. — *Lorsque $n \to \infty$, la suite u_n converge fortement dans V, vers la solution u du problème* (2.1)*; et ceci est valable pour tout choix de $z_n \in K$ et vérifiant* (2.7).

Démonstration.

1^er^ point : La suite $J(u_n)$ est décroissante.

Dans le cas où $(g_n, p_{n+1}) = 0$, nous savons que $u_n = u$; dans le cas où $(g_n, p_{n+1}) > 0$, à partir de la formule (2.4) on obtient

$$J(u_{n+1}) \leqslant J(u_n)$$

or,

$$J(u_n) \geqslant J(u) \qquad \forall n$$

et donc

$$J(u_0) \geqslant \ldots \geqslant J(u_n) \geqslant \ldots \geqslant J(u) \tag{2.10}$$

La suite $J(u_n)$, décroissante et bornée, est convergente. Une première conséquence est la suivante (il suffit d'utiliser la coercivité de A) :

$$\begin{cases} \|u_n\| \leqslant c_1 & \forall n \\ c_1 = \text{constante} > 0, & \text{indépendante de } n \end{cases} \tag{2.11}$$

Une deuxième conséquence est :

$$\lim_{n\to\infty} (J(u_n) - J(u_{n+1})) = 0 \tag{2.12}$$

2^e^ point : On va montrer que $\lim\limits_{n\to\infty} (g_n, p_{n+1}) = 0$;

1^er^ cas : $\rho_{n+1} = \rho^m_{n+1}$; Alors comme cela a été vu au n° 1 (voir 1.6),

$$J(u_n) - J(u_{n+1}) = \frac{(g_n, p_{n+1})^2}{(Ap_{n+1}, p_{n+1})} \tag{2.13}$$

Les suites z_n et u_n étant bornées, la suite p_n est donc bornée; d'où compte tenu de la continuité de A

$$J(u_n) - J(u_{n+1}) \geqslant c_2 (g_n, p_{n+1})^2 \tag{2.14}$$

2^e^ cas : $\rho^1_{n+1} = \rho_{n+1}$. donc $\rho^1_{n+1} \leqslant \rho^m_{n+1}$.

Rappelons que

$$1 \leqslant \rho^1_{n+1}$$

d'où

$$(2.15) \qquad 1 \leqslant \rho^1_{n+1} \leqslant \rho^m_{n+1} = \frac{(g_n, p_{n+1})}{(Ap_{n+1}, p_{n+1})}$$

et puisque

$$J(u_n) - J(u_n + \rho p_{n+1}) = 2\rho(g_n, p_{n+1}) - \rho^2(Ap_{n+1}, p_{n+1})$$

on a

$$J(u_n) - J(u_n + p_{n+1}) \leqslant J(u_n) - J(u_n + \rho^1_{n+1} p_{n+1})$$

c'est-à-dire

$$2(g_n, p_{n+1}) - (Ap_{n+1}, p_{n+1}) \leqslant J(u_n) - J(u_{n+1})$$

Avec (2.15) cela donne

$$(2.16) \qquad (g_n, p_{n+1}) \leqslant J(u_n) - J(u_{n+1})$$

Finalement en combinant (2.13) avec soit (2.14), soit (2.16) il vient

$$(2.17) \qquad \lim\,(g_n, p_{n+1}) = 0$$

3e point : Convergence de la suite u_n.

On a la relation algébrique

$$(2.18) \qquad (A(u_n - u), u_n - u) + \{ (Au, u_n - u) - (f, u_n - u) \} + \{ (f, u_n - u) - (Au_n, u_n - u) \} = 0$$

d'après (2.1′), on a

$$(Au, u_n - u) - (f, u_n - u) \geqslant 0$$

et par suite

$$(A(u_n - u), u_n - u) \leqslant (Au_n - f, u_n - u)$$

ou encore

$$(2.19) \qquad (A(u_n - u), u_n - u) \leqslant (g_n, u - u_n)$$

1er cas : $u \in K_n$, alors par la définition même de z_n puis de p_{n+1}

$$(g_n, u - u_n) \leqslant (g_n, p_{n+1})$$

et donc

$$(2.20) \qquad (A(u_n - u), u_n - u) \leqslant (g_n, p_{n+1})$$

2e cas : $u \notin K_n$, et alors $\|u - u_n\| > r$; alors

$$(g_n, u - u_n) = \frac{\|u - u_n\|}{r} \left(g_n, \frac{r}{\|u - u_n\|}(u - u_n) \right) \leqslant \frac{\|u - u_n\|}{r}(g_n, p_{n+1})$$

et puisque u et u_n sont bornés

$$(2.21) \qquad \begin{cases} (g_n, u - u_n) \leqslant c_3(g_n, p_{n+1}) \\ c_3 = \text{constante} > 0, \quad \text{indépendante de } n \end{cases}$$

Finalement avec (2.17) et soit (2.20), soit (2.21) il vient

$$\lim_{n\to\infty} (A(u_n - u), u_n - u) = 0$$

ce qui avec l'hypothèse de coercivité entraîne

$$\lim_{n\to\infty} \|u_n - u\| = 0$$

2.3. Remarques diverses

REMARQUE 2.1.

Il n'y aurait aucune difficulté à obtenir un théorème analogue au théorème 2.1 en prenant

$$\rho_{n+1} = \min\left(\overset{m}{\rho}_{n+1}, \overset{1}{\rho}_{n+1}, \frac{r}{\|p_{n+1}\|}\right)$$

(c'est-à-dire qu'on introduit dans le cas précédent la contrainte

$$\|u_{n+1} - u_n\| \leqslant r);$$

alors pour r assez petit u_{n+1} est approximativement le point de K_n pour lequel $J(v)$ est minimum. L'influence de la contrainte $\|u_{n+1} - u_n\| \leqslant r$, ne se fera bien entendu sentir que pour les premiers itérés.

REMARQUE 2.2.

Si K est borné, on peut ne pas faire intervenir r en le choisissant assez grand : alors $K = K_n$. On est alors ramené à la méthode d'Haugazeau [11].

REMARQUE 2.3.

Si K n'est pas borné, on peut se ramener au cas précédent. En effet, soit $u_0 \in K$, alors

$$J(u) \leqslant J(u_0)$$

et cela conduit avec la coercivité à :

$$\|u\| \leqslant c(u_0)$$

Soit alors l'ensemble convexe fermé $\tilde{K}$ défini par

$$v \in \tilde{K} \Leftrightarrow v \in K, \qquad \|v\| \leqslant c(u_0)$$

Il est clair qu'on peut remplacer le problème initial par le suivant : déterminer $u \in \tilde{K}$ tel que

$$J(u) \leqslant J(v) \qquad \forall v \in \tilde{K}.$$

REMARQUE 2.4.

Dans la condition (2.7) que vérifie z_n, on a introduit K_n, définie par

$$v \in K_n \Leftrightarrow v \in K, \qquad \|v - u_n\| \in B_r$$

où B_r désigne la boule centrée à l'origine et de rayon r. Il est évidemment possible de changer cette boule en une autre boule définie à l'aide d'une norme équivalente à la norme initiale. Cela pourra être intéressant dans le cas de la dimension finie.

REMARQUE 2.5 : *Stabilité du choix de z_n.*

Introduisons $w_n \in K$ tel que

$$\begin{cases} (g_n, w_n) > (g_n, u_n) \\ g_n \in K, \\ \lim\limits_{n \to \infty} \|z_n - w_n\| = 0 \end{cases}$$

où z_n satisfait comme précédemment à (2.7).

Dans la construction précédente de la suite u_n, nous remplaçons z_n par w_n. On peut reproduire textuellement le 1er point et le 2e point de la démonstration du théorème 2.1; on arrivera à :

$$\lim_{n \to \infty} (g_n, w_n - u_n) = 0$$

ce qui avec les hypothèses faites sur w_n entraîne

$$\lim (g_n, z_n - u_n) = 0$$

et on peut alors utiliser le 3e point, d'où la convergence de la nouvelle suite u_n vers u.

Au point de vue pratique, on pourra obtenir z_n par une méthode itérative : on construira une suite z_n^p qui converge vers z_n lorsque $p \to \infty$; l'élément w_n pourra être alors un certain élément z_n^p de la suite.

S'il n'est pas possible de trouver $w_n \in K$ tel que

$$(g_n, w_n) > (g_n, u_n)$$

c'est qu'alors

$$(g_n, u_n - w) \geqslant 0 \qquad \forall w \in K$$

avec (2.1'), cela entraîne

$$u_n = u.$$

2.4. Le choix de z_n

Rappelons les relations importantes dans la construction de u_{n+1} :

$$\begin{aligned} & p_{n+1} = z_n - u_n, \qquad z_n \in K \\ & (g_n, z_n) \geqslant (g_n, v) \qquad \forall v \in K_n \\ & u_{n+1}(\rho) = u_n + \rho \cdot p_{n+1} \\ & J(u_{n+1}(\rho)) = J(u_n) - 2\rho(g_n, p_{n+1}) + \rho^2(Ap_{n+1}, p_{n+1}) \end{aligned}$$

Nous allons indiquer maintenant plusieurs choix de z_n :

2.4.1. La méthode du gradient modifiée localement

On désigne par v_n un élément de K_n tel que

$$(g_n, v_n) \geqslant (g_n, v) \qquad \forall v \in K_n \tag{2.22}$$

cet élément existe, mais n'est pas nécessairement unique. Nous prenons

$$z_n = v_n \tag{2.23}$$

et grâce à (2.22) la méthode est convergente.

Lorsque K est borné et r suffisamment grand pour qu'il n'intervienne pas de façon explicite, nous retrouvons la méthode d'Haugazeau [11].

Lorsque le convexe K est défini par un nombre fini de contraintes linéaires et que r est suffisamment petit nous retrouvons la méthode du gradient projeté de Rosen [15].

2.4.2. La méthode du gradient conjugué modifiée

Nous posons :

$$\begin{cases} z_n(\lambda) = v_n + \lambda p_n \\ \lambda = 0 \text{ si } (g_n, p_n) \leqslant 0, \lambda \geqslant 0 \text{ si } (g_n, p_n) > 0. \end{cases} \tag{2.24}$$

et alors

$$(g_n, z_n(\lambda)) = (g_n, v_n) \geqslant (g_n, v) \qquad \forall v \in K_n$$

Nous choisirons λ « au mieux » c'est-à-dire de façon que $z_n(\lambda) \in K$ et que $(A(v + \lambda p_n - u_n), (v_n + \lambda p_n - u_n)$ soit minimum. Lorsque $\lambda = 0$, il s'agit de la méthode du 2.4.1; en choisissant le « meilleur » nous ne pouvons qu'améliorer la convergence de la méthode.

2.4.3. Méthode du gradient axial modifiée

Nous sommes maintenant dans le cas de la dimension finie : $V = R^q$, une base de V est désignée par $e_1, \ldots, e_q$; nous introduisons les points

$$u_{n+\frac{i}{q}}, \; i = 1, \ldots, q.$$

Nous poserons

$$z_{n+\frac{i+1}{q}}(\lambda) = v_{n+\frac{i}{q}} + \lambda e_{i+1}$$

et nous choisirons λ de façon que $z_{n+\frac{i+1}{q}} \in K$, $\lambda(e_{i+1}, g_n) \geqslant 0$ $\Big($pour que $\Big(g_{n\frac{i}{q}}, z_{n+\frac{i+1}{q}}\Big) \geqslant \Big(g_{n+\frac{i}{q}}, v\Big)$ $\forall v \in K_{n+\frac{i}{q}}\Big)$ et que

$$\Big(A(z_{n+\frac{i+1}{q}}(\lambda) - u_{n+\frac{i}{q}}), z_{n+\frac{i+1}{q}}(\lambda) - u_{n+\frac{i}{q}}\Big)$$

soit minimum.

Il est clair [à partir du développement de $J(u_{n+1}(\rho))$] que cette méthode est une amélioration de la méthode du 2.4.1.

Remarque 2.6.

Les hypothèses sur z_n sont

$$\begin{cases} z_n \in K & \\ \|z_n\| \leqslant c & \forall n \\ (g_n, z_n) \geqslant (g_n, v) & \forall v \in K_n \end{cases}$$

et donc en particulier, lorsque $v = v_n$ (voir 2.4.1)

$$\begin{cases} z_n \in K & \\ \|z_n\| \leqslant c & \forall n \\ (g_n, z_n) \geqslant (g_n, v_n) & \end{cases}$$

et grâce à (2.22) il y a équivalence entre les 2 groupes précédents de relations.

Situation des divers éléments.

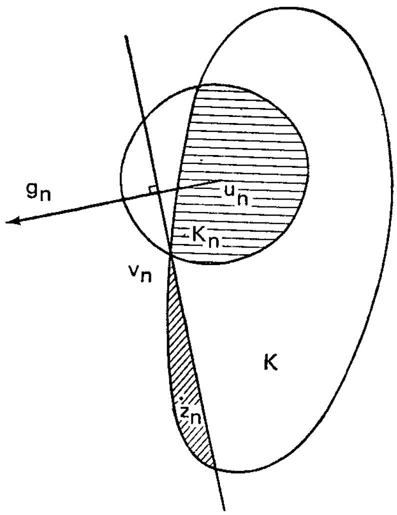

Figure 1

3. APPLICATION AU CONTROLE OPTIMAL

Nous allons nous placer dans le cadre de [13]. Pour faciliter la compréhension du texte nous allons reproduire ici ce cadre :

Tous les espaces utilisés sont des espaces de Hilbert réels; de façon générale $(\ ,\ )_X$ désigne le produit scalaire dans X et $(\ ,\ )$ le produit scalaire entre X et son dual X' non identifié à X.

Le système à contrôler est décrit par $\Lambda \in \mathcal{L}(Y, \Phi)$ avec : Λ est un isomorphisme de Y sur Φ.

Soit $\mathcal{U}$ l'espace des contrôles, $E \in \mathcal{L}(\mathcal{U}, \Phi)$, l donné dans Φ; pour un contrôle $u \in \mathcal{U}$, l'état $y(u)$ du système est donné par :

$$\Lambda y(u) = 1 + Eu, \qquad y(u) \in \gamma \tag{3.1}$$

Soit $L \in \mathcal{L}(Y, Z)$: l'observation est $Ly(u)$; la fonction coût est :

$$J(u) = (M(Ly(u) - z_d), \quad Ly(u) - z_d)_Z + (Nu, u)_U \tag{3.2}$$

où

$$\begin{cases} z_d \text{ est donné dans } Z \\ M \in \mathcal{L}(Z, Z) \\ N \in \mathcal{L}(U, U) \end{cases} \tag{3.3}$$

Soit K l'ensemble des contrôles admissibles; on suppose : K est un ensemble convexe fermé dans U. Un contrôle u est dit optimal si

$$J(u) = \operatorname*{Inf}_{v \in K} J(v)$$

Un système adjoint est défini par la donnée d'un nouvel espace $\mathcal{T}$ et des opérateurs Λ_1 et L_1 vérifiant :

Λ_1 est un isomorphisme de $\mathcal{T}$ sur Z

$L_1 \in \mathcal{L}(\mathcal{T}, \Phi)$

$$(L_1 p, \Lambda y) = (\Lambda_1 p, Ly) \qquad \forall p \in \mathcal{T} \quad , \quad \forall y \in Y \tag{3.4}$$

l'état adjoint est donné par

$$\Lambda_1 p(u) = M(Ly(u) - z_d), \qquad p(u) \in \mathcal{T} \tag{3.5}$$

Nous faisons les hypothèses suivantes

(3.6) M et N sont des opérateurs symétriques

(3.7) $G^*MG + N$ est défini positif dans $\mathcal{L}(U, U)$

où $G = L\Lambda^{-1}E$.

Alors le théorème suivant est démontré dans [13].

Théorème 3.1. — i) *Il existe un contrôle optimal et un seul.*

ii) *Le contrôle optimal est caractérisé par :*

$$(3.8) \qquad (M(Ly(u) - z_d), Ly(v) - Ly(u)) + (Nu, v - u) \geqslant 0 \qquad \forall v \in K$$

ou sous une autre forme (équivalente; on introduit les opérateurs $L_1, \ldots$) :

$$(3.8') \qquad (E^* L_1 p(u) + Nu, v - u) \leqslant 0 \qquad \forall v \in K$$

⇒ *Nous nous proposons d'appliquer la méthode du gradient modifiée localement (voir n° 2.4.1) à l'approximation du contrôle optimal u.*

Cela est en effet possible, car après élimination de $y(u)$, on retrouve le cadre n° 2, toutes les hypothèses étant satisfaites; nous avons

$$(3.9) \quad \begin{cases} J(u + \rho(v - u)) = J(u) \\ + \rho \, \{ (M(Ly(u) - z_d), Ly(v) - Ly(u)) + N(u, v - u) \} \\ + \rho^2 \, \{ (M(Ly(v) - Ly(u)), Ly(v) - Ly(u)) + (N(v - u, v - u) \} \end{cases}$$

Nous donnons maintenant la construction de la suite u_n :

i) u_0 est donné dans K,

ii) on détermine v_n en introduisant r et $K_n = \{ v \mid v \in K, \|v - u_n\| \leqslant r \}$; v_n est alors défini par

$$(3.10) \qquad (M(Ly(u_n) - z_d), Ly(v_n)) + (Nu_n, v_n) \leqslant (M(Ly(u_n) - z_d), Ly(v)) + (Nu_n, v)$$

iii) on pose :

$$u_{n+1}(\rho) = u_n + \rho(v_n - u_n)$$

et on détermine ρ_n de façon que :

$$(3.11) \qquad \begin{cases} J(u_{n+1}(\rho_n)) \leqslant J(u_{n+1}(\rho)) & \forall u_{n+1}(\rho) \in K \\ u_{n+1}(\rho_n) \in K \end{cases}$$

On pose

$$(3.12) \qquad u_{n+1} = u_{n+1}(\rho_n)$$

et le théorème 2.1 conduit au

Théorème 3.2. — *La suite u_n converge fortement vers le contrôle optimal u lorsque $n \to +\infty$.*

Nous allons examiner maintenant les calculs à effectuer pour obtenir v_n, ρ_n, u_{n+1} à partir de u_n.

Nous avons, grâce à (3.5) :

$$(M(Ly(u) - z_d), Ly(v)) + (Nu, v) = (\Lambda_1 p(u), Ly(v)) + (Nu, v)$$

puis grâce à (3.4)

$$(\Lambda_1 p(u), Ly(v)) + (Nu, v) = (L_1 p(u), Ay(v)) + (Nu, v)$$

puis grâce à (3.1)

$$(L_1 p(u), \Lambda y(v)) + (Nu, v) = (L_1 p(u), Y) + (L_1 p(u), Ev) + (Nu, v)$$

Finalement

$$(3.13) \qquad (M(Ly(u) - z_d), Ly(v)) + (Nu, v) = (L_1 p(u), Ev) + (Nu, v) + (L_1 p(u), 1)$$

Si nous posons

$$y_n = y(u_n)$$

$$p_n = p(u_n)$$

la relation (3.10) se traduit, en utilisant (3.13) par

(3.14)

$$(L_1 p_n, Ev_n) + (Nu_n, v_n) \leqslant (L_1 p_n, Ev) + (Nu_n, v) \qquad \forall v \in K_n$$

ou encore par :

$$(3.14') \quad (E^* L_1 p_n + Nu_n, v_n)_U \leqslant (E^* L_1 p_n + Nu_n, v)_U \qquad \forall v \in K_n$$

Ainsi le calcul de v_n nécessite la connaissance de p_n et donc de y_n.

Supposons résolu le problème de la recherche de v_n vérifiant (3.14'). Nous devons maintenant calculer ρ_n : on sait qu'il s'introduit alors ρ_n^m vérifiant :

$$(3.15) \begin{cases} \{ (M(Ly_n - z_d), Ly(v_n) - Ly_n) + (Nu_n, v_n - u_n) \} + \\ + \rho_n^m \{ (ML(y(v_n) - y_n, L(y(v_n) - y_n + (N(v_n - u_n), v_n - u_n) \} = 0 \end{cases}$$

Nous sommes donc amenés à calculer $y(v_n)$ (ou mieux : $y(v_n) - y_n$) solution de :

$$(y(v_n) - y(u_n)) = E(v_n - u_n)$$

Après le calcul de ρ_n il vient trivialement u_{n+1}.

En résumé :

— pour chaque itéré, il faut résoudre deux problèmes directs (calculs de $y(u_n)$ et de $y(v_n)$) et un problème adjoint (calcul de $p(u_n)$);

— de plus il faut résoudre le problème (3.14').

BIBLIOGRAPHIE

[1] A. AUSLENDER et F. BRODEAU, *Convergence d'un algorithme de Frank et Wolfe appliqué à un problème de contrôle.* Revue Française d'Informatique et de Recherche Opérationnelle – 1968, n°7-R1.

[2] BALAKRISHNAN and HSIEN, *Optimum system synthesis,* Conference. Dayton, Ohio, September 1962.

[3] A. D. BOOTH, *Numerical methods,* London Butterworths Scientific Publication, 1957.

[4] H. BREZIS et M. SIBONY, *Méthode d'approximation et d'itération pour les opérateurs monotones* (à paraître).

[5] F. E. BROWDER, Existence and uniquences theorems for solutions of non linear Bounday value problems. *Proceedings of Symposia in Applied Math.,* vol. XVII, A.M.S., 1965.

[6] J. W. DANIEL, The conjugate gradient method for linear and non linear operator equations, *SIAM Num. Anal.,* vol. 4, n° 1, 1967.

[7] V. F. DEM' JANOV and A. M. RUBINOV, On the problem of minimization of a smooth functional with convex constraints, *Soviet Mathematics,* vol. 8, n° 1, 1965.

[8] R. FLETCHER and M. J. D. POWELL, *The computer journal,* 1963, 6, 163-168.

[9] R. FLETCHER and C. M. REEVES, Function minimization by conjugate gradients, *Computing Journal,* 1964, 7.

[10] M. FRANK and P. WOLFE, On algorithm for quadratic programming, *Naval Research Logistics Quaterly,* 1956.

[11] Y. HAUGAZEAU, Sur la minimisation des formes quadratiques avec contraintes, *C. R. Acad. Sci.,* Paris, 2 novembre 1966.

[12] M. HESTENES and E. STIEFFEL, Method of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Stand.,* Sect. B, 49, 1952.

[13] J. L. LIONS, Sur le contrôle optimal de systèmes décrits par des équations aux dérivées partielles, *C. R. Acad. Sc. Paris,* 7, 14, 21 novembre 1966.

[14] B. T. POLJAK, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Soviet. Math.,* 1966, t. 166, n° 2.

[15] J. B. ROSEN, The gradient projection method for non linear programming *SIAM,* vol. 8, n° 1, 1960.

[16] H. L. STEIN, Gradient methods in the solution of systems of linear equations, *Nat. Bur. of Stand. NAML,* Rep. 52-7, 1951.

[17] M. VALADIER, Extension d'un algorithme de Frank et Wolfe, *Revue Française de recherche opérationnelle,* n° 36, 1965.