Applied Mathematics and Computation 250 (2015) 309–319

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

An Accelerated Double Step Size model in unconstrained optimization

Milena J. Petrović

University of Priština, Faculty of Science and Mathematics, Lole Ribara 29, 28000 Kosovska Mitrovica, Serbia

ARTICLE INFO

Keywords:
Line search
Gradient descent methods
Newton method
Convergence rate

ABSTRACT

This work presents a double step size algorithm with accelerated property for solving nonlinear unconstrained optimization problems. Using the inexact line search technique, as well as the approximation of the Hessian by an adequate diagonal matrix, an efficient accelerated gradient descent method is developed. The proposed method is proven to be linearly convergent for uniformly convex functions and also, under some specific conditions, linearly convergent for strictly convex quadratic functions. Numerical testings and comparisons show that constructed scheme exceeds some known iterations for unconstrained optimization with respect to all three tested properties: number of iterations, CPU time and number of function evaluations.

1. Introduction and preliminaries

In this research the nonlinear unconstrained optimization problem

$$\min f(x), \quad x \in \mathbb{R}^n \tag{1.1}$$

is considered, where f is a differentiable function.

The iteration that is suggested has the general form

$$x_{k+1} = x_k + \alpha_k s_k + \beta_k d_k. \tag{1.2}$$

Here, x_{k+1} represents a new iterative point, x_k the previous one, positive real quantities α_k and β_k are two step sizes, while vectors s_k and d_k generate search directions. Each of the two mentioned step sizes are chosen in a way that the value of the objective function decreases through the iteration:

$$f(x_k + \alpha_k s_k + \beta_k d_k) < f(x_k).$$

This request is reached by means of the inexact *line search procedures*. Using these procedures, the derived iterative step provides non-increasing the function value (see [3,8,9,11,15,16,20,23]). In this work Armijo's backtracking line search technique is used for obtaining the optimal step length.

Furthermore, determination of the search direction is based on the descent condition. Actually, if l_k is the search direction in the kth iteration, then the next inequality has to be fulfilled:

$$g_k^T l_k \leqslant 0, \tag{1.3}$$

E-mail address: milena.petrovic@pr.ac.rs

http://dx.doi.org/10.1016/j.amc.2014.10.104

where g_k is the gradient vector at the point x_k. There are many schemes with differently defined search direction vectors and step length parameters (see [4,6–8,14,17,18,21,22]). Some of those ideas have directly influenced the development of the method described in this paper.

Certainly, one of the contributions that is of the great importance for this research is the algorithm presented in [21], stated in this paper as *SM* method. An interesting approach for calculating the accelerated parameter from [21] is applied on iteration defined in Section 3.

In [13] *double direction* form of the iteration for unconstrained optimization is investigated. Derived method is called *Accelerated Double Direction Method*, in short *ADD* method. The main result of this research is the reduction in the number of iterations needed to minimize tested functions. Another concept of research interesting to examine is *double step size* form of iteration for the same problem. The iterative scheme described in this paper arrived from this idea. Derived method is denoted as *Accelerated Double Step Size method*, or shortly *ADSS* method.

In further, the following notation is used:

$$g(x) = \bigtriangledown f(x), \quad G(x) = \bigtriangledown^2 f(x), \quad g_k = \bigtriangledown f(x_k), \quad G_k = \bigtriangledown^2 f(x_k), \tag{1.4}$$

where $\bigtriangledown f(x)$ denotes the gradient of f and $\bigtriangledown^2 f(x)$ denotes the Hessian of f. Conventionally, x^T denotes the transpose of x.

The rest of the paper is organized as follows. In Section 2 a short analysis of accelerated gradient descent methods is given. Section 3 contains detailed presentation of constructing the *ADSS* method, deriving of the value of necessary multiplying factor in correctly defined approximation of the Hessian inverse – ie the acceleration parameter γ and the *ADSS* algorithm. Linear convergency of the *ADSS* method for uniformly convex functions and for strictly convex quadratics under some specific conditions, is proved in Section 4. Finally, numerical results for the large number of parameters in test functions and comparison with the *SM* method and the *ADD* method are illustrated in Section 5.

2. Backgrounds

The most common form of iterative scheme for multi variable unconstrained minimization problem (1.1) is

$$x_{k+1} = x_k + t_k d_k, \tag{2.1}$$

where t_k is the step length parameter in the kth iteration.

A sort of subclass of mentioned methods is based on the iterative principle

$$x_{k+1} = x_k - \theta_k t_k g_k. \tag{2.2}$$

This specific scheme, originally introduced in [1], has an acceleration parameter denoted by θ_k. In [21], the authors named such a type of iterative schemes as *accelerated gradient descent methods*. Generally, mentioned acceleration parameter improves the behavior of gradient descent algorithms.

Newton's method with the line search has the general form

$$x_{k+1} = x_k - t_k G_k^{-1} g_k. \tag{2.3}$$

Combining an acceleration property from (2.2) with features of Newton's method (2.3) where the inverse of Hessian is approximated by appropriately defined diagonal matrix, a new sort of accelerated gradient descent methods with better characteristics is provided.

With Quasi-Newton method, evaluation of Hessian in every iterative step is avoided in some way. As known fact, Quasi-Newton methods satisfy quasi-Newton equation:

$$S_{k+1} y_k = s_k, \tag{2.4}$$

where $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$. Here, S_k is symmetric $n \times n$ matrix and presents an approximation of the Hessian inverse. Starting from the point of so defined matrix S_k, in [22], the next iteration is defined by

$$x_{k+1} = x_k - t_k S_k g_k, \tag{2.5}$$

but with requirements that S_k is positive definite matrix not necessarily satisfying quasi-Newton equation. So defined scheme presents a modified Newton method, [22].

An idea of creating a method of the form (2.2) by constructing the acceleration parameter θ_k using the properties of matrix S_k from (2.5) is presented in [21]. The main point is in substitution of symmetric matrix S_k with diagonal scalar matrix defined by

$$S_k = \gamma_k^{-1} I, \quad \gamma_k \in R, \tag{2.6}$$

which reduces the scheme (2.5) into the accelerated gradient descent method with the line search, so called *SM* method (see [13,21]):

$$x_{k+1} = x_k - t_k \gamma_k^{-1} g_k. \tag{2.7}$$