

A New Inexact Line Search Method for Convex Optimization Problems

Aliyu Usman Moyi

Department of Mathematics
 Faculty of Science, Universiti Putra Malaysia
 43400 Serdang, Malaysia
 Email: aliyumoyik@yahoo.com

Wah June Leong

Department of Mathematics
 Faculty of Science, Universiti Putra Malaysia
 43400 Serdang, Malaysia
 Email: wjleong@science.upm.edu.my (corresponding author)

Abstract—In general one can say that line search procedure for the steplength and search direction are two important elements of a line search algorithm. The line search procedure requires much attention because of its far implications on the robustness and efficiency of the algorithm. The purpose of this paper is to propose a simple yet effective line search strategy in solving unconstrained convex optimization problems. This line search procedure does not require the evaluation of the objective function. Instead, it forces reduction in gradient norm on each direction. Hence it is suitable for problems when function evaluation is very costly. To illustrate the effectiveness of our line search procedure, we employ this procedure together with the symmetric rank one quasi-Newton update and test it against the same quasi-Newton method with the well-known Armijo line search. Numerical results on a set of standard unconstrained optimization problems showed that the proposed procedure is superior to the Armijo line search.

Index Terms—Line search algorithm, convex optimization, gradient descent, Armijo condition.

I. INTRODUCTION

This paper concerns the numerical methods for solving the following unconstrained optimization problem

$$\min f(x); x \in R^n \quad (1)$$

where $f : R^n \rightarrow R$ is assumed to be at least twice continuously differentiable and n denotes the dimension of the problem. Usually problem (1) is solved iteratively through a line search scheme;

$$x_{k+1} = x_k + \alpha_k d_k \quad (2)$$

where d_k is the search direction in which most line search algorithms required $g_k^T d_k < 0$ such that it is a descent direction [13], [15] and $\alpha_k > 0$ is the steplength. There are several approaches for which the steplength α_k can be computed. The steplength selection guarantees the global convergence of the algorithm which in turn helps to design powerful, efficient and stable algorithm in practice [14]. The steplength α_k , is selected at each iteration so that we can find the global solution of the minimization problem:

$$\min_{\alpha \in R} f(x_k + \alpha d_k) \quad (3)$$

This type of line search is known as the exact line search, which many researchers believe that it is time-consuming to

be carried out or impossible to find in practical computation. Therefore we opt to use the inexact line searches to identify the steplength that will ensure a substantial reduction in f at minimum cost [13]. Among the most frequently used inexact line searches methods as pointed out in [11] is the Armijo line search [2] which assume $\alpha_k = \rho^z \bar{\alpha}$ for which

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \rho^z \bar{\alpha} g_k^T d_k \quad (4)$$

where $0 < \rho \leq 1$ and $\bar{\alpha} > 0$ are constants and z is the smallest non-negative integer such that (4) holds.

Various choices on the search directions contribute to different line search methods. Among them is the quasi-Newton method, which takes the search direction as

$$d_k = -B_k^{-1} g_k, \quad (5)$$

where the matrix B_k is some approximation to the Hessian matrix at x_k . Particularly, we consider the SR1 update for the Hessian approximation,

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{s_k^T (y_k - B_k s_k)}, \quad (6)$$

where $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$.

Throughout if $H = B^{-1}$, the inverse update respected to SR1 is given by

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{y_k^T (s_k - H_k y_k)}. \quad (7)$$

The SR1 update makes a symmetric rank-one change to the previous Hessian approximation B_k . Compared with other secant updates, the SR1 update is simpler and may require less computation per iteration when its unfactored forms are used (factored updates are those in which a Cholesky factor of B_k is updated at each iteration)[13]. Although SR1 update possesses desirable features, it has some major drawbacks, namely it does not necessarily generates a positive definite matrix even if the current update is positive definite, and the denominator in the SR1 update may become zero. Despite these shortcomings, we are interested in the SR1 formula because a simple safeguards seems to adequately prevent the breakdown of the update as well the loss of positive definiteness, see [8]-[10], [5]. The SR1 update has been a focus of interest particularly after the influential work of

[3], [12] thus providing a spirit of encouragement to many researchers to seek some modified schemes of update which possess not only the features of original SR1 update but also good numerical stability.

Our primary concern in this paper is to find a simple yet effective line search method for the implementation of symmetric rank-one update in solving unconstrained optimization problems. The line search procedure that we are considering does not require the evaluation of the objective function. Instead, it forces reduction in gradient norm on each direction and thus, it would be a good alternative when function evaluation is very costly. The rest of the paper is organized as follows: In section 2 we introduce the algorithm for our new inexact line search method, and in section 3 we present the numerical results of the proposed method and some discussions and we conclude the paper in section 4.

II. INEXACT LINE SEARCH METHOD THAT GUARANTEES DESCENT IN GRADIENT NORM

Since, for line search methods of the form (2) we cannot guarantee that the method converges to the solution x^* of (1) in most practical situation. The best type of global convergence result that can be obtained is that the sequence $\{x_k\}$ is attracted by stationary points, i.e. to prove the following limit instead

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (8)$$

This limit guarantees at least that x_k is close to the set of the stationary points where $g(x) = 0$. Thus, it is reasonable to introduce a strategy that ensures monotonically decrease in gradient norm on each iteration of the line search algorithm. Therefore, we propose a line search procedure as follows: Find an acceptable steplength α_k , in which the following condition holds:

$$\|g(x_k + \alpha_k d_k)\| \leq \sigma \|g(x_k)\|, \quad (9)$$

where $\sigma \in (0, 1)$.

A natural question arises at this point would be on whether such steplength exists or not. The answer of the above question can be delineated by the following lemma.

Lemma 2.1

Suppose that the objective function $f(x)$ is twice continuously differentiable and the gradient $g(x)$ is Lipschitz continuous, that is there exists a positive constant L such that, for all $x, y \in R^n$,

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad (10)$$

where $\|\cdot\|$ denotes the Euclidean norm. Then, for all steplength α of (2) that satisfies (9) will also satisfy

$$\alpha_k \geq \frac{(1 - \sigma)\|g_k\|}{L\|d_k\|}. \quad (11)$$

Proof.

Since the gradient $g(x)$ is Lipschitz continuous for all $x \in R^n$,

then

$$\begin{aligned} (1 - \sigma)\|g_k\| \leq \|g_k\| - \|g_{k+1}\| &\leq \|g(x_k) - g(x_{k+1})\| \\ &\leq L\|x_{k+1} - x_k\| \\ &= \alpha_k L\|d_k\|. \end{aligned}$$

This immediately lead to (11).

In what follows, we give the description of our new gradient descent symmetric rank one algorithm as:

Algorithm 2.1 GD

- S1: Given an initial point x_0 , and an initial positive definite matrix $H_0 = I$, set $k = 0$.
- S2: If the convergence criterion $\|g_k\| \leq \varepsilon$ is achieved, then stop.
- S3: Compute a quasi-Newton direction $d_k = -H_k g_k$, where H_k is given by (7).
- S4: Find α_k , such that at each iteration the gradient descent line search condition (9) holds, where $\sigma \in (0, 1)$.
- S5: Set $x_{k+1} = x_k + \alpha_k d_k$.
- S6: Update $H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{y_k^T (s_k - H_k y_k)}$ whenever $g_k^T d_k < 0$ and $|y_k^T (s_k - H_k y_k)| \geq r \|y_k\| \|s_k - H_k y_k\|$, where $r \in (0, 1)$ is a preset constant. Otherwise, set $H_{k+1} = H_k$.
- S7: Set $k := k + 1$, and go to S2.

III. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we present some numerical results from an implementation of our new algorithm for solving unconstrained optimization problems, we evaluate and compare the performance of our new line search algorithm with the Armijo line search for the symmetric rank-one update restart with identity matrix [7] and restart by skipping non-positive definite updates [6]. All the experiments were implemented on a PC using MATLAB 7.0, with double precision arithmetic. For each test function, we performed six numerical experiments with variable dimensions $20 \leq n \leq 1000$. As regards the stopping criteria used in our experiments, in all the algorithms, convergence is assumed if

$$\|g_k\| \leq \varepsilon, \quad \text{where } \varepsilon = 10^{-4}.$$

We forced the algorithm to stop whenever the number of iterations exceeds 2000, and the symbol "—" is used to represent the failure. Test functions are the standard unconstrained optimization problems obtained from [1].

Also in order to compare the performance between these line search methods we use the Dolan and Morè performance profile proposed in [4]. Figure 1-2 show the performance profile plots for these line search methods in terms of number of iterations and function evaluations respectively for the symmetric rank-one method. In overall, the results showed

that the gradient descent line search (GD) method is the best performer corresponding to the top curve on the plot.

Furthermore, from the numerical experiments, a total of 576 runs were performed. Based on the results, SR1 method using the Armijo line search solves 80% of the test problems while the SR1 implemented with gradient descent line search method solves 94% of the test problems considered, which clearly shows that gradient descent line search strategy exhibits a superior numerical performance in comparison with the Armijo line search algorithm.

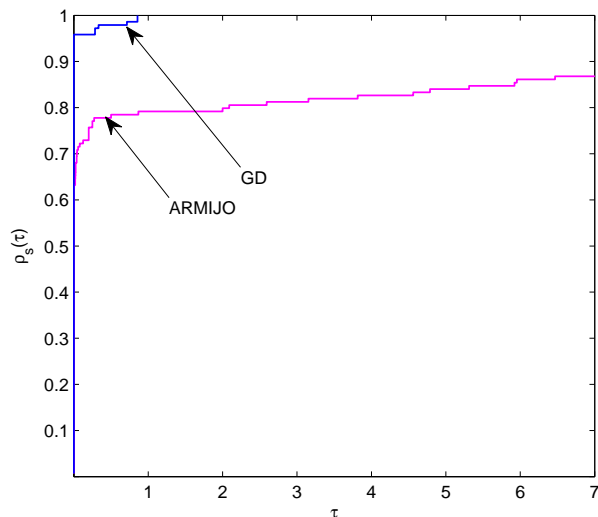


Fig. 1. Performance profiles based on iteration calls

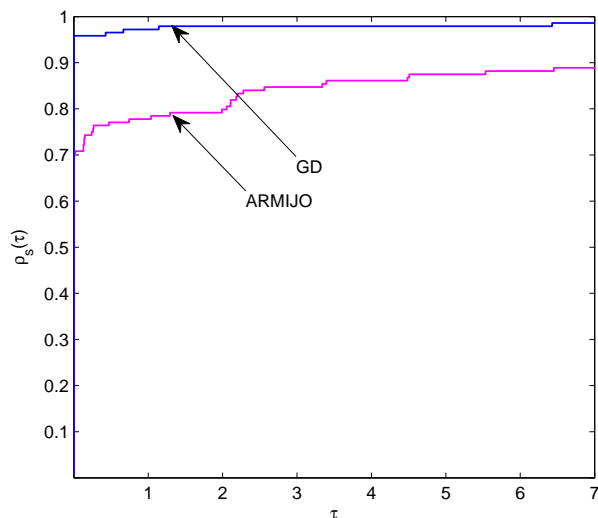


Fig. 2. Performance profiles based on function/gradient evaluations

IV. CONCLUSION

In this paper, we have presented a new line search strategy for the solution implementation of the SR1 updating scheme. Numerical testing provides strong indication that in all the

tested problems, our new inexact line search strategy performs significantly better both in terms of function/gradient evaluations and iterations than the Armijo line search method, Hence we can conclude that, a simple yet effective line search strategy can be derived to improve the performance of the line search algorithm.

ACKNOWLEDGMENT

The authors would like to thank University Putra Malaysia for providing travel fund to attend the 2013 International Conference on Applied Mathematics and Computational Methods

REFERENCES

- [1] N. Andrei, *An unconstrained optimization test functions collection*, Adv. Model. Optim. 10 (2008) 147-161
- [2] L. Armijo, *Minimization of functions having Lipschitz continuous first partial derivatives*, Pacific J. Math. 16 (1966) 13.
- [3] A. R. Conn, N. I. M. Gould and Ph. Toint, *Convergence of quasi-Newton matrices generated by the symmetric rank one update*, Math. program. 48 (1991) 549-560.
- [4] E. D. Dolan and J. J. Morè, *Benchmarking optimization software with performance profiles*, Math. Program. 91 (2) (2002) 201-213.
- [5] M. Farid and W. J. Leong, *An improved multi-step gradient-type method for large scale optimization*, Comput. Math. Appl. 61 (2011), 3312-3318.
- [6] H. F. H. Khalfan, R. H. Byrd and R. B. Schnabel, *A theoretical and experimental study of the symmetric rank one update*, SIAM J. Optim. 3 (1993) 1-24.
- [7] W. J. Leong and M. A. Hassan, *A restarting approach on symmetric rank one update for unconstrained optimization*, Comp. Optim. Appl. 42 (2009), 327-334.
- [8] W. J. Leong and M. A. Hassan, *A new gradient method via least change secant update*, Int. J. Comput. Math. 88 (2011), 816-828.
- [9] W. J. Leong and M. A. Hassan, *Positive-definite memoryless symmetric rank one method for large-scale unconstrained optimization*, Pac. J. Optim. 7 (2011), 29-41.
- [10] W. J. Leong, M. A. Hassan and M. W. Yusuf, *A matrix-free quasi-Newton method for solving large-scale nonlinear systems*, Comput. Math. Appl. 62 (2011), 2354-2363.
- [11] W. J. Leong and B. S. Goh, *Convergence and stability of line search methods for unconstrained optimization*, Acta Appl. Math. (2013) DOI 10.1007/s10440-012-9796-x.
- [12] F. Modarres, M. A. Hassan and W. J. Leong, *Improved Hessian approximation with modified secant equations for symmetric rank-one method*, J. Comp. Appl. Math. 235 (2011), 2423-2431.
- [13] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer series in Operation Research, New York: Springer-Verlag, 2006.
- [14] Z. J. Shi, *Convergence of line search methods for unconstrained optimizations*, Appl. Math. Comput. 157 (2004) 393405.
- [15] Z. J. Shi, *Convergence of quasi-Newton method with new inexact line search*, J. Math. Anal. Appl. 315 (2006) 120131.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US