ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIV, 3 (1974)

S. GRZEGÓRSKI and H. POPŁAWSKA (Lublin)

ON A METHOD OF OPTIMIZATION

Morrison [1] has developed a method for minimizing a non-linear function under non-linear constraints in the form of equalities by solving a sequence of problems of unconstrained minimization. Here we consider a generalization of this method to the case of minimizing a functional on a Banach space under constraints of mixed type, i.e. in the form of a system of equalities and inequalities.

1. Let f be a real functional on a Banach space B_1.

PROBLEM 1 (the constrained minimum). We choose an element $\bar{x} \in B_1$ such that the functional f reaches the minimum at this point under the constraints $g_i(x) \geqslant 0$, $i = 1, 2, \ldots, n$, and $A(x) = 0$, where g_i are functionals on B_1 and A is the operator from B_1 into a Banach space B_2.

We formulate the unconstrained minimization problem which we will use to solve Problem 1.

PROBLEM 2 (the unconstrained minimum). Let

$$F(x, L, M) = \big(f(x) - M\big)^2 + P(x, L),$$

where

$$P(x, L) = h[A(x)] + \sum_{i=1}^{n} \lambda_i \big(g_i(x) - l_i^2\big)^2,$$

h is a functional defined on a Banach space B_2 satisfying the conditions

$$h[0] = 0, \qquad \bigwedge_{y_1, y_2 \in B_2} \big(\|y_1\| < \|y_2\| \Rightarrow h[y_1] < h[y_2]\big),$$

L denotes the vector $(l_1, l_2, \ldots, l_n)$, $l_i \in R$, and M and λ_i are given constants, $\lambda_i > 0$, $i = 1, 2, \ldots, n$. We choose a point $x^* \in B_1$ and a vector L^* such that

$$F(x^*, L^*, M) = \min_{x \in B_1, L \in R^n} F(x, L, M).$$

Let us write

(i) $$s(M) = f(x^*).$$

2. Let M be an optimistic estimate of $f(\bar{x})$, i.e. suppose that $M \leqslant f(\bar{x})$.

THEOREM 1. *If (x^*, L^*) is a solution of Problem 2, and $\bar{x}$ is a solution of Problem 1, then $f(x^*) \leqslant f(\bar{x})$.*

Proof. Let $\bar{L} = (\bar{l}_1, \bar{l}_2, \ldots, \bar{l}_n)$, $\bar{l}_i = \sqrt{g_i(\bar{x})}$. Since (x^*, L^*) is a solution of Problem 2, we have

$$F(\bar{x}, \bar{L}, M) \geqslant F(x^*, L^*, M)$$

and

$$\big(f(\bar{x})-M\big)^2+P(\bar{x}, \bar{L}) \geqslant \big(f(x^*)-M\big)^2+P(x^*, L^*).$$

Since $P(\bar{x}, \bar{L}) = 0$ and $P(x^*, L^*) \geqslant 0$, therefore

$$\big(f(\bar{x})-M\big)^2 \geqslant \big(f(x^*)-M\big)^2+P(x^*, L^*) \geqslant \big(f(x^*)-M\big)^2.$$

Using the condition $f(\bar{x})-M \geqslant 0$, we have $|f(x^*)-M| \leqslant f(\bar{x})-M$, whence $f(x^*) \leqslant f(\bar{x})$, q.e.d.

THEOREM 2. *If $M = f(\bar{x})$ and $\bar{L}$ is defined as in the proof of Theorem 1, then $(\bar{x}, \bar{L})$ is a solution of Problem 2 and x^* is a solution of Problem 1.*

Proof. If $M = f(\bar{x})$, then, for any $x \in B_1$ and $L \in R^n$,

$$F(\bar{x}, \bar{L}, M) = \big(f(\bar{x})-M\big)^2+P(\bar{x}, \bar{L}) = 0$$
$$\leqslant \big(f(x)-M\big)^2+P(x, L) = F(x, L, M).$$

This shows that $(\bar{x}, \bar{L})$ is a solution of Problem 2. On the other hand, by the definition of (x^*, L^*),

$$0 \leqslant F(x^*, L^*, M) \leqslant F(\bar{x}, \bar{L}, M) = 0.$$

Hence $f(x^*) = M$ and $P(x^*, L^*) = 0$, and from this

$$A(x^*) = 0 \quad \text{and} \quad g_i(x^*) = l_i^{*2} \geqslant 0$$

and

$$f(x^*) = M = f(\bar{x}) \leqslant f(x) \quad \text{for any } x \in B_1.$$

This shows that x^* is a solution of Problem 1, q.e.d.

THEOREM 3. *The function $s(M)$ defined by* (i) *is monotonically non-decreasing on the set R.*

Proof. Let $M_1 < M_2$ and (x_i^*, L_i^*) with $M = M_i$, $i = 1, 2$, be the solution of Problem 2. Then

$$F(x_1^*, L_1^*, M_1) \leqslant F(x, L, M_1) \quad \text{and} \quad F(x_2^*, L_2^*, M_2) \leqslant F(x, L, M_2).$$

Assuming $x = x_2^*$ and $L = L_2^*$ in the right-hand side of the first inequality, and $x = x_1^*$ and $L = L_1^*$ in the second one and adding both these, we have

$$-2f(x_1^*)M_1-2f(x_2^*)M_2 \leqslant -2f(x_2^*)M_1-2f(x_1^*)M_2,$$

whence

$$f(x_1^*)(M_2-M_1) \leqslant f(x_2^*)(M_2-M_1).$$

Since $M_1 < M_2$, we get $s(M_1) = f(x_1^*) \leqslant f(x_2^*) = s(M_2)$, q.e.d.

3. We give an algorithm which allows to obtain the solution of Problem 1 as the limit of a sequence of solutions of Problem 2. To do this we construct a sequence of numbers $\{M_r\}_{r=1,2,\ldots}$. Let (x_r^*, L_r^*) denote the solution of Problem 2 for $M = M_r$.

ALGORITHM. We take an arbitrary $M_1 \leqslant f(\bar{x})$. Let $r = 1, 2, \ldots$

1. We solve Problem 2 for $M = M_r$, i.e. we choose (x_r^*, L_r^*) such that

$$F(x_r^*, L_r^*, M_r) = \min_{x\in B_1, L\in R^n} F(x, L, M_r).$$

2. We evaluate $M_{r+1} = M_r + \sqrt{F(x_r^*, L_r^*, M_r)}$.

3. If $M_{r+1} = M_r$, we finite this process, otherwise we go back to 1.

THEOREM 4. *If* $M_r \leqslant f(\bar{x})$, *then* $M_{r+1} \leqslant f(\bar{x})$.

Proof. We have $F(x_r^*, L_r^*, M_r) \leqslant F(x, L, M_r)$. In particular, when $x = \bar{x}$ and $L = \bar{L}$, $F(x_r^*, L_r^*, M_r) \leqslant F(\bar{x}, \bar{L}, M_r)$. Hence

$$F(x_r^*, L_r^*, M_r) \leqslant \big(f(\bar{x})-M_r\big)^2.$$

Since $M_r \leqslant f(\bar{x})$, this implies that $\sqrt{F(x_r^*, L_r^*, M_r)} \leqslant f(\bar{x})-M_r$, whence

$$M_r + \sqrt{F(x_r^*, L_r^*, M_r)} \leqslant f(\bar{x}),$$

and thus $M_{r+1} \leqslant f(\bar{x})$, q.e.d.

Assume that d is a non-negative number. Let

$$T_d = \big\{x\colon \bigvee_{L\in R^n} P(x, L) = d\big\},$$

let $\bar{x}_d$ be the solution of the problem of minimizing $f(x)$ on the set T_d, and

$$m(d) = \min_{x\in T_d} f(x).$$

We prove the following

LEMMA. $m(0) = f(\bar{x}_0) = f(\bar{x})$.

In fact,

$$x \in T_0 = \big\{x\colon \bigvee_{L\in R^n} P(x, L) = 0\big\}$$

if and only if $h[A(x)] = 0$ and there exist $l_1, l_2, \ldots, l_n$ such that $g_i(x) = l_i^2$. Evidently, $\bar{x}$ satisfies these conditions. Since $\bar{x}$ is a minimum of $f(x)$ on B_1 under the weaker conditions $g_i(x) \geqslant 0$, we have

$$f(\bar{x}) = \min_{x\in T_0} f(x) = m(0), \qquad \text{q.e.d.}$$

Let us write $d_k = P(x_k^*, L_k^*)$.

THEOREM 5. *If $m(d)$ is a continuous function of d for $d = 0$, then the sequences $\{M_k\}_{k=1,2...}$ and $\{f(x_k^*)\}_{k=1,2...}$ converge to $f(\bar{x})$.*

Proof. Let $\varepsilon > 0$. Since $m(d)$ is continuous, there exists a $\delta > 0$ such that

$$|m(d) - m(0)| < \varepsilon \quad \text{for } |d| < \delta.$$

Of course, $\{M_k\}_{k=1,2...}$ is convergent, since it is non-decreasing and bounded from above (Theorem 4). Hence there exists a k_0 such that

$$M_{k+1} - M_k < \min(\sqrt{\varepsilon}, \sqrt{\delta}) \quad \text{for } k > k_0,$$

i.e.

$$F(x_k^*, L_k^*, M_k) < \min(\varepsilon, \delta).$$

In other words, $F(x_k^*, L_k^*, M_k) \to 0$ as $k \to \infty$. Thus

$$|f(x_k^*) - M_k| \to 0 \quad \text{and} \quad P(x_k^*, L_k^*) \to 0 \quad \text{as } k \to \infty.$$

In particular, for $k > k_0$,

$$d_k = P(x_k^*, L_k^*) < \min(\varepsilon, \delta) \leqslant \delta,$$

which, in view of the Lemma, implies

$$|f(\bar{x}_{d_k}) - f(\bar{x})| = |m(d_k) - m(0)| < \varepsilon$$

and, finally,

$$f(\bar{x}) - \varepsilon < f(\bar{x}_{d_k}).$$

Hence $x_k^* \in T_{d_k}$, and

$$f(\bar{x}_{d_k}) = \min_{x \in T_{d_k}} f(x),$$

therefore, $f(\bar{x}_{d_k}) \leqslant f(x_k^*)$ and $f(\bar{x}) - \varepsilon < f(x_k^*)$. Thus $f(\bar{x}) - f(x_k^*) < \varepsilon$. From Theorem 1 we have $f(x_k^*) \leqslant f(\bar{x})$. Thus

$$0 \leqslant f(\bar{x}) - f(x_k^*) < \varepsilon \quad \text{for } k > k_0.$$

That is, $f(x_k^*) \to f(\bar{x})$ as $k \to \infty$. Since $|f(x_k^*) - M_k| \to 0$ as $k \to \infty$, also $M_k \to f(\bar{x})$ as $k \to \infty$, q.e.d.

In practice, for sufficiently large k, we obtain only a certain approximation of $f(\bar{x})$. Thus the point x_k^* will satisfy the conditions of Problem 1 with some tolerance. To satisfy these conditions with a better accuracy one can try to change (to increase) the values λ_i during the iterative process.

Reference

[1] D. D. Morrison, *Optimization by least squares*, SIAM J. Numer. Anal. 5 (1968), p. 83-88.

DEPARTMENT OF NUMERICAL METHODS
M. CURIE-SKŁODOWSKA UNIVERSITY
20-031 LUBLIN

Received on 12. 5. 1973

S. GRZEGÓRSKI i H. POPŁAWSKA (Lublin)

O PEWNEJ METODZIE OPTYMALIZACYJNEJ

STRESZCZENIE

W pracy [1] podana jest metoda warunkowego poszukiwania minimum funkcji nieliniowej wielu zmiennych z ograniczeniami równościowymi. Pozwala ona uzyskać rozwiązanie zadania na minimum warunkowe, przechodząc do granicy w ciągu rozwiązań zadań na minimum bezwarunkowe.

Podane tu uogólnienie polega na skonstruowaniu analogicznego ciągu w przypadku minimalizacji funkcjonałów, określonych w przestrzeniach Banacha, z ograniczeniami typu mieszanego (równościowego i nierównościowego).
