Revue française d'informatique et de recherche opérationnelle. Série rouge

Auslender

Méthodes du second ordre dans les problèmes d'optimisation avec contraintes

Revue française d'informatique et de recherche opérationnelle. Série rouge, tome 3, nº R2 (1969), p. 27-42

<http://www.numdam.org/item?id=M2AN_1969__3_2_27_0>

R.I.R.O.
(3e année, N° R-2/1969, p. 27-42)

METHODES DU SECOND ORDRE DANS LES PROBLEMES D'OPTIMISATION AVEC CONTRAINTES

par AUSLENDER

Résumé. — *Il est présenté une théorie concernant des méthodes du second ordre pour résoudre les problèmes d'optimisation avec contraintes dans des espaces normés. En particulier il est proposé trois types de méthodes et il est donné des théorèmes de convergence dans lesquels on s'attache à préciser la rapidité de convergence. Enfin il est montré en application comment la théorie développée peut s'appliquer à la résolution des problèmes de contrôle.*

Les méthodes de gradients fournissent rapidement des suites minimisantes c'est-à-dire des suites dont les valeurs de la fonction à optimiser convergent rapidement vers la solution optimale.

Dans les espaces normés, ces suites ne convergent pas en général pour la topologie de la norme vers une solution optimale et lorsqu'elles convergent on constate expérimentalement que cette convergence est lente.

Ceci provient du fait que les directions admissibles utilisées sont obtenues à partir d'approximation du premier ordre sur la fonction f à optimiser. D'où l'idée d'obtenir des directions admissibles à partir d'approximation du second ordre.

Cette idée étant plus évidente lorsque le problème d'optimisation est sans contraintes, a été largement développée pour des problèmes sans contraintes (3), (4), (5), (6), (7).

Dans (7) Poljack généralise la méthode de Newton à des problèmes d'optimisation avec contraintes. Ceci n'est guère satisfaisant car l'algorithme ne converge que si le point de départ est choisi près de l'optimum.

Nous nous proposons ici de développer des méthodes du second ordre appliquées à des problèmes d'optimisation sous contraintes.

Au paragraphe 2 on propose trois types de méthodes qui recouvrent en particulier les méthodes existantes citées ci-dessus.

Au paragraphe 3 on donne des théorèmes de convergence dans lesquels on s'attache à préciser la rapidité de convergence.

Au paragraphe 4 on montre en application comment la théorie développée peut s'appliquer pour la résolution des problèmes de contrôle.

I. NOTATIONS ET RAPPELS

Soient :

X, un espace vectoriel normé.

X^* le dual topologique.

$L(X, X^*)$ l'ensemble des applications linéaires continues de X dans X^*.

Q un convexe borné complet de X et x_0 un point de Q.

f une fonction numérique définie et continûment dérivable sur X.

f' sa dérivée, $\langle f'(u), v\rangle$ la valeur de $f'(u)$ au point v.

σ un scalaire de l'intervalle ouvert $]0, 1/2[$.

$B(x, \rho) = \{ y \in Q : \|x - y\| \leqslant \rho \}$.

J l'isomorphisme canonique de X dans X^* lorsque X est Hilbertien.

A tout point x de Q on associe un opérateur $T_x \in L(X, X^*)$ symétrique, c'est-à-dire qui vérifie quel que soit u et v appartenant à Q la relation :

$$\langle T_x \cdot u, v\rangle = \langle T_x \cdot v, u\rangle$$

où $T_x \cdot u$ indique la valeur de T_x au point u.

On rappelle (8) :

1) qu'une fonction numérique g est fortement convexe sur Q s'il existe $\gamma > 0$ tels que quels que soient x_1, x_2 appartenant à Q, quel que soit λ appartenant à $[0, 1]$ on ait :

$$g(\lambda x_1 + (1-\lambda)x_2) \leqslant g(x_1) + (1-\lambda)g(x_2) - \lambda(1-\lambda)\gamma \|x_1 - x_2\|^2$$

2) pour qu'une fonction numérique g soit fortement convexe sur Q il faut et il suffit qu'il existe $\gamma > 0$ tel que quels que soient x_1, x_2 appartenant à Q on ait :

si g est dérivable :

$$g(x_2) - g(x_1) \geqslant \langle g'(x_1), x_2 - x_1\rangle + \gamma \|x_2 - x_1\|^2$$

si g est deux fois dérivable :

$$\langle g''(x_1) \cdot x_2 - x_1, x_2 - x_1\rangle \geqslant 2\gamma \|x_2 - x_1\|^2$$

On suppose désormais que $f^* = \inf(f(x) \mid x \in Q)$ est fini et l'on se propose de résoudre *le problème P d'optimisation :*

$P =$ « Trouver une suite $\{ x_n \}$ de Q telle que $f^* = \lim f(x_n)$ si $n \to \infty$ ».

On rappelle alors :

1) qu'une condition nécessaire pour qu'un point x^* vérifie $f(x^*) = f^*$ est qu'il soit stationnaire, c'est-à-dire qu'il vérifie la relation :

$$\langle f'(x^*), x^* \rangle = \min (\langle f'(x^*), x \rangle \mid x \in Q)$$

2) si X est un espace de Banach reflexif et si f est fortement convexe alors il existe un et un seul point x^* qui réalise le minimum de f sur Q et vers lequel toute suite $\{ x_n \}$ minimisante converge.

II. PRESENTATION DES METHODES

Posons :

$$f_n(x) = \langle f'(x_n), x - x_n \rangle + \frac{1}{2} \langle T_{x_n} \cdot x - x_n, x - x_n \rangle \qquad (2.0)$$

Méthode 1 : A partir d'un point x_0 de Q on définit les suites $\{ x_n \}$, $\{ \bar{x}_n \}$ de Q de la façon suivante :

1) $x_n \to \bar{x}_n$ par la relation

$$f_n(\bar{x}_n) = \min (f_n(x) \mid x \in Q) \quad , \quad \bar{x}_n \in Q \qquad (2.1)$$

2) $x_n, \bar{x}_n \to x_{n+1}$ par la relation :

$$x_{n+1} = \bar{x}_n \qquad (2.2)$$

Méthode 2 : A partir d'un point x_0 de Q on définit les suites $\{ x_n \}$, $\{ \bar{x}_n \}$ de Q de la façon suivante :

1) $x_n \to \bar{x}_n$ par la relation 2.1.

2) soient

$$c_n = \langle f'(x_n), \bar{x}_n - x_n \rangle$$

$$\Delta(x_n, \bar{x}_n, \rho) = f(\bar{x}_n + \rho(\bar{x}_n - x_n)) - f(x_n)$$

Si $c_n = 0$ alors x_n est un point stationnaire, l'algorithme s'arrête.

Si $c_n \neq 0$ posons $g(x_n, \bar{x}_n, \rho) = \dfrac{\Delta(x_n, \bar{x}_n, \rho)}{\rho \cdot c_n}$

Alors $x_n, \bar{x}_n \to \rho_n$ par la relation :

si $g(x_n, \bar{x}_n, 1) \geqslant \sigma \rho_n = 1$ sinon $\rho_n \in [0, 1]$ et vérifie $\sigma \leqslant g(x_n, \bar{x}_n, \rho_n) \leqslant 1 - \sigma$ (2.3)

3) $x_n, \bar{x}_n, \rho_n \to x_{n+1}$ par la relation :

$$x_{n+1} = x_n + \rho_n(\bar{x}_n - x_n) \qquad (2.4)$$

Méthode 3 : A partir du point x_0 de Q on définit les suites $\{ x_n \}$, $\{ \bar{x}_n \}$ de Q de la façon suivante :

1) $x_n \to \bar{x}_n$ par la relation 2.1

2) $x_n, \bar{x}_n \to x_{n+1}$ par la relation :

$$f(x_{n+1}) = \min \left(f(x) \mid x \in [x_n, \bar{x}_n] \right) \quad , \quad x_{n+1} \in [x_n, \bar{x}_n] \tag{2.5}$$

On fait désormais les hypothèses A, B, C :

A : Pour tout x de Q il existe un point $\bar{x}$ de Q tel que :

$$\langle f'(x), \bar{x} - x \rangle + \frac{1}{2} \langle T_x \cdot \bar{x} - x, \bar{x} - x \rangle$$
$$= \min \left(\langle f'(x), u - x \rangle + \frac{1}{2} \langle T_x \cdot u - x, u - x \rangle \mid u \in Q \right) \tag{2.6}$$

B : Pour tout x de Q il existe un point $\bar{\bar{x}}$ tel que :

$$\langle f'(x), \bar{\bar{x}} - x \rangle = \min \left(\langle f'(x), u - x \rangle \mid u \in Q \right) \quad , \quad \bar{\bar{x}} \in Q \tag{2.7}$$

C : Il existe des scalaires $M, m (M \geqslant m \geqslant 0)$ tel que $\forall\, x, y \in Q$ on ait :

$$m \, \| y - x \|^2 \leqslant \langle T_x \cdot y - x, y - x \rangle \leqslant M \, \| y - x \|^2$$

REMARQUES

— si f est convexe et deux fois continuement dérivable dans Q et si $T_x = f''(x_o)$ ou si $T_x = f''(x)$ et f'' fortement bornée sur Q alors l'hypothèse C est vérifiée. Dans ce cas si f est fortement convexe alors m est > 0

— si X est un espace de Banach réflexif alors l'hypothèse B est vérifiée et si l'hypothèse C est vérifiée, alors l'hypothèse A l'est aussi.

Définition 1 : Une suite $\{ x_n \}$ de Q est dite stationnaire s'il existe une suite $\{ \bar{\bar{x}}_n \}$ associée à $\{ x_n \}$ par l'intermédiaire de 2.7 telle que si :

$$a_n = \langle f'(x_n), \bar{\bar{x}}_n - x_n \rangle. \tag{2.8}$$

alors :

$$\lim a_n = 0 \qquad \text{si} \qquad n \to \infty$$

Proposition 1 : Si f est convexe, toute suite $\{ x_n \}$ de Q qui est stationnaire est une suite minimisante.

Démonstration :

Soit $\{ y_m \}$ une suite de Q minimisante. Il existe une suite $\{ \varepsilon_m \}$ de scalaires positifs associée à la suite $\{ y_m \}$ telle que :

$$\lim \varepsilon_m = 0 \qquad \text{si} \qquad m \to \infty \ ;$$
$$f^* \geqslant f(y_m) - \varepsilon_m \ \forall\, m.$$

Si f est convexe alors :

$$0 \leqslant f(x_n) - f^* \leqslant f(x_n) - f(y_m) + \varepsilon_m$$
$$\leqslant \langle f'(x_n), x_n - y_m \rangle + \varepsilon_m \leqslant \langle f'(x_n), x_n - \bar{\bar{x}}_n \rangle + \varepsilon_m$$

Par conséquent :

$$\lim f(x_n) = f^*$$

Soit $\{ x_n \}$ une suite de Q, $\{ \bar{x}_n \}$ une suite associée à $\{ x_n \}$ par la relation 2.1 et c_n la suite de scalaire associée à $\{ x_n, \bar{x}_n \}$ par la relation :

$$c_n = \langle f'(x_n), \bar{x}_n - x_n \rangle \qquad (2.9)$$

Théorème 1 : Si pour une suite $\{ x_n \}$ il existe une suite $\{ \bar{x}_n \}$ définie par 2.1 telle que $\lim c_n = 0$ si $n \to \infty$ alors la suite $\{ x_n \}$ est une suite stationnaire.

Démonstration :

Puisque $\bar{x}_n$ réalise le minimum de f_n sur Q, $\bar{x}_n$ est un point stationnaire pour la fonction f_n et par conséquent on a la relation :

$$\langle f'_n(\bar{x}_n), \bar{x}_n - x_n \rangle \leqslant 0$$

c'est-à-dire :

$$\langle f'(x_n), \bar{x}_n - x_n \rangle + \langle T_{x_n} \cdot \bar{x}_n - x_n, \bar{x}_n - x_n \rangle \leqslant 0 \qquad (2.10)$$

Si l'on pose

$$\bar{b}_n = \langle T_{x_n} \cdot \bar{x}_n - x_n, \bar{x}_n - x_n \rangle$$

Alors la relation 2.10 entraîne :

$$\lim \bar{b}_n = 0 \qquad \text{si } n \to \infty$$

Soit $\{ \bar{\bar{x}}_n \}$ une suite quelconque associée à $\{ x_n \}$ par 2.8 et posons :

$$\bar{\bar{b}}_n = \langle T_{x_n} \cdot \bar{\bar{x}}_n - x_n, \bar{\bar{x}}_n - x_n \rangle$$

Si $\lim a_n \neq 0$ lorsque $n \to \infty$ alors il existe $\varepsilon > 0$, $j_0(\varepsilon)$ et une sous-suite $\{ n_j \}$ telle que si $j > j_0$ alors $-a_{n_j} \geqslant \varepsilon$

Posons

$$y_{n_j} = x_{n_j} + \lambda_{n_j}(\bar{\bar{x}}_{n_j} - x_{n_j}) \qquad \lambda_{n_j} \in [0, 1]$$

Puisque

$$c_{n_j} + \frac{1}{2} \bar{b}_{n_j} \leqslant \langle f'(x_{n_j}), y_{n_j} - x_{n_j} \rangle + \frac{1}{2} \langle T_{x_{n_j}} \cdot y_{n_j} - x_{n_j}, y_{n_j} - x_{n_j} \rangle$$

Alors :

$$c_{n_j} + \frac{1}{2} \bar{b}_{n_j} \leqslant \lambda_{n_j} a_{n_j} + \frac{\lambda_{n_j}^2}{2} \bar{\bar{b}}_{n_j}.$$

D'après l'hypothèse C si $a = \sup_j \bar{b}_{n_j}$ alors a est fini.

Par conséquent :

$$-\left(c_{n_j} + \frac{1}{2}\bar{b}_{n_j}\right) \geqslant \lambda_{n_j}\left(\varepsilon - \frac{\lambda_{n_j} \cdot a}{2}\right)$$

En choisissant $\lambda_n = \min\left(\frac{\varepsilon}{a}, 1\right)$ il existe alors $\varepsilon_1 > 0$ tel que :

$$-\left(c_{n_j} + \frac{1}{2}\bar{b}_{n_j}\right) \geqslant \varepsilon_1 > 0$$

En passant alors à la limite on obtient une contradiction. La suite $\{ x_n \}$ est donc stationnaire.

III. CONVERGENCE DES METHODES

1) Convergence des méthodes du type 1

On définit trois hypothèses D, E, F

— D : « $f''(x_0)$ existe ».

— E : « m est strictement positif ».

Dans ce cas on remarque que pour tout x_n il existe un et un seul point $\bar{x}_n$ défini par 2.1.

— Soient $A_{x_0\rho}$, $B_{x_0\rho}$, $C_{x_0\rho}$ les ensembles suivants :

$$A_{x_0\rho} = \{ \alpha' \in \bar{R}_+ : \|T_x - f''(x_0)\| \leqslant \alpha' \quad \forall\, x \in B(x_0, \rho) \}$$

$$B_{x_0\rho} = \{ \beta' \in \bar{R}_+ : \|f'(x) - f'(y) - f''(x_0) \cdot x - y\| \leqslant \beta' \, \|x - y\| \ \forall\, x, y \in B(x_0, \rho) \}$$

$$C_{x_0} = \{ x \in Q \quad : f_0(x) = \min\left(f_0(u) \mid u \in Q\right) \}$$

Soient $\alpha_{x_0\rho}$, $\beta_{x_0\rho}$ des éléments de $A_{x_0\rho}$ et $B_{x_0\rho}$ et η_{x_0} le scalaire défini par :

$$\eta_{x_0} = \sup\left(\|x - x_0\| \mid x \in C_{x_0}\right)$$

Posons alors $K_{x_0\rho} = \dfrac{2}{m}(\alpha_{x_0\rho} + \beta_{x_0\rho})$

$$r_{x_0\rho} = \eta_{x_0}/1 - K_{x_0\rho}$$

L'hypothèse F s'énonce alors :

« x_0 est tel qu'il existe $\rho_{x_0} > 0$, des scalaires $\alpha_{x_0\rho_{x_0}}$, $\beta_{x_0\rho_{x_0}}$ tels que :

$$K_{x_0\rho_{x_0}} < 1$$

$$r_{x_0\rho_{x_0}} \leqslant \rho_{x_0}$$

Théorème 2 : Si les hypothèses D, E, F sont vérifiées alors la suite $\{ x_n \}$ définie par les relations 2.1, 2.2 est stationnaire et converge fortement vers un point x^*. La rapidité de convergence est alors donnée par la formule :

$$\|x^* - x_n\| \leqslant K^n_{x_0\rho_{x_0}} \times r_{x_0\rho_{x_0}} \tag{3.2}$$

Démonstration :

a) Montrons que la suite $\{ x_n \}$ appartient à $B(x_0, \rho_{x_0})$. Raisonnons par récurrence :

Bien évidemment x_0, x_1 appartiennent à $B(x_0, \rho_{x_0})$. Supposons qu'il en soit ainsi jusqu'à l'ordre n et montrons que x_{n+1} appartient à $B(x_0, \rho_{x_0})$.

D'après 2.10 on a :

$$f_n(x_{n+1}) \leqslant -\frac{1}{2} < T_{x_n} \cdot x_{n+1} - x_n, \quad x_{n+1} - x_n > \leqslant -\frac{m}{2} \|x_n - x_{n+1}\|^2$$

Ce qui entraîne :

$$\|x_{n+1} - x_n\|^2 \leqslant -\frac{2}{m} f_n(x_{n+1})$$

D'autre part :

$$f_n(x_{n+1}) \geqslant \langle f'(x_n), x_{n+1} - x_n \rangle$$

comme :

$$f'(x_n) = f'(x_{n-1}) + T_{x_{n-1}} \cdot (x_n - x_{n-1})$$
$$+ (f''(x_0) - T_{x_{n-1}}) \cdot (x_n - x_{n-1}) + r_0$$

avec

$$\|r_0\| \leqslant \beta_{x_0\rho_{x_0}} \|x_n - x_{n-1}\|,$$

et comme x_{n-1} appartient à $B(x_0, \rho_{x_0})$ alors :

$$f'(x_n) = f'(x_{n-1}) + T_{x_{n-1}} \cdot (x_n - x_{n-1}) + r_1$$

avec

$$\|r_1\| \leqslant (\alpha_{x_0\rho_{x_0}} + \beta_{x_0\rho_{x_0}}) \|x_n - x_{n-1}\|$$

Puisque $x_{n+1} = \bar{x}_n$ et vérifie 2.1 on a alors :

$$-f_n(x_{n+1}) \leqslant -\langle f'(x_n), x_{n+1} - x_n \rangle$$
$$= -\langle f'(x_{n-1}) + T_{x_{n-1}} \cdot (x_n - x_{n-1}), x_{n+1} - x_n \rangle + \langle r_1, x_{n+1} - x_n \rangle$$

Et puisque : $f'_{n-1}(x_n) = f'(x_{n-1}) + T_{x_{n-1}} \cdot (x_n - x_{n-1})$ alors :

$$-f_n(x_{n+1}) \leqslant -\langle f'_{n-1}(x_n), x_{n+1} - x_n \rangle - \langle r_1, x_{n+1} - x_n \rangle$$

Comme :

$$\langle f'_{n-1}(x_n), x_{n+1} - x_n \rangle \geqslant 0$$

puisque d'après 2.1 : $x_n = \overline{x_{n-1}}$

On a alors :

$$-f_n(x_{n+1}) \leqslant -\langle \iota_1, x_{n+1} - x_n \rangle$$
$$\leqslant (\alpha_{x_0\rho_{x_0}} + \beta_{x_0\rho_{x_0}}) \|x_{n+1} - x_n\| \|x_n - x_{n-1}\|$$

Par conséquent :

$$\|x_{n+1} - x_n\| \leqslant 2\left(\frac{\alpha_{x_0\rho_{x_0}} + \beta_{x_0\rho_{x_0}}}{m}\right) \times \|x_n - x_{n-1}\| \leqslant K^n_{x_0\rho_{x_0}} \|x_1 - x_0\|$$

On a alors :

$$\|x_{n+1} - x_0\| \leqslant \sum_{v=0}^{v=n} \|x_{v+1} - x_v\| \leqslant \sum_{v=0}^{n} K^v_{x_0\rho_{x_0}} \eta_{x_0} \leqslant \frac{\eta_{x_0}}{1 - K_{x_0\rho_{x_0}}}$$

ce qui montre que :

$$x_{n+1} \in B(x_0, \rho_{x_0})$$

b) On vient de voir que la suite $\{ x_n \}$ est une suite de Cauchy. Comme Q est complet elle converge alors vers un point x^*. Comme :

$$\|x_{v+p} - x_v\| = \sum_{i=1}^{p} \|x_{v+i} - x_{v+i-1}\| \leqslant K^v_{x_0\rho_{x_0}} \times \sum_{i=0}^{p-1} K^i_{x_0\rho_{x_0}} \|x_1 - x_0\|$$
$$\leqslant K^v_{x_0\rho_{x_0}} \times \frac{\eta_{x_0}}{1 - K_{x_0\rho_{x_0}}}$$

On a :

$$\|x^* - x^v\| \leqslant K^v_{x_0\rho_{x_0}} \times r_{x_0\rho_{x_0}}$$

c) On a montré aussi que :

$$\|x_{n+1} - x_n\|^2 \leqslant -\frac{2}{m} f_n(x_{n+1})$$
$$\leqslant \frac{2}{m} (\alpha_{x_0\rho_{x_0}} + \beta_{x_0\rho_{x_0}}) \|x_{n+1} - x_n\| \cdot \|x_n - x_{n-1}\|$$

Ceci entraîne que : $\lim f_n(x_{n+1}) = 0$ si $n \to \infty$.

Le théorème 1 entraîne alors que la suite $\{ x_n \}$ est stationnaire.

Proposition 2 : Si

— X est un espace de Banach reflexif,

— f est fortement convexe sur Q et deux fois dérivable,

— au point x^* solution optimale de P $T_{x^*} = f''(x^*)$ et $m > 0$,

— S'il existe $\rho > 0$, $l > 0$ tel que quels que soient x, y appartenant à $B(x^*, \rho)$ on ait :

$$\|T_x - f''(y)\| \leqslant l \|x - y\|$$

Alors il existe ρ^* tel que pour tout x_0 de $B(x^*, \rho^*)$ les hypothèses D, E, F sont vérifiées.

Démonstration :

1) Les hypothèses D et E sont évidemment vérifiées.

2) Pour tout $x_0 \in B(x_0, \rho/2)$ on a si $\rho_1 < \rho/2$ pour tout $x, y \in B(x_0, \rho_1)$

$$\|T_x - f''(y)\| \leqslant 2l\rho_1$$

Par conséquent si l'on pose :

$$\rho_{x_0} \leqslant \min\left(\frac{m}{24l}, \rho/2\right)$$

$$\alpha_{x_0\rho_{x_0}} = 2l\rho_{x_0}; \qquad \beta_{x_0\rho_{x_0}} = 4l\rho_{x_0}$$

On constate que quel que soit x_0 appartenant à $B(x^*, \rho/2)$ on a :

$$K_{x_0\rho_{x_0}} \leqslant \frac{1}{2}$$

3) Montrons maintenant la proposition :

$\forall\, \varepsilon > 0 \quad \exists\, \eta(\varepsilon) \quad$ tel que $\quad \forall\, x_0 \in B(x^*, \rho) \quad$ vérifiant $\|x^* - x_0\| \leqslant \eta(\varepsilon)$

alors $$\|x^* - x_1\| \leqslant \varepsilon \tag{1}$$

Les topologies envisagées pour la démonstration de ce point sont des topologies fortes.

α) Posons

$$F_x(u) = \langle f'(x), u - x \rangle + \frac{1}{2} \langle T_x \cdot u - x, u - x \rangle$$

et soit $y(x)$ le point de Q qui réalise le minimum de F_x sur Q. Si l'on remarque que x^* réalise le minimum de F_{x^*} sur Q on constate alors que la proposition (1) sera démontrée si l'on prouve que la continuité au point x^* de l'application $x \rightarrow T_x$ de Q dans $L(X, X^*)$ entraîne la continuité au point x^* de l'application $x \rightarrow y(x)$ de Q dans Q.

β) Pour cela montrons d'abord au point x^* la continuité de la fonction numérique $x \rightarrow F_x(y(x))$ définie dans Q.

D'après les hypothèses de la proposition il existe une constante $K > 0$ telle que $\forall\, x \in B(x^*, \rho)\ \forall\, u \in Q$ on ait :

$$\|F_x(u) - F_{x^*}(u)\| \leqslant K\, \|x - x^*\|$$

Soit alors $\{ x_n \}$ une suite de points de $B(x^*, \rho)$ convergeant fortement vers x^*; on a alors la relation :

$$F_{x_n}(y(x_n)) \leqslant F_{x_n}(x^*)$$

et par conséquent :

$$\overline{\lim}\, F_{x_n}(y(x_n)) \leqslant \overline{\lim}\, F_{x_n}(x^*) = F_{x^*}(x^*)$$

D'autre part :

$$F_{x^*}(x^*) \leqslant F_{x^*}(y(x_n)) \Rightarrow F_{x^*}(x^*) \leqslant \underline{\lim}\, F_{x^*}(y(x_n)) = \underline{\lim}\, F_{x_n}(y(x_n))$$

ce qui entraîne :

$$F_{x^*}(x^*) = \lim F_{x_n}(y(x_n)) \quad \text{si} \quad n \to \infty \tag{2}$$

La relation (2) entraîne alors que :

$$F_{x^*}(x^*) = \lim F_{x^*}(y(x_n)) \quad \text{si} \quad n \to \infty$$

Comme F_{x^*} est fortement convexe on a :

$$F_{x^*}(y(x_n)) - F_{x^*}(x^*) \geqslant m \, \|y(x_n) - x^*\|^2$$

d'où la proposition (1).

4) Pour que x_0 soit tel que $r_{x_0 \rho_{x_0}} \leqslant \rho_{x_0}$ il suffit que x_0 appartienne à $B(x^*, \rho/2)$ et vérifie :

$$\eta_{x_0} \leqslant \frac{1}{2} \min\left(\frac{m}{24l}, \rho/2\right)$$

Posons :

$$\varepsilon = \frac{1}{4} \min\left(\frac{m}{24l}, \rho/2\right)$$

En choisissant $\rho^* = \min(\varepsilon, \eta(\varepsilon), \rho/2)$ [$\eta(\varepsilon)$ est définie par la proposition (1)] alors la proposition 2 est démontrée.

Remarque : Les hypothèses de la proposition 2 deviennent particulièrement claires si $T_x = f''(x)$.

2) Convergence des méthodes de type 2

Théorème 3 : Les suites $\{ x_n \}$, $\{ \bar{x}_n \}$, $\{ \rho_n \}$ définies par les relations 2.1, 2.3, 2.4 sont bien définies. La suite $\{ f(x_n) \}$ est décroissante.

Si f' est uniformément continue sur Q alors la suite $\{ x_n \}$ est stationnaire.

Démonstration :

1) si $c_n = 0$ alors d'après le théorème 1 x_n est un point stationnaire.

si $c_n \neq 0$ il existe $\xi \in [x_n, x_n + \rho(\bar{x}_n - x_n)]$ tel que :

$$g(x_n, \bar{x}_n, \rho) = 1 + \frac{\langle f'(\xi) - f'(x_n), \bar{x}_n - x_n \rangle}{c_n} \tag{3.2}$$

si $\rho = 0$, $g(x_n, \bar{x}_n, 0) = 1$; si $g(x_n, \bar{x}_n, 1) \leqslant \sigma$ alors la continuité de f entraîne la continuité de l'application $\rho \to g(x_n, \bar{x}_n, \rho)$ et par conséquent il existe ρ_n tel que :

$$\sigma \leqslant g(x_n, \bar{x}_n, \rho_n) \leqslant 1 - \sigma \quad , \quad \rho_n \in [0, 1]$$

Puisque $-c_n$, $g(x_n, \bar{x}_n, \rho_n)$ sont positifs, $\Delta(x_n, \bar{x}_n, \rho_n)$ est négatif et par conséquent la suite $\{f(x_n)\}$ est décroissante.

2) Pour montrer que la suite $\{x_n\}$ est stationnaire il suffit de montrer d'après le théorème 1 que :

$$\lim c_n = 0 \qquad \text{si} \qquad n \to \infty$$

S'il n'en était pas ainsi il existerait une sous-suite $\{n_j\}$ telle que :

$$\lim c_{n_j} \neq 0 \qquad \text{si} \qquad j \to \infty$$

Dans ce cas la sous-suite $\{\rho_{n_j}\}$ associée converge vers 0.

En effet :

$$-(f(x_{n_j+1}) - f(x_{n_j})) \geqslant -\sigma c_{n_j} \rho_{n_j}$$

Comme il existe $\varepsilon_1 > 0$ $j_0(\varepsilon)$ tel que si $j > j_0(\varepsilon) - c_{n_j} > \varepsilon_1$ et comme la suite $\{f(x_n)\}$ est décroissante et bornée inférieurement on voit que

$$f(x_{n_j}) - f(x_{n_j+1}) \geqslant \sigma \varepsilon_1 \rho_{n_j} \Rightarrow \lim \rho_{n_j} = 0 \qquad \text{si} \qquad j \to \infty$$

Soit ξ_n la suite associée à $\{x_n, \bar{x}_n, \rho_n\}$ par 3.2.

Posons

$$b_n = |\langle f'(\xi_n) - f'(x_n), \bar{x}_n - x_n \rangle|$$

Puisque f' est uniformément continue alors $\lim |\rho_{n_j}| = 0$ si $j \to \infty$ entraîne puisque Q est borné :

$$\lim b_{n_j} = 0 \qquad \text{si} \qquad j \to \infty$$

D'autre part :

$$1 - \sigma \geqslant 1 + b_{n_j}/c_{n_j} \geqslant 1 - b_{n_j}/\varepsilon_1$$

Ce qui est en contradiction avec le fait que $0 < \sigma < 1/2$ et que

$$\lim b_{n_j} = 0 \text{ si } j \to \infty.$$

Théorème 4 : Reprenons les hypothèses du théorème 3 et supposons de plus :

— X est un espace de Hilbert.

— f est fortement convexe dans Q et deux fois continuement dérivable.

— $T_x = U_x + t^2 J$ où :

U_x est un opérateur symétrique de $L(X, X^*)$ tel que :

— $\langle U_x \cdot y - x, y - x \rangle \geqslant m_1 \|y - x\|$, $m_1 > 0$,

— si x^* est l'unique point qui réalise le minimum de f sur Q alors $U_{x^*} = f''(x^*)$.

— Il existe des scalaires strictement positifs ρ, l tels que $\forall\, x, y \in B(x^*, \rho)$ on ait :

$$\|U_x - f''(y)\| \leqslant l\, \|x - y\|$$

Alors il existe des scalaires strictement positifs t_0, γ, $d(d < 1)$ des entiers N_0, $N(N > N_0)$ tels que si ρ_n, x_n sont définis par les relations 2.1, 2.3, 2.4 et

1) si $|t| < t_0$, $n > N_0$ alors $\rho_n = 1$;

2) si de plus $n > N$ alors :

$$\|x^* - x_n\| \leqslant \gamma d^j \qquad j = N - n$$

Démonstration :

1) Montrons qu'il existe N_0, t'_0 tel que si $|t| \langle t'_0, n \rangle N_0$ alors $\rho_n = 1$.

D'après le théorème 3 on a :

$$\lim c_n = 0 \qquad \text{si} \qquad n \to \infty$$

Ceci entraîne :

$$\lim f_n(\bar{x}_n) = 0 \qquad \text{si} \qquad n \to \infty$$

D'autre part comme dans le cas *a*) de la démonstration du théorème 2 on a :

$$\|\bar{x}_n - x_n\| \leqslant \frac{-2f_n(\bar{x}_n)}{m_1}$$

et par conséquent :

$$\lim \|\bar{x}_n - x_n\| = 0 \qquad \text{si} \qquad n \to \infty$$

D'autre part :

$$g(x_n, \bar{x}_n, 1) = 1 + \frac{\langle T_{x_n} \cdot \bar{x}_n - x_n, \bar{x}_n - x_n \rangle}{2c_n} + \frac{\langle (f''(\xi_n) - T_{x_n}) \cdot \bar{x}_n - x_n, \bar{x}_n - x_n \rangle}{2c_n}$$

ξ_n appartenant à $[x_n, \bar{x}_n]$.

D'après la relation 2.10 si $\bar{b}_n = \langle T_{x_n} \cdot \bar{x}_n - x_n, \bar{x}_n - x_n \rangle$ alors

$$\bar{b}_n \leqslant - c_n$$

D'autre part :

$$\bar{b}_n \geqslant m_1 \|\bar{x}_n - x_n\|^2$$

Par conséquent :

$$\|g(x_n, \bar{x}_n, 1) - 1\| \leqslant \frac{1}{2} + \frac{1}{2m_1}\|f''(\xi_n) - T_{x_n}\|$$

Puisque f est fortement convexe, la suite $\{ x_n \}$ qui est minimisante converge fortement vers x^* et comme $\lim \|x_n - \bar{x}_n\| = 0$ si $n \to \infty$ il existe donc N'_0 tel que pour $n > N'_0$ on ait :

$$x_n, \bar{x}_n \in B(x^*, \rho)$$

Par suite :

$$\|g(x_n, x_n, 1) - 1\| \leqslant \frac{1}{2} + \frac{l}{2m_1}(\|\xi_n - x_n\|) + \frac{t^2}{2m_1}$$

Il existe donc t'_0 tel que $\forall\ |t|\ \langle t'_0$ on ait :

$$\underline{\lim}\ g(x_n, \bar{x}_n, 1) > \sigma$$

Par conséquent il existe bien N_0, t'_0 tel que si $|t|\ \langle t'_0, n \rangle\ N_0$ alors $\rho_n = 1$.

2) En reprenant les mêmes arguments que ceux utilisés dans la démonstration de la proposition 2 on montre qu'il existe t''_0, ρ^* tel que

$$\forall\ x_0 \in B(x^*, \rho^*),\ |t| \leqslant t''_0$$

il existe ρ_{x_0} telles que les hypothèses D, E, F soient vérifiées.

Soit N l'entier positif supérieur à N_0 tel que $\forall\ n > N$ on ait :

$$x_n \in B(x^*, \rho^*).$$

Alors pour tout $|t| \leqslant t_0 = \inf(t'_0, t''_0), n > N$ la méthode 2 coïncide avec la méthode 1 et si l'on pose $x_0 = x_N$ les hypothèses D, E, F du théorème 2 sont alors vérifiées. D'où le théorème 4.

3) Convergence des méthodes de type 3

Théorème 5 : On suppose qu'il existe $\alpha > 0$, $K > 0$ tel que

$$\forall\ x, y \in Q, \qquad \forall\ \lambda \in [0, 1] \qquad \text{on ait la relation :}$$

$$|f(x + \lambda(y - x)) - f(x) - \lambda \langle f'(x), y - x \rangle| \leqslant K\lambda^{1+\alpha},$$

Alors les suites $\{ x_n \}$, $\{ \bar{x}_n \}$ définies par les relations 2.1, 2.5 sont bien définies. La suite $\{ f(x_n) \}$ est décroissante et la suite $\{ x_n \}$ est stationnaire.

Démonstration :

a) f étant continue, le point x_{n+1} est bien défini à partir de x_n, $\bar{x}_n$.

D'autre part la suite $\{ f(x_n) \}$ est décroissante par construction.

b) Montrons que $\lim c_n = 0$ si $n \to \infty$. Le théorème 1 montre alors que la suite $\{ x_n \}$ est une suite stationnaire.

S'il n'en n'était pas ainsi il existerait alors $\varepsilon > 0$ j_0 et une sous-suite $\{ n_j \}$ telle que pour $j > j_0$ on ait :

$$-c_{n_j} > \varepsilon.$$

La relation 2.5 entraîne alors :

$$f(x_{n_j}) - f(x_{n_j} + \lambda(\bar{x}_{n_j} - x_{n_j})) \leqslant f(x_{n_j}) - f(x_{n_j+1})\ \forall\ \lambda \in [0, 1]$$

Comme $\{ f(x_n) \}$ est une suite décroissante on a :

$$\underline{\lim}\ \{ f(x_{n_j}) - f(x_{n_j} + \lambda(\bar{x}_{n_j} - x_{n_j})) \} \leqslant 0$$

c'est-à-dire

$$\underline{\lim} \{ \lambda \langle f'(\bar{x}_{n_j}), x_{n_j} - x_{n_j} \rangle + \omega(\lambda, \bar{x}_{n_j}, x_{n_j}) \} \geqslant 0$$

avec

$$|\omega(\lambda, x_n, \bar{x}_n)| \leqslant K\lambda^{1+\alpha}, \alpha > 0$$

Par conséquent

$$\underline{\lim} \langle f'(x_{n_j}), \bar{x}_{n_j} - x_{n_j} \rangle \geqslant 0$$

d'où la contradiction.

IV. APPLICATIONS A LA THEORIE DU CONTROLE

Soient :

— $X = L_2^m[0, T]$ l'espace des classes d'équivalence des fonctions mesurables à valeurs dans R^m et dont les composantes sont de carré intégrables sur $[0, T]$ pour la mesure de Lebesgue, espace muni de la norme Hilbertienne habituelle.

— A, B, h des matrices (n, n), (n, m), $(n, 1)$ dont les termes sont des fonctions mesurables bornées sur $[0, T]$,

— Φ une fonction numérique définie sur R^n et trois fois continûment différentiable,

— F une fonction numérique définie sur $[0, T] \times R^n$ trois fois continûment différentiable,

— R une application mesurable est bornée de $[0, T]$ dans l'ensemble des matrices (m, m),

— M un sous-ensemble convexe compact de R^m,

— I la matrice unité de R^m.

Soient alors :

— $Q = \{ u \in X : u(t) \in M \ pp \text{ sur } [0, T] \}$

— $B(u, \rho) = \{ v \in Q : \|u - v\| \leqslant \rho \}$

— f la fonction numérique définie par

$$f(u) = \Phi(x_u(T)) + \int_0^T F(t, x_u(t)) + u(t) \cdot R(t) \cdot u(t) \, dt \qquad (4.1)$$

où x_u correspondant à u par 4.2

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + h(t) \qquad (4.2)$$

$$x(0) = x_0$$

On recherche un point $u^* \in Q$ tel que $f(u^*) = (\min f(u) \mid u \in Q)$. On constate alors :

1) f admet une dérivée d'ordre trois f''' et f' est uniformément continue sur Q.

2) si $T_u = 0$ ou si $T_u = J$ alors les hypothèses A, B, C et celles du théorème 5 sont vérifiées.

3)

— si $\forall\, t \in [0, T]$ l'application $x \to F(t, x)$ est convexe,

— si Φ est convexe,

— s'il existe $m \geqslant 0$ [resp $m > 0$] tel que $\forall\, s \in [0, T]$, $\forall\, u \in R^m$ on ait :

$$u^t R(s) u \geqslant m \, \|u\|^2$$

Alors f est convexe [resp fortement convexe].

a) Si f est convexe et si $T_u = f''(u) + t^2 J$ ou si $T_u = f''(u_0)$ alors les hypothèses A, B, C et les hypothèses des théorèmes 3 et 5 sont vérifiées.

b) Si f est fortement convexe alors

α) si $T_u = f''(u) + t^2 J$ il existe t_0 tel que si $|\, t \,| < t_0$ les hypothèses du théorème 4 sont vérifiées.

De plus il existe ρ^* tel que si $u_0 \in B(u^*, \rho^*)$ alors les hypothèses du théorème 2 sont vérifiées.

β) Si $T_u = f''(u_0)$ $u^* \neq u$, $T_{u^*} = f''(u^*)$ alors il existe ρ^* tel que si $u_0 \in B(u^*, \rho^*)$ les hypothèses du théorème 2 sont vérifiées.

D'autre part si l'on pose :

$$h_{1n}(t) = \Phi_x(x_{u_n}(t)) \quad ; \quad h_{2n}(t) = \Phi_{xx}(x_{u_n}(t)) \quad ; \quad h_{3n}(t) = F_x(t, x_{u_n}(t), u_n(t)),$$

$$h_{4n}(t) = 2[R(t) + I\alpha^2] \cdot u_n(t) \quad ; \quad h_{5n}(t) = F_{xx}(t, x_{u_n}(t), u_n(t))$$

$$h_{6n}(t) = 2(R(t) + \alpha^2 I)$$

alors si $T_u = f''(u) + \alpha^2 J$ on a :

$$\begin{aligned} {}_n(u) = h_{1n}(T) \cdot (x_u(T) - x_{u_n}(T)) + \frac{1}{2} (x_u(T) - x_{u_n}(T))^t \cdot h_{2n}(T) \cdot (x_u(T) - x_{u_n}(T)) \\ + \int_0^T h_{3n}(s) \cdot (x_u(s) - x_{u_n}(s)) + h_{4n}(s) \cdot (u(s) - u_n(s)) \, ds \\ + \frac{1}{2} \int_0^T (x_u(s) - x_{u_n}(s))^t \cdot h_{5n}(s) \cdot (x_u(s) - x_{u_n}(s)) + (u(s) - u_n(s))^t \\ \cdot h_{6n}(s) \cdot (u(s) - u_n(s)) \, ds. \end{aligned}$$

Le point $\bar{u}_n$ qui réalise le minimum de f_n sur Q est alors obtenu par des algorithmes de programmation quadratique du type [1].

BIBLIOGRAPHIE

(1) AUSLENDER-BRODEAU, Convergence d'un algorithme de Frank et Wolfe appliqué à un problème de contrôle, *R.I.R.O.*, n° 1, 1968.

(2) FRANK et WOLFE, An algorithm for quadratic progamming, *Naval Research Logistic Quaterly*.

(3) GOLDSTEIN, Minimizing functionnals on normed linear spaces, *Siam Control*, vol. 4, n° 1, 1966.

(4) KANTOROVITCH-AKILOV, Functionnal Analysis in normed spaces, New York Mac Millan, 1964.

(5) MARQUART, An algorithm for least spaces estimations of non linear parameters, *J. Soc. Indust. App. Math.*, 11, 1963, 431-441.

(6) PERAYA, Iterative methods for solving non linear least squares, *Problems Siam J. Numerical Anal.*, 4, 1967, 27-36.

7) POLJAK-LEVITIN, Constrained minimizations problems, *Zh. vychisl Mat mat Fiz*, 6.5, 787-823, 1966.

(8) POLJAK-LEVITIN, BT Dok Akad Nauk USSR, 168, 5, 997-1000, 1966.