Control
and Cybernetics

VOL. 7 (1978) No. 3

Method of functions basis embeding in optimization problem

by

STANISŁAW M. GRZEGÓRSKI

Numerical Analysis Department
Institute of Mathematics
Marie Curie-Skłodowska University, Lublin

In the paper an iterative method of minimizing a functional without constraints is presented. The method consists in replacing the problem of minimization of a functional defined on an infinitely dimensional space by an equivalent sequence of minimizations of functions defined on a m-dimensional space, m being fixed. The convergence of the method is discussed.

1. Introduction

Let X be a Banach space and let a linearly independent set of elements $\{e_1, e_2, ...\}$ of X be given, such that the linear spanned by $\{e_1, e_2, ...\}$ is dense in X. In addition, let an integer $m \geq 1$, a set $\Omega \subset R^m$ and a function $\varphi: \Omega \to X$ be given, such that

$$e_i \in \varphi(\Omega) \quad \text{for all} \quad i=1, 2, \dots . \tag{1}$$

Usually Ω is either R^m or a convex cone in R^m.

In this paper we will present a method of unconstrained minimization of a functional f defined on X.

In Sec. 2 we will define this method iteratively

$$x_{i+1}=x_i-\alpha_i \varphi(s_i),\ i=0, 1, 2, \dots \tag{2}$$

where $\alpha_i \in R$, $s_i \in \Omega$ are calculated by the minimization of functions

$$h_i(\alpha, s) \overset{\mathrm{df}}{=} f(x_i-\alpha \varphi(s)),\ \alpha \in R,\ s \in \Omega . \tag{3}$$

In the Ritz's method one has to increase the dimension of auxiliary problem, whereas in the method presented here the minimization still proceeds on the set $R \times \Omega \subset R^{m+1}$, m being fixed.

In Sec. 3 we will give a theorem on the weak convergence of the sequence $f'(x_i)$ as $i \to \infty$, where $f'(x)$ denotes the Fréchet-differential at x. In Sec. 4 we will present an example of sufficient conditions for the convergence of the sequence x_i. Some examples of a natural embeding of the set $\{e_1, e_2, ...\}$ in a finite-dimensional manifold will be given in Sec. 5. The material presented is taken from the doctoral dissertation ([1] Pt. II) written under the supervision of Dr. Ś. Ząbek.

2. Algorithm

Let X be a Banach space and let a linearly independent set of elements $\{e_1, e_2, ...\}$ of X be given, such that the lineal spanned by $\{e_1, e_2, ...\}$ is dense in X.

Let there be given a point $x_0 \in X$ and a functional $f: X \to R$ and let

$$\inf_{x \in X} f(x) = d > -\infty . \tag{4}$$

Let us consider the following:

Problem A. Find a point $\tilde{x} \in X$ such that

$$f(\tilde{x}) \leqslant d(1+\varepsilon) \tag{5}$$

where ε is a fixed non-negative number.

As a particularly case we may consider the problem:

find such $\bar{x} \in X$ if it exists, that $f(\bar{x}) = d$.

In addition, let there be given, an integer $m \geqslant 1$, a set $\Omega \subset R^m$ and a function $\varphi: \Omega \to X$, satisfying (1).

Now we can defined the following sets:

$$W_0 = [x \in X: f(x) \leqslant f(x_0)] \tag{6}$$

and

$$Q = [x \in X: x = \pm \varphi(s)/\|\varphi(s)\|, s \in \Omega, \varphi(s) \neq 0] . \tag{7}$$

We assume that

$$W_0 \text{ is bounded}, \tag{8}$$

$$\text{the functional } f \text{ is Fréchet-differentiable on } X, \tag{9}$$

$$f'(x) \text{ satisfies the Lipschitz's condition on } \operatorname{conv}(W_0) . \tag{10}$$

First, for a given f and $x_0 \in X$ we formulate a method which generates a sequence $x_k \in X$, $k=1, 2, ...$, such that

$$f'(x_k) \xrightarrow[k \to \infty]{} 0 \text{ (weakly*)} .$$

We fix $0 < \varepsilon < 1$.

Algorithm. For a given x_k, $k=0, 1, 2, ...$ we choose $y_k \in Q$ satisfying the inequality

$$f'(x_k)\, y_k \geqslant (1-\varepsilon) \sup_{y \in Q} f'(x_k)\, y \tag{11}$$

and define $r_k \in R$, such that

$$r_k = \min \{r > 0 : f'(x_k - r y_k)\, y_k = 0\}. \tag{12}$$

Then we select $t_k \in R$ and $z_k \in Q$ satisfying the relation

$$f(x_k - t_k z_k) \leqslant f(x_k - r_k y_k) \tag{13}$$

for example, $t_k = r_k$, $z_k = y_k$, and we take

$$x_{k+1} = x_k - t_k z_k. \tag{14}$$

The introduction of ε into (11) assures the existence of y_k in each iteration and also quarantees the convergence of the numerical realization of the method if ε denotes a relative error generated in the calculation of $f'(x_k)\, y_k$.

If $f'(x_k) \neq 0$, then it follows from (8) and (10) that there exists r_k defined by (10). On the other hand, if $f'(x_k) = 0$, then the necessary condition for extremum is satisfied.

It is easy to verify that the method defined by (2), (3) is a particular case of the algorithm defined by (11)—(14).

3. Weak convergence of the algorithm

In this section we will give a theorem on the weak* convergence of the sequence $f'(x_k)$.

THEOREM 1. Let the functional f satisfy the conditions (4), (8)—(10), and the sequence x_k be defined by (11)—(14), then

$$f'(x_k) \to 0 \text{ (weakly*) as } k \to \infty.$$

Proof. First we shall demonstrate that

$$f'(x_k)\, y_k \to 0 \quad \text{as} \quad k \to \infty.$$

The contradiction of this result will be proved here. We shall assume that there exist $\{k_i\}$, such that

$$f'(x_{k_i})\, y_{k_i} \geqslant \delta > 0 \quad \text{(eqs. (8), (9))}.$$

From (8) and (10) it follows that there exists a sequence $q_i \in (0, r_{k_i})$, i=0, 1, 2, ..., such that

$$f'(x_{k_i} - q_i y_{k_i})\, y_{k_i} = \delta/2.$$

Observe that

$$f(x_0) - d \geqslant \sum_{i=0}^{l} [f x_{k_i}) - f(x_{k_i+1})] \geqslant \sum_{i=0}^{l} [f(x_{k_i}) - f(x_{k_i} - q_i y_{k_i})] =$$

$$= \sum_{i=0}^{l} f'(x_{k_i} - \Theta_i q_i y_{k_i})\, q_i y_{k_i} \geqslant \frac{\delta}{2} \sum_{i=0}^{l} q_i, \qquad \Theta_i \in (0, 1).$$

Hence $q_i \to 0$ for $i \to \infty$ and also

$$f'(x_{k_i})\, y_{k_i} - f'(x_{k_i} - q_i\, y_{k_i})\, y_{k_i} \geqslant \delta - \frac{\delta}{2} = \frac{\delta}{2} > 0\,.$$

By virtue of (8) we get

$$f'(x_{k_i})\, y_{k_i} - f'(x_{k_i} - q_i\, y_{k_i})\, y_{k_i} \xrightarrow{i\to\infty} 0\,.$$

The contradiction of this statement demonstrates that

$$f'(x_k)\, y_k \to 0 \quad \text{as} \quad k \to \infty\,.$$

Due to inequality (9) we have

$$f'(x_k)\, y_k \geqslant \left| f'(x_k) \frac{e_j}{\|e_j\|} \right| (1-\varepsilon) \quad \text{for} \quad j=1,2,\ldots, k=0,1,2,\ldots$$

which implies

$$f'(x_k)\, e_j \to 0 \quad \text{as} \quad k \to \infty \quad \text{for} \quad j=1,2,\ldots\,.$$

Consequently

$$\sum_{j=1}^{m} b_k^j f'(x_k)\, e_j \to 0 \quad \text{as} \quad k \to \infty \quad \text{for} \quad m \in N \tag{13}$$

if b_k^j's are uniformly bounded.

Assumption (8) implies the boundness of $f'(x)$ on W_0. Applying (13) we can say that

$$f'(x_k)\, v \to 0 \quad \text{as} \quad k \to \infty \quad \text{for} \quad v \in X$$

which means

$$f'(x_k) \to 0 \text{ (weakly*) as } k \to \infty\,.$$

4. Convergence of the algorithm

Now we can formulate the sufficient conditions for the strong convergence of the sequence $x_k \in X$, such that

$$f(x_k) \geqslant f(x_{k+1}) \quad \text{for} \quad k=0,1,2,\ldots \tag{14}$$

$$f'(x_k) \to 0 \text{ (weakly*) as } k \to \infty \tag{15}$$

to the point $\bar{x} \in X$, such that $f(\bar{x}) < f(x)$ for $x \in X, x \neq \bar{x}$.

THEOREM 2. Let X denote a reflexive Banach space, let f be continuously Fréchet--differentiable on X and let there exists, a constant $c>0$, such that

$$f(x) - f(y) \geqslant (f'(y), x-y) + c\,\|x-y\|^2 \quad \text{for} \quad x, y \in X\,. \tag{16}$$

Moreover we assume that the sequence $x_k \in X$ satisfies the conditions (14), (15).

Then there exists a limit

$$\lim_{k\to\infty} x_k = \bar{x} \in X$$

and

$$f'(\bar{x})=0, f(\bar{x})<f(x) \quad \text{for} \quad x\in X, x\neq\bar{x}.$$

Proof. Since f is weakly lower-semicontinuous on X and

$$f(x_i)\to\infty \quad \text{if} \quad ||x_i||\to\infty,$$

then due to the Generalized Theorem of Weierstrass there exists $\bar{x}\in X$ such that $f(\bar{x})\leqslant f(x)$ for $x\in X$.

By the definition of Fréchet-differential we have $f'(\bar{x})=0$. If $f(x_1)\leqslant f(x)$ for $x\in X$, and $f(x_2)\leqslant f(x)$ for $x\in X$, then

$$0=f(x_1)-f(x_2)\geqslant(f'(x_2), x_1-x_2)+c\,||x_1-x_2||^2\geqslant 0.$$

Hence $x_1=x_2=\bar{x}$.

Let X_n denote a linear subspace spanned by $\{e_1, e_2, ..., e_n\}$. By virtue of (15) we have

$$\lim_{k\to\infty}(f'(x_k), v_k^n)=0, \quad \text{for} \quad n=1, 2, 3, ...,$$

if the sequences $v_k^n\in X_n$ are bounded.

Applying (16) we get

$$f(x_k+v_k^n)-x(x_k)\geqslant(f'(x_k)\,v_k^n)+c\,||v_k^n||^2,\ k=0, 1, 2, ...$$
$$n \text{ — fixed.}$$

Hence

$$\lim_{k\to\infty} f(x_k+v_k^n)\geqslant \lim_{k\to\infty} f(x_k)=f^*\geqslant f(\bar{x}).$$

We choose sequences $\bar{v}_k^n\in X_z$ for $n=1, 2, ...,$ such that

$$\lim_{n\to\infty}\bar{v}_k^n=\bar{x}-x_k \quad \text{for} \quad k=0, 1, 2, ...$$

(such sequences exist because a linear subspace spanned by $(e_1, e_2, ...)$ is dense in X). Then

$$f(\bar{x})\leqslant f^*\leqslant \lim_{k\to\infty}\ \inf_{v^n\in X_n} f(x_k+v^n)\leqslant \lim_{k\to\infty}\ \lim_{n\to\infty} f(x_k+\bar{v}_k^n)=f(\bar{x}),$$

which means

$$\lim_{k\to\infty} f(x_k)=f(\bar{x}).$$

Using (16) we obtain

$$f(x_k)-f(\bar{x})\geqslant(f'(\bar{x}), x_k-\bar{x})+c\,||\bar{x}-x_k||^2,\ k=0, 1, 2, ...$$

and now it follows that

$$\lim_{k\to\infty} x_k=\bar{x}\in X$$

and $f(\bar{x})<f(x)$ for $x\in X, x\neq\bar{x}$. Q.E.D.

Observe that the assumption (16) in Theorem 2 can also be replaced by some other conditions used by Loridan [2], [3], for example, instead of the condition (16) we may assume that f is uniformly convex.

THEOREM 2 can be applied to the problem:

$$\text{let } H_1, H_2 \text{ be Hilbert spaces, } A \in L(H_1, H_2).$$

Then

$$f(x) = \|Ax - y_0\|^2$$

satisfies the assumptions of Theorem 2 and the minimization of functional f yields a solution of the equation $Ax = y_0$.

5. Examples

We give here some examples of sets $\Omega \subset R^m$ and functions $\varphi: \Omega \to X$, such that $e_i \in \varphi(\Omega)$ for $i = 1, 2, \ldots$.

Example 1. Let $X = L^2[a, b]$, $a > 0$ with $e_i = t^i$, $i = 0, 1, 2, \ldots$, $t \in [a, b]$, then we assume

$$m = 1, \ \Omega = R, \quad \varphi(s) = t^s, \quad s \in \Omega, \quad t \in [a, b]$$

and obviously

$$\varphi(i) = e_i \quad \text{for} \quad i = 0, 1, 2 \ldots.$$

Example 2. Let $X = L^2[0, \pi]$ with linearly independent functions $1, \cos t, \sin t, \cos 2t, \sin 2t, \ldots$, then we assume

$$m = 3, \ \Omega = R^3 \quad \text{and} \quad \varphi(s) = \cos(s_1 t) + s_2 \sin(s_3 t)$$

$$t \in [0, \pi], \ s = (s_1, s_2, s_3), \ s_i \in R, \ i = 1, 2, 3.$$

Example 3. Let $X = [x(t) \in H_1[0, 1] : x(0) = 0]$, $e_k = t \exp(kt)$ for $k = 0, 1, 2, \ldots$, $t \in [0, 1]$, where $H_1[0, 1]$ — a Sobolev space.

Let

$$\begin{cases} f(x) = \int_0^1 \{[x'(t)]^2 + [x(t)]^2 + 2t^2 x(t)\}\, dt, \\ x(0) = 0. \end{cases}$$

Problem: minimize $f(x)$, $x \in X$.

The solution of this problem is

$$\bar{x}(t) = (2 - c) \exp t + c \exp(-t) - t^2 - 2$$
$$f(\bar{x}) = -0.051100855\ldots,$$

where $c = 1.11353988\ldots$.

We take $\Omega=(-\infty, \infty)$ and $\varphi(s)=t \exp(st)$. In this case the new method consists in the minimization of the function

$$h_k(\alpha, s)=f\left(x_k(t)+\alpha t \exp(st)\right)$$
$$t \in [0, 1], \quad \text{for} \quad k=0, 1, 2, \ldots$$

and

$$x_k(t)=x_{k-1}(t)-\alpha_k t \exp(s_k t) \quad \text{for} \quad k=1, 2, \ldots,$$

where

$$x_0(t)\equiv 0 \quad \text{for} \quad t \in [0, 1]$$
$$h_k(\alpha_k, s_k)\leqslant h_k(\alpha, s) \quad \text{for} \quad \alpha, s \in (-\infty, \infty).$$

We have performed the computations and have obtained

i	α_i	s_i	$f(x_k)$
1	−0.278314	−0.42	−0.04984
2	0.312676−7	12.62	−0.05046
3	0.809892−1	−4.68	−0.05083
4	−0.113865−1	−0.30	−0.05093

Similar examples can be constructed in the spaces of functions of several variables.

Acknowledgments. I would like to thank Prof. K. Goebel and Prof. A. P. Wierzbicki for helpful comments on this paper.

References

1. S. Grzegórski, Numerical methods of minimizing a functional defined on a Banach space and their applications to the problems of calculus of variations and optimal control. (Numeryczne metody minimalizacji funkcjonału określonego na przestrzeni Banacha i ich zastosowanie do zadań rachunku wariacyjnego i sterowania optymalnego). Doctoral dissertation. Univ. MSC in Lublin, Now. 1975.
2. P. Loridan, Sur un procède d'optimisation utilisant simultanement les méthodes de pénalisation et des variations localles. *SIAM J. Control* **11** (1973).
3. P. Loridan, Sur la minimisation de fonctionnelles convexes par pénalisation. R.I.R.O. 5, R-1 (1971).

Received, July 1977.

Metoda zanurzania bazy funkcyjnej w zagadnieniach optymalizacji

Przedstawiono metodę iteracyjną minimalizacji funkcjonału bez ograniczeń. Metoda ta polega na zastąpieniu zadania minimalizacji funkcjonału określonego na przestrzeni nieskończenie wielowymiarowej równoważnym ciągiem zadań minimalizacji funkcji określonych na przestrzeniach m-wymiarowych, przy czym m jest liczbą ustaloną. Omówiono też zbieżność metody.

Метод погружения функционального базиса в оптимизационных задачах

В работе представлен итерационный метод минимизации функционала без ограничений. Метод состоит в замене задачи минимизации функционала, определенного на бесконечном многомерном пространстве, эквивалентной последовательностью задач минимизации функций определенных на m-мерных пространствах, причем m является определенным числом. Рассматривается сходимость метода.