CONVERGENCE RATE OF THE GRADIENT DESCENT
METHOD WITH DILATATION OF THE SPACE

N. Z. Shor

1. INTRODUCTION

We consider a convex function \( f(x) \), defined throughout the Euclidean space \( \mathbb{R}^n \) and having the following property:

\[
\lim_{||x|| \to \infty} f(x) = +\infty.
\]  

(1)

It follows from property (1) that the domain of the minima \( m^* \) of \( f(x) \) is a closed bounded set. Suppose the function \( f(x) \) takes the value \( m^* \) on \( m^* \).

The method of generalized descent with dilatation of the space was proposed in [1] to minimize convex functions (the method is abbreviated as GDDS). The basis of this method was laid for the special case of the step length and the space dilatation being regular, the value \( m^* \) being assumed known.

In this paper we recall the fundamental definitions and results of [1] and investigate in detail the algorithm when \( m^* \) is unknown in advance.

Then we obtain estimates of the convergence rate of the GDDS method and consider the practical problems of its use.

2. FUNDAMENTAL DEFINITIONS AND RESULTS

The generalized gradient of the function \( f(x) \) at the point \( x_0 \) is the vector \( \hat{g}_f(x_0) \), satisfying the following inequality for arbitrary \( x \in \mathbb{R}^n \):

\[
f(x) - f(x_0) \geq (\hat{g}_f(x_0), x - x_0).
\]  

(2)

Suppose we are given the vector \( \eta \neq 0 \). The generalized derivative of \( f(x) \) in the direction of \( \eta \) at the point \( x_0 \) is the number \( f_\eta(x_0) \) which is the projection of \( \hat{g}_f(x_0) \) in the direction of \( \eta \):

\[
f_\eta(x_0) = \frac{\left(\hat{g}_f(x_0), \eta\right)}{||\eta||}.
\]  

(3)

Taking \( \eta = x - x_0 \), we obtain

\[
f_{x - x_0}(x_0) \cdot ||x - x_0|| \leq f(x) - f(x_0)
\]  

(4)

or

\[
f_{x - x_0}(x) \cdot ||x - x_0|| \geq f(x) - f(x_0).
\]  

(5)

Let \( B \) be a linear operator, defined on \( \mathbb{R}^n \). Consider the function

\[
\psi(y) = f(By);
\]  

(6)

\( \phi(y) \) is a convex function and the generalized gradient of \( \phi(y) \) can be calculated from the following equation:

\[
\hat{g}_\phi(y_0) = B^* \hat{g}_f(x_0),
\]  

(7)

where \( x_0 = By_0 \) and \( B^* \) is the operator conjugate with \( B \).

Suppose we are given a vector \( \xi, \|\xi\| = 1 \). The dilatation operator \( R_\alpha(\xi) \) in the direction of \( \xi \) with coefficient \( \alpha \) is the linear operator which operates as follows on an arbitrary vector \( x \in \mathbb{R}^n \):

\[
R_\alpha(\xi)x = \alpha x_\xi + (x - x_\xi),
\]

where \( x_\xi = (x, \xi)\xi \).

The dilatation coefficient of the vector \( x \neq 0 \) under the action of the operator \( A \) is the quantity \( k_x(A) \) defined by the equation

\[
k_x(A) = \frac{\|Ax\|}{\|x\|}. \tag{9}
\]

If \( A \) and \( B \) are two operators, then

\[
k_x(AB) = k_{x(A)}k_{x(B)}. \tag{10}
\]

For the dilatation operator \( R_\alpha(\xi) \) we have

\[
k_x(R_\alpha(\xi)) = \sqrt{1 + (\alpha^2 - 1) \frac{(x, \xi)^2}{\|x\|^2}}. \tag{11}
\]

The algorithm for the generalized gradient method with dilatation of the space (the GDDS algorithm) is the algorithm for minimizing convex functions described below.

I. Prerequisites for the Application of the GDDS Algorithm.

For a given convex function \( f(x) \) there is an algorithm for computing \( \hat{g}_f(x) \) at an arbitrary point \( x \in \mathbb{R}^n \); operators are specified for computing the sequences of positive numbers \( \{h_k\} \) and \( k = 1, 2, \ldots \).

II. The 0 Step of the Algorithm.

We choose an initial approximation \( x = x_0 \) and a matrix \( B_0 = A_0^{-1} = E \).

III. The \((k + 1)\)th Step of the Algorithm, \( k = 0, 1, \ldots \).

We compute

\[
\hat{g}_f(x_k);
\]

\[
\hat{g}_{q_k}(y_k) = B_{q_k}^* \hat{g}_f(x_k),
\]

Note, \( q_k(y) = \frac{1}{2} (B_k y); y_k = A_k x_k; \)

\[
\hat{g}_{q_k}(y_k);
\]

\[
\epsilon_{k+1} = \frac{\hat{g}_{q_k}(y_k)}{\|g_{q_k}(y_k)\|};
\]

\[
h_{k+1};
\]

\[
\alpha_{k+1};
\]

\[
x_{k+1} = x_k - B_{h_{k+1}} \epsilon_{k+1};
\]

\[
R_{q_{k+1}}^{-1}(\epsilon_{k+1}) = R_{q_{k+1}}^{-1}(\epsilon_{k+1});
\]

\[
B_{k+1} = B_{h_{k+1}} R_{q_{k+1}}^{-1} (\epsilon_{k+1}).
\]

The various versions of the GDDS algorithm differ from one another in the methods of calculating the sequences \( \{h_k\} \) and \( \{\alpha_k\} \). In [1] we considered in detail the case when \( \{\alpha_k\} \) and \( \{h_k\} \) are chosen knowing the value of the function at the minimum point and proved that this method converged for certain assumptions. We chose \( h_{k+1} \) as follows:

\[
h_{k+1} = \frac{\gamma (f(x_k) - m)}{\|g_{q_k}(y_k)\|}, \tag{16}
\]

103