

**A subgradient method with space dilation
for minimizing convex functions**

by

PHAM TRONG QUAT

Automatic Control Institute
Technological University of Warsaw
Nowowiejska 15/19
00-665 Warszawa, Poland

In this paper an implementable algorithm using the operation of space dilation for finding the minimum of a convex, not necessarily differentiable function f is investigated. The method is based on combining, modifying and extending the nonsmooth optimization works of Shor [15] and Wolfe [16]. Global convergence of the algorithm is established. the algorithm is conceptually simple and easy to implement.

1. Introduction

In what follows, we shall be concerned with an algorithm for solving the following optimization problem

$$\min_{x \in R^n} f(x), \quad (1)$$

where f is a real-valued function defined on R^n . It is assumed throughout this paper that the function f is continuous, convex and

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty. \quad (2)$$

This problem is "nonsmooth" in the sense that the function f need not be differentiable everywhere.

The difficulties in minimizing a nonsmooth function are well discussed in [16], where an implementable descent algorithm for a convex function is given. This algorithm for a quadratic function coincides with the conjugate gradient method [6] and, hence, has finite termination in this case, as do algorithms of Lemarechal [10] or of Shor [14].

The descent approach for convex function of Bertsekas and Mitter [1] has been made implementable by Lemarechal [9], and has been extended in theory to locally

Lipschitz function by Goldstein [5]. Descent algorithms for min-max objectives, which are also difficult to implement, are given in Demjanov [2] and Goldstein [4]. The readily implementable algorithms of Kiwiel [7], [8] calculate stationary points for nonsmooth minimization problems. These methods generalize several efficient algorithms for simpler problems due to Pshenichny and Danilin [13] and Lemarechal [11]. Owing to suitable strategies for selecting and aggregating the past subgradient information, Kiwiel's algorithms have flexible storage requirements and work per iteration that can be controlled by the user.

Shor's algorithms [14] are nondescent methods. Shor suggested to transform the space metric at each iteration so as to accelerate convergence of the subgradient method. He used operators of space dilation of the following type. Let $s \in R^n$, $\|s\| = 1$, $\alpha > 0$. Then a linear operator $R_\alpha(s)$ such that

$$R_\alpha(s) x = x + (\alpha - 1) s s^T x$$

is referred to as the space-dilation operator acting in the direction s with the coefficient α .

Our algorithm combines, extends and modifies ideas of Wolfe [16] and Shor [15]. We use the operator of space dilation for finding direction at each iteration and our choice of step size is based on Wolfe's idea. Global convergence of the algorithm is established. The algorithm is conceptually simple and easy to implement. In particular, it does not require the solution of an auxiliary problem for generating search directions as in [7], [9], [16]. Hence it can be used for solving large-scale problems.

In Section 2 we present the algorithm, while its convergence is discussed in Section 3. In Section 4 we present a simple numerical example.

2. Algorithm

The algorithm uses positive parameters δ , m_1 , m_2 , β_1 , $\beta_2 < 1$ satisfying

$$m_2 < m_1 < 0.5, \quad (3)$$

$$m_1/(1-m_1) \leq \beta_1 < 1 \quad (4)$$

and a sequence of positive numbers $\{\delta_k\}$ satisfying

$$\delta_k \rightarrow 0 \text{ as } k \rightarrow +\infty. \quad (5)$$

Initially we have a starting point x^0 and some $g_0 \in \partial f(x^0)$, where $\partial f(x)$ denotes the set of all subgradients of f at x . Suppose a point x^k is known. To find the next point x^{k+1} the algorithm realizes the following iterative process.

STEP 0: Set $x_{k,0} = x^k$, $s_{k,0} = g_k$, where $g_k \in \partial f(x_k)$, and $\varepsilon_k = \max\{1/\sqrt{f(x^{k-1}) - f(x^k)}, \delta_k\}$. Set $i=0$.

STEP 1: If $\|s_{k,i}\| \leq \varepsilon_k$, set $x^{k+1} = x_{k,i}$, increase k by 1 and go to Step 0; otherwise, go to Step 2.

STEP 2: (Line search: see below for details). Find $\tau_{k,i} \geq 0$ and some $g_{k,i+1} \in \partial f(x_{k,i} - \tau_{k,i} s_{k,i})$ such that

$$f(x_{k,i} - \tau_{k,i} s_{k,i}) \leq f(x_{k,i}) - m_2 \tau_{k,i} \|s_{k,i}\|^2 \quad (6)$$

and

$$\langle g_{k,i+1}, s_{k,i} \rangle \leq m_1 \|s_{k,i}\|^2. \quad (7)$$

Set $x_{k,i+1} = x_{k,i} - \tau_{k,i} s_{k,i}$ and go to Step 3.

STEP 3: If $\|x^k - x_{k,i+1}\| > \delta$ or $f(x^k) - f(x_{k,i+1}) > \delta$, set $x^{k+1} = x_{k,i+1}$, increase k by 1 and go to Step 0; otherwise, go to Step 4.

STEP 4: If $\langle g_{k,i+1}, s_{k,i} - g_{k,i+1} \rangle \geq 0$ go to Step 5, otherwise, go to Step 6.

STEP 5: Set

$$\xi_{k,i+1} = (g_{k,i+1} - s_{k,i}) / \|g_{k,i+1} - s_{k,i}\|,$$

$$s_{k,i+1} = R_{\beta_1}(\xi_{k,i+1}) s_{k,i},$$

increase i by 1 and go to Step 1.

STEP 6: Set

$$\xi_{k,i+1} = (s_{k,i} - g_{k,i+1}) / \|s_{k,i} - g_{k,i+1}\|,$$

$$q_0 = g_{k,i+1} \text{ and } j = 0.$$

i) Set

$$q_{j+1} = R_{\beta_2}(\xi_{k,i+1}) q_j.$$

ii) If

$$\|q_{j+1}\|^2 \leq \|s_{k,i}\|^2 \left(1 + \frac{(\beta_1^2 - 1)(1 - 2m_1) \varepsilon_k^2}{\|s_{k,i} - g_{k,i+1}\|^2} \right) = \gamma_k, \quad (8)$$

set $s_{k,i+1} = q_{j+1}$, increase i by 1 and go to Step 1; otherwise, increase j by 1 and go to i).

Condition (7) can be verified by the following lemma.

LEMMA 1 (see [3]). Let f be a convex function on R^n , and s a fixed vector of R^n satisfying the inequality

$$f(x - \tau s) \geq f(x) - m_1 \tau \|s\|^2 \quad (7')$$

for some $m_1 > 0$, $\tau > 0$. Then we have

$$\langle g, s \rangle \leq m_1 \|s\|^2$$

for any $g \in \partial f(x - \tau s)$.

LINE SEARCH: For given $x_{k,i}$ and $s_{k,i}$, with $\|s_{k,i}\| \neq 0$, suppose that the directional derivative

$$f'(x_{k,i} - s_{k,i}) = \max \{ \langle g, -s_{k,i} \rangle : g \in \partial f(x_{k,i}) \}$$

and the vector $g' \in \text{Argmax} \{ \langle g, -s_{k,i} \rangle : g \in \partial f(x_{k,i}) \}$ are computed. Now let us consider the following two cases.

FIRST CASE: $f'(x_{k,i} - s_{k,i}) \geq -m_1 \|s_{k,i}\|^2$.

Then the vector g' satisfies inequality (7). In this case we set $\tau_{k,i} = 0$.

SECOND CASE: $f'(x_{k,i} - s_{k,i}) < -m_1 \|s_{k,i}\|^2$.

By (2) and (3), there exists $\tau > 0$ satisfying conditions (6) and (7'). Then we can find $\tau_{k,i}$ as follows.

Define

$$L = \{ \tau > 0 : f(x_{k,i} - \tau s_{k,i}) \leq f(x_{k,i}) - m_2 \tau \|s_{k,i}\|^2 \},$$

$$R = \{ \tau > 0 : f(x_{k,i} - \tau s_{k,i}) \geq f(x_{k,i}) - m_1 \tau \|s_{k,i}\|^2 \}.$$

Choose $\tau > 0$. Set $\tau_m = 0$, $\tau_M = 0$.

(a) If $\tau \in L \setminus R$ go to (b).

If $\tau \in L \cap R$ go to (d).

If $\tau \in R \setminus L$ go to (c).

(b) If $\tau_m = 0$, replace τ_m by τ and τ by 2τ ; otherwise, replace τ_m by τ and τ by $(\tau_M - \tau_m)/2$; and go to (a).

(c) Replace τ_M by τ and τ by $(\tau_M - \tau_m)/2$, and go to (a).

(d) Set $\tau_{k,i} = \tau$ and stop.

Using the proof of Lemma 1 in [16] it is easy to see that the above process is finite.

The following results show that the algorithm is well-defined.

LEMMA 2. If $s, g \in R^n$, $\xi = (s-g)/\|s-g\|$, $\langle g, s-g \rangle < 0$ and $\langle s, g-s \rangle < 0$ then $\langle R_\beta(\xi)g, s - R_\beta(\xi)g \rangle < 0$ for any $\beta \in (0, 1)$.

Proof. Denote $\lambda = (\beta-1) \langle g, s-g \rangle / \|s-g\|^2$. From the assumption it follows that $\lambda > 0$ and $\langle s, g \rangle - \langle s, s \rangle < 0$. Thus $\langle g, g \rangle - \langle g, s \rangle < \langle g, g \rangle - 2\langle s, g \rangle + \langle s, s \rangle$, so $-\langle g, s-g \rangle \leq \|s-g\|^2$, and finally $-\langle g, s-g \rangle / \|s-g\|^2 < 1$. Hence $\lambda < 1$. From the definition of the operator of space dilation we obtain

$$s = R_\beta(\xi)g = (s-g) - [(\beta-1) \langle g, s-g \rangle / \|s-g\|^2] (s-g) = (1-\lambda) (s-g)$$

and $R_\beta(\xi)g = (1-\lambda)g + \lambda s$. Then we have

$$\begin{aligned} \langle R_\beta(\xi)g, s - R_\beta(\xi)g \rangle &= \langle (1-\lambda)g + \lambda s, (1-\lambda)(s-g) \rangle = \\ &= (1-\lambda) \langle \lambda(s-g) + g, s-g \rangle = (1-\lambda) [\langle g, s-g \rangle + \lambda \|s-g\|^2] = \\ &= (1-\lambda) [\langle g, s-g \rangle + (\beta-1) \langle g, s-g \rangle] = \beta(1-\lambda) \langle g, s-g \rangle < 0. \end{aligned}$$

This completes the proof. ■

From Lemma 1 it follows that s' lies between $s_{k,i}$ and $g_{k,i+1}$. Therefore

$$\|h\|^2 < \|s'\|^2 < \gamma_k. \quad (11)$$

From inequalities (11) and (9) we see that there exists j such that

$$\|q_j\|^2 < \gamma_k.$$

This completes the proof. ■

LEMMA 4. $s_{k,i}$ is a convex combination of two points $s_{k,i-1}$ and $g_{k,i}$, for any k and $i=1, 2, \dots, t(k)$.

Proof. From the proofs of Lemmas 2 and 3 it is evident that in the case when Step 6 is performed, $s_{k,i}$ is always a convex combination of $s_{k,i-1}$ and $g_{k,i}$. It remains to consider the case where Step 5 of the algorithm is carried out. In this case we have

$$s_{k,i} = R_{\beta_1}(\xi_{k,i}) s_{k,i-1},$$

where

$$\xi_{k,i} = (g_{k,i} - s_{k,i-1}) / \|g_{k,i} - s_{k,i-1}\|.$$

From the definition of the operator of space dilation we obtain

$$\begin{aligned} s_{k,i} &= s_{k,i-1} + (\beta_1 - 1) \langle s_{k,i-1}, \xi_{k,i} \rangle \xi_{k,i} = \\ &= s_{k,i-1} + [(\beta_1 - 1) \langle s_{k,i-1}, g_{k,i} - s_{k,i-1} \rangle / \|g_{k,i} - s_{k,i-1}\|^2] (g_{k,i} - \\ &\quad - s_{k,i-1}) = (1 - \lambda) s_{k,i-1} + \lambda g_{k,i}, \end{aligned}$$

where

$$\lambda = (\beta_1 - 1) \langle s_{k,i-1}, g_{k,i} - s_{k,i-1} \rangle / \|g_{k,i} - s_{k,i-1}\|^2.$$

We shall show that $\lambda \in [0, 1]$. We have

$$\langle s_{k,i-1}, g_{k,i} - s_{k,i-1} \rangle \leq m_1 \|s_{k,i-1}\|^2 - \|s_{k,i-1}\|^2 = -(1 - m_1) \|s_{k,i-1}\|^2.$$

Thus

$$\lambda \geq (1 - \beta_1) (1 - m_1) \|s_{k,i-1}\|^2 / \|g_{k,i} - s_{k,i-1}\|^2 > 0.$$

Besides

$$\begin{aligned} \langle s_{k,i-1}, g_{k,i} - s_{k,i-1} \rangle &= -\|s_{k,i-1}\|^2 + \langle s_{k,i-1}, g_{k,i} \rangle \geq \\ &\geq -\|s_{k,i-1}\|^2 + 2 \langle s_{k,i-1}, g_{k,i} \rangle - m_1 \|s_{k,i-1}\|^2 = \\ &= -\|s_{k,i-1} - g_{k,i}\|^2 - m_1 \|s_{k,i-1}\|^2 + \|g_{k,i}\|^2. \end{aligned}$$

Thus

$$\begin{aligned} \lambda &\leq (1 - \beta_1) [1 + (m_1 \|s_{k,i-1}\|^2 - \|g_{k,i}\|^2) / \|s_{k,i-1} - g_{k,i}\|^2] \leq \\ &\leq (1 - \beta_1) [1 + m_1 \|s_{k,i-1}\|^2 / \|s_{k,i-1} - g_{k,i}\|^2]. \end{aligned}$$

On the other hand

$$\begin{aligned} \|s_{k,i-1} - g_{k,i}\|^2 &= \|s_{k,i-1}\|^2 + \|g_{k,i}\|^2 - 2 \langle s_{k,i-1}, g_{k,i} \rangle \geq \\ &\geq (1-2m_1) \|s_{k,i-1}\|^2 + \|g_{k,i}\|^2 \geq (1-2m_1) \|s_{k,i-1}\|^2. \end{aligned}$$

Hence

$$\lambda \leq (1-\beta_1) [1+m_1/(1-2m_1)] = (1-\beta_1) (1-m_1)/(1-2m_1).$$

From the assumption that $m_1(1-m_1) < \beta_1 < 1$ it follows that

$$\lambda \leq [1-m_1/(1-m_1)] (1-m_1)/(1-2m_1) = 1.$$

This completes the proof.

LEMMA 5. The process of finding x^{k+1} is finite for any k .

Proof. It is easy to observe that the process of calculating x^{k+1} could be stopped with the equality $x^{k+1} = x_{k, \tau(k)}$ as soon as we have one of the following three cases:

- a) $\|x^k - x_{k, \tau(k)}\| > \delta$,
- b) $f(x^k) - f(x_{k, \tau(k)}) > \delta$,
- c) $\|s_{k, \tau(k)}\| \leq \varepsilon_k$.

Let us assume that for some fixed k and all i cases a) and b) do not occur. Then we shall prove that there exists some finite $\tau(k)$ sufficiently large such that $\|s_{k, \tau(k)}\| \leq \varepsilon_k$. Combining (2), Lemma 4 and the local boundeness of the subdifferentials, it is easily seen that there exists $b > 0$ such that

$$\begin{aligned} \|g_{k,i}\| &\leq b \quad \text{for all } i, \\ \|s_{k,i}\| &\leq b \quad \text{for all } i. \end{aligned}$$

Assume, for contradiction purposes, that $\|s_{k,i}\| > \varepsilon_k$ for all i . Step 6 yields

$$\|s_{k,i+1}\|^2 \leq \|s_{k,i}\|^2 [1 + (\beta_1^2 - 1) (1 - 2m_1) \varepsilon_k^2 / \|s_{k,i} - g_{k,i+1}\|^2] \quad (12)$$

for all i .

In Step 5 we have

$$\|s_{k,i+1}\|^2 = \|R_{\beta_1}(\xi_{k,i+1}) s_{k,i}\|^2 = \|s_{k,i}\|^2 + (\beta_1^2 - 1) \langle s_{k,i}, \xi_{k,i+1} \rangle^2$$

and from inequality (10) it follows that in this case we also have (12). From inequality (12) one obtains

$$\begin{aligned} \|s_{k,i+1}\|^2 &\leq \|s_{k,i}\|^2 \left[1 + \frac{(\beta_1^2 - 1) (1 - 2m_1) \varepsilon_k^2}{\|g_{k,i+1}\|^2 - 2 \langle g_{k,i+1}, s_{k,i} \rangle + \|s_{k,i}\|^2} \right] \leq \\ &\leq \|s_{k,i}\|^2 [1 + (\beta_1^2 - 1) (1 - 2m_1) \varepsilon_k^2 / 4b^2] \end{aligned}$$

for all i . This yields

$$\begin{aligned} \|s_{k,i}\| &\leq \|s_{k,0}\| [1 + (\beta_1^2 - 1) (1 - 2m_1) \varepsilon_k^2 / 4b^2]^{i/2} \leq \\ &\leq b [1 + (\beta_1^2 - 1) (1 - 2m_1) \varepsilon_k^2 / 4b^2]^{i/2}. \end{aligned}$$

If

$$i > 2(1n\varepsilon_k - 1nb)/1n [1 - (1 - \beta_1^2) (1 - 2m_1) \varepsilon_k^2 / 4b^2]$$

then $1n \|s_{k,i}\| \leq 1n \varepsilon_k$, i.e. $\|s_{k,i}\| \leq \varepsilon_k$. This completes our proof. ■

The above lemmas show that our algorithm generates an infinite sequence of points $\{x^k\}$. In the next section we shall prove that this sequence minimizes f .

3. Convergence of the algorithm

THEOREM. Assume that Conditions (2) — (5) are satisfied and let $\{x^k\}_{k=0}^{\infty}$ be the sequence generated by the algorithm. Then

$$\lim_{k \rightarrow +\infty} f(x^k) = f^* = \min_{x \in R^n} f(x).$$

Proof. From the description of the algorithm and Condition (2) it is evident that the sequence $\{f(x^k)\}_{k=0}^{\infty}$ is nonincreasing and bounded from below, therefore there exists

$$\lim_{k \rightarrow \infty} f(x^k) = f_{\infty} \geq f^* = \min_{x \in R^n} f(x), \quad (13)$$

Then

$$\varepsilon_k = \max \left\{ \sqrt{f(x^{k-1}) - f(x^k)}, \delta_k \right\} \rightarrow 0 \text{ as } k \rightarrow \infty. \quad (14)$$

Let us now prove that there exists an infinite subset of indices $K \subset N$ such that

$$\eta^k = \max \{ \|x^k - x_{k,i}\| : i = 1, 2, \dots, t(k) \} \rightarrow 0, \quad (15)$$

as $k \rightarrow \infty$, $k \in K$.

Let us observe that there must exist an infinite subset $K \subset N$ such that

$$f(x^k) - f(x^{k+1}) \leq f(x^{k-1}) - f(x^k) \text{ for all } k \in K. \quad (16)$$

Indeed, if this was not true one would have $f(x^k) \rightarrow -\infty$, a contradiction. We shall establish (15) for $k \in K$.

From the description of the algorithm we have

$$\begin{aligned} 0 \leq \|x^k - x_{k,i}\| &= \left\| \sum_{t=0}^{i-1} (x_{k,t} - x_{k,t+1}) \right\| = \left\| \sum_{t=0}^{i-1} \tau_{k,t} s_{k,t} \right\| \leq \sum_{t=0}^{i-1} \tau_{k,t} \|s_{k,t}\| = \\ &= \sum_{t=0}^{i-1} \tau_{k,t} \|s_{k,t}\|^2 / \|s_{k,t}\| \leq \frac{1}{\varepsilon_k} \sum_{t=0}^{i-1} \tau_{k,t} \|s_{k,t}\|^2 \leq \frac{1}{m_2 \varepsilon_k} \sum_{t=0}^{i-1} [f(x_{k,t}) - \\ &\quad - f(x_{k,t+1})] = \frac{1}{m_2 \varepsilon_k} [f(x_{k,0}) - f(x_{k,i})], \end{aligned}$$

for $i = 1, 2, \dots, t(k)$.

Thus

$$0 \leq \|x^k - x_{k,i}\| \leq \frac{1}{m_2 \varepsilon_k} [f(x^k) - f(x^{k+1})], \quad (17)$$

for $i = 1, 2, \dots, t(k)$. From (16) and (17) we get for $k \in K$ and $i = 1, \dots, t(k)$

$$0 \leq \|x^k - x_{k,i}\| \leq \frac{1}{m_2 \varepsilon_k} [f(x^{k-1}) - f(x^k)].$$

By the definition of ε_k , since $f(x^{k-1}) - f(x^k) \rightarrow 0$ as $k \rightarrow +\infty$, the right-hand side of the above inequality tends to 0, as $k \rightarrow +\infty$. This establishes (15).

From the description of the algorithm it is clear that $g_{k,i} \in \partial f(x^k; \eta^k)$, $i=0, 1, \dots, t(k)$, where $\eta^k \rightarrow 0$, as $k \rightarrow +\infty$, $k \in K$, from (15). Here $\partial f(x, \eta)$ denotes the set of all η -subgradients of f at x defined by

$$\partial f(x, \eta) = \text{conv} \bigcup \{ \partial f(y) : \|y - x\| \leq \eta \}.$$

On the basis of Lemma 4 it follows that $s_{k,i} \in \partial f(x^k, \eta^k)$ for $k \in K$. From Assumption (2) it is easy to see that there exist x' and an infinite subset $K' \subset K$ such that $x^k \rightarrow x'$ as $k \rightarrow +\infty$ and $k \in K'$. By (14) and (15), we have

$$s_{k,t(k)} \rightarrow 0 \text{ as } k \rightarrow +\infty, k \in K',$$

where $s_{k,t(k)} \in \partial f(x^k, \eta^k)$. We know that the map $(x, \eta) \rightarrow \partial f(x, \eta)$ is upper semi-continuous on R^n (see [12]). Therefore $0 \in \partial f(x', 0) = \partial f(x')$, which means that x' is a minimum point of f and

$$\lim_{\substack{k \rightarrow +\infty \\ k \in K'}} f(x^k) = f(x') = f^*. \quad (18)$$

From (13) and (18) we obtain

$$\lim_{k \rightarrow +\infty} f(x^k) = f^* = \min_{x \in R^n} f(x).$$

The theorem is proved.

From the proof of the above theorem we have the following.

COROLLARY. *In addition to the assumptions of the above theorem, suppose that f is strongly convex on R^n . Then the sequence $\{x^k\}_{k=0}^\infty$ converges to the minimum point x^* of f on R^n .*

4. Example

The objective function to be minimized is

$$f(x) = \max \{ f'(x), f''(x) \},$$

where $x \in R^2$, $f'(x) = 4x_1^2 + (x_2 - 4)^2$, $f''(x) = (2x_1 - 4)^2 + x_2^2$. For this problem, the optimal solution is $x^* = (1, 2)$ with $f(x^*) = 8$. Let LF denote the number of calculations of the function values and LG that of the function subgradients.

Our algorithm used the starting point $x^0 = (2, 0)$ with $f(x^0) = 32$. For $\beta_1 = \beta_2 = 0.3$, $m_1 = 0.23$, $m_2 = 0.17$, $\delta_k = \frac{1}{\sqrt{k}}$, we obtained

$$\begin{aligned} k &= 18, \quad \sum_{k=0}^{18} t(k) = 67, \quad f(x^{18}) = 8.0001309, \\ x_1^{18} &= 1.0020049, \quad x_2^{18} = 2.0039851, \\ \text{LF} &= 673, \quad \text{LG} = 155. \end{aligned}$$

References

- [1] BERTSEKAS D.P., MITTER S.K. A Descent Numerical Method for Optimization Problems with Nondifferentiable Const Functionals. *SIAM J. Control.* 11 (1973), 637-652.
- [2] DEMJANOV V.F. Algorithms for Some Minimax Problems. *J. Comput. System Sci.* 2 (1968), 342-380.
- [3] DEMJANOV V.F., VASILIEV L.V. Nondifferentiable Optimization (In Russian). Moscow, 1981.
- [4] GOLDSTEIN A.A. Optimization with Corners. In O. L. Mangasarian, R.R. Meyer and S.M. Robinson, eds., *Nonlinear Programming 2*. New York, Academic Press, 1975, 215-230.
- [5] GOLDSTEIN A.A. Optimization of Lipschitz Continuous Functions. *Mathematical Programming* 13, (1977), 14-22.
- [6] HESTENES M.R., STIEFEL E. Methods of Conjugate Gradients for Solving Linear Systems. *J. Res. Nat. Bur. Standards.* 49 (1952), 409-436.
- [7] KIWIŁ K.C. Efficient Algorithms for Nonsmooth Optimization and Their Applications. Doctoral Thesis. Politechnika Warszawska, 1982 (In Polish).
- [8] KIWIŁ K.C. An Aggregate Subgradient Method for Nonsmooth Convex Minimization. *Mathematical Programming* 27 (1983), 320-341.
- [9] LEMARECHAL C. An Algorithm for Minimizing Convex Functions. In J. L. Rosenfeld, ed. *Information Processing*. Amsterdam, North-Holland, 1974, 552-556.
- [10] LEMARECHAL C. An Extension of Davidon Methods to Nondifferentiable Problems. In M.L. Balinski and P. Wolfe, eds., *Nondifferentiable Optimization*. Math. Programming Study 3, Amsterdam, North-Holland, 1975, 95-109.
- [11] LEMARECHAL C. Nonsmooth Optimization and Descent Methods. RR-78-4, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1978.
- [12] MIEFLIN R. An Algorithm for Constrained Optimization with Semismooth Functions. *Mathematics of Operations Research* 2 (1977), 191-207.
- [13] PSHENICHNY B.N., DANILIN Ju.M. Numerical Methods for Extremum Problems (In Russian). Moscow, Nauka, 1975.
- [14] SHOR N.Z. Methods for Minimizing Nondifferentiable Functions and Their Applications. Kiev, Naukova Dumka, 1979.
- [15] SHOR N.Z., ZHURBENKO N.G. A Minimization Method Using Space Dilation in the Direction of the Difference of Two Successive Gradients. *Kibernetika* 7 (3) (1971), 51-59.
- [16] WOLFE P. Method of Conjugate Subgradients for Minimizing Nondifferentiable Functions. In M.L. Balinski and P. Wolfe, eds., *Nondifferentiable Optimization*. Math. Programming Study 3, Amsterdam, North-Holland, 1975, 145-173.

Received, March 1984.

Subgradientowa metoda z rozciąganiem przestrzeni dla minimizacji wypukłych funkcji

W artykule rozpatrywana jest subgradientowa metoda z rozciąganiem przestrzeni dla minimizacji wypukłych funkcji. Przedstawiono model algorytmu oraz udowodniono zbieżność przy ogólnych założeniach dotyczących funkcji i parametrów opisujących metodę. Metoda daje się łatwo zaprogramować. Zamieszczono wyniki obliczeń prostego przykładu numerycznego.

Субградиентный метод с растяжением пространства для минимизации выпуклых функций

В статье рассматривается субградиентный метод с растяжением пространства для минимизации выпуклых функций. Представлена модель алгоритма и доказана сходимость при общих предположениях, касающихся функций и параметров, описывающих метод. Для метода несложно разработать программы. Приведены результаты вычислений для простого численного примера.