

PAPER • OPEN ACCESS

New scaling on the gradient method

To cite this article: Zexian Liu *et al* 2019 *J. Phys.: Conf. Ser.* **1168** 052002

View the [article online](#) for updates and enhancements.

You may also like

- [Linear convergence of CQ algorithms and applications in gene regulatory network inference](#)
Jinhua Wang, Yaohua Hu, Chong Li et al.
- [Majorize–minimize linesearch for inversion methods involving barrier function optimization](#)
E Chouzenoux, S Moussaoui and J Idier
- [Stochastic EM methods with variance reduction for penalised PET reconstructions](#)
Željko Kereta, Robert Twyman, Simon Arridge et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

New scaling on the gradient method

Zexian Liu^{1,2,a}, Hongwei Liu², Chuanmei He², Jun Yang³ and Ming Li²

¹ School of Mathematics and Computer Science, Hezhou University, 542899, Hezhou, China

² School of Mathematics and Statistics, Xidian University, 710126, Xi'an, China

³ School of Mathematics and Information Science, Xianyang Normal University, 712000, Xianyang, China

^aCorresponding author: liuzexian2008@163.com

Abstract. Motivated by the idea of Barzilai and Borwein (BB) method, we introduce a concept of approximately optimal stepsize and present a new gradient method with approximately optimal stepsize for unconstrained optimization. We construct a new quadratic approximation model to generate an approximately optimal stepsize. We then use the two BB stepsizes to truncate it and treat the resulted approximately optimal stepsize as the new stepsize for gradient method. Moreover, for the nonconvex case, we also design two approximate models to generate approximately optimal stepsizes for gradient method. We analyze the convergence of the proposed method. The numerical results show that the proposed method is very promising.

1. Introduction

Consider the following unconstrained optimization problem:

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

where $f: R^n \rightarrow R$ is continuously differentiable and its gradient is denoted by g .

Throughout this paper, $f_k = f(x_k)$, $g_k = g(x_k)$

and $\|\cdot\|$ denotes the Euclidean norm.

The gradient method takes the form:

$$x_{k+1} = x_k - \alpha_k g_k,$$

where α_k is the stepsize. It is accepted widely that the steepest descent method, where the stepsize is determined by

$$\alpha_k^{SD} = \arg \min_{\alpha > 0} f(x_k - \alpha g_k), \quad (2)$$

is badly affected by ill conditioning and thus converges very slowly. Since the BB method [1] was proposed by Barzilai and Borwein in 1988, the interest for the gradient method has been renewed. The BB method is in essence a gradient method, the stepsize of which is given by

$$\alpha_k^{BB_1} = \frac{\|s_{k-1}\|^2}{s_{k-1}^T y_{k-1}} \text{ or } \alpha_k^{BB_2} = \frac{s_{k-1}^T y_{k-1}}{\|y_{k-1}\|^2},$$



where $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1}$.

Due to its simplicity, numerical efficiency and low storage, the BB method has enjoyed great developments in the past decades [2,3]. Raydan [4] extended the BB method to unconstrained optimization by incorporating the nonmonotone line search (GLL line search) [5], and the numerical results in [4] showed that the BB method was competitive to several famous conjugate gradient algorithms. Many BB-like method has been developed [6-9].

Since the stepsize α_k^{SD} is not efficient for gradient method, we introduce a class of stepsize called approximately optimal stepsize [10] for gradient method.

Definition 1 Suppose that f is continuously differentiable and $\phi_k(\alpha)$ is an approximate model of $f(x_k - \alpha g_k)$. A positive constant α^* is called **approximately optimal stepsize** associated to $\phi_k(\alpha)$ for gradient method, if α^* satisfies $\alpha^* = \arg \min_{\alpha > 0} \phi_k(\alpha)$.

It is clear that the effectiveness of approximately optimal stepsize will rely on the approximate mode $\phi_k(\alpha)$ greatly. Can one develop more suitable approximate model to generate efficient approximately optimal stepsizes? The is the purpose of the paper.

In this paper, we present a efficient gradient method with approximately optimal stepsize for unconstrained optimization. If $s_{k-1}^T y_{k-1} > 0$, we construct a new quadratic approximate model to generate an approximately optimal stepsize, and use the two well-known BB stepsize to truncate it and treat the resulted approximately optimal stepsize as a new stepsize for gradient method. Otherwise, we also design two new approximate models to derive two efficient approximately optimal stepsizes for gradient method. We also analyze the global convergence of the proposed method. We illustrate some numerical results which show that the proposed method is not only superior to the BB method and the SBB4 method [9] but also is competitive to two well-know software CG_DESCENT (5.3) [11] and CGOPT [12] for the given test problem set.

The rest of this paper is organized as follows. In Section 2, we exploit some different approximate models to generate some approximately optimal stepsizes. In Section 3, we present an efficient gradient method with approximately optimal stepsize for unconstrained optimization, and prove the convergence of the proposed method. In Section 4, we do some numerical experiments to examine the effectiveness of the proposed method. Conclusions are given in the last section.

2. Derivation of Approximately Optimal Stepsize

In this section we develop several approximate models to generate some approximately optimal stepsizes for gradient method.

Case I $s_{k-1}^T y_{k-1} > 0$

Consider the following quadratic model:

$$\phi_k^1(\alpha) = f(x_k) - \alpha \|g_k\|^2 + \frac{1}{2} \alpha^2 g_k^T B_k g_k,$$

where B_k is a symmetric and positive definite approximation to the Hessian matrix. Taking care of computational cost, storage cost and approximate effect, B_k is generated by imposing the BFGS formula [13] on a scalar matrix $\lambda_k I$, where $\lambda_k > 0$. Clearly, the scalar λ_k will have an important influence on the effectiveness of approximately optimal stepsize. Based on the idea of the BB method,

we use the combination of $\frac{1}{\alpha_k^{BB_1}} I$ and $\frac{1}{\alpha_k^{BB_2}} I$:

$$D_k = \left[(1-t_k) \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} + t_k \frac{\|y_{k-1}\|^2}{s_{k-1}^T y_{k-1}} \right] I,$$

to approximate the Hessian matrix, where $t_k \in [0, 1]$. Clearly, the choice of t_k is key to the approximate effect of the Hessian matrix. In this paper, we determine t_k as

$$t_k = \begin{cases} \frac{(\mathbf{s}_{k-1}^T \mathbf{y}_{k-1})^2}{\|\mathbf{s}_{k-1}\|^2 \|\mathbf{y}_{k-1}\|^2}, & \text{if } \|\mathbf{g}_k\|_2^2 \leq \xi_1, \\ \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{s}_{k-1}\| \|\mathbf{y}_{k-1}\|}, & \text{otherwise,} \end{cases}$$

where $\xi_1 > 0$.

It is well-known that the secant equation $B_k \mathbf{s}_{k-1} = \mathbf{y}_{k-1}$ is the core of quasi-Newton methods. However, the standard secant equation is not true for the D_k generally. It is accepted that the scalar matrix D_k is too simple to approximate the Hessian matrix very well. Therefore, in order to allow D_k to satisfy the secant equation, we update D_k by the BFGS update formula

$$B_k = D_k - \frac{D_k \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T D_k}{\mathbf{s}_{k-1}^T D_k \mathbf{s}_{k-1}} + \frac{\mathbf{y}_{k-1} \mathbf{y}_{k-1}^T}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}$$

It is clear that if $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} > 0$, then B_k symmetric and positive definite and satisfy the secant equation.

Imposing $\frac{d\phi_k^1}{d\alpha} = 0$, by the positive definiteness of B_k we obtain the approximately optimal stepsize:

$$\bar{\alpha}_k^{AOS(1)} = \frac{\|\mathbf{g}_k\|^2}{\left((1-t_k) \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{s}_{k-1}\|^2} + t_k \frac{\|\mathbf{y}_{k-1}\|^2}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}} \right) \left(\|\mathbf{g}_k\|^2 - \frac{(\mathbf{g}_k^T \mathbf{s}_{k-1})^2}{\|\mathbf{s}_{k-1}\|^2} \right) + \frac{(\mathbf{g}_k^T \mathbf{y}_{k-1})^2}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}}.$$

It is observed that the bound $[\alpha_k^{BB_2}, \alpha_k^{BB_1}]$ for the approximately optimal stepsize $\bar{\alpha}_k^{AOS(1)}$ is very preferable in the numerical experiments. Together with the success of the BB stepsizes $\alpha_k^{BB_1}$ and $\alpha_k^{BB_2}$, it is reasonable to impose the bound $[\alpha_k^{BB_2}, \alpha_k^{BB_1}]$ on the approximately optimal stepsize $\bar{\alpha}_k^{AOS(1)}$. As a result, in practice we take the truncated form of the approximately optimal stepsize $\bar{\alpha}_k^{AOS(1)}$:

$$\alpha_k^{AOS(1)} = \min \left\{ \alpha_k^{BB_1}, \max \left\{ \alpha_k^{BB_2}, \bar{\alpha}_k^{AOS(1)} \right\} \right\} \quad (3)$$

as the new stepsize for gradient method.

Case II $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \leq 0$

When $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \leq 0$, in most BB-like methods the stepsize is set simply to $\alpha_k = 10^{30}$. It is too simple to cause expensive computational cost for searching a suitable stepsize for gradient method.

Suppose for the moment that f is twice continuously differentiable, let us consider the second order Taylor expansion:

$f(x_k - \alpha \mathbf{g}_k) = f(x_k) - \alpha \mathbf{g}_k^T \mathbf{g}_k + \frac{1}{2} \alpha^2 \mathbf{g}_k^T \nabla^2 f(x_k) \mathbf{g}_k + o(\alpha^2)$. For a small $\tau_k > 0$, denote

$$\rho_k = \frac{|\mathbf{g}_k^T (\mathbf{g}_k - \mathbf{g}(x_k - \tau_k \mathbf{g}_k))|}{\tau_k}. \quad (4)$$

We use ρ_k to approximate $\mathbf{g}_k^T \nabla^2 f(x_k) \mathbf{g}_k$, which implies a new approximate model :

$$\phi_k^2(\alpha) = f(x_k) - \alpha \mathbf{g}_k^T \mathbf{g}_k + \frac{1}{2} \alpha^2 \rho_k.$$

Similarly, by imposing $\frac{d\phi_k^2}{d\alpha} = 0$, we obtain the approximately optimal stepsize:

$$\alpha_k^{AOS(2)} = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\rho_k}.$$

To obtain the stepsize $\alpha_k^{AOS(2)}$, it has the cost of an extra gradient evaluation, which may result in great computational cost if the gradient evaluation is evoked frequently. To avoid this case, we turn to consider \mathbf{g}_{k-1} . Since

$$\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} = -\alpha_{k-1} \mathbf{g}_{k-1}^T (\mathbf{g}_k - \mathbf{g}_{k-1}) = \alpha_{k-1} (\|\mathbf{g}_{k-1}\|^2 - \mathbf{g}_{k-1}^T \mathbf{g}_k) < 0,$$

we have that $\|\mathbf{g}_{k-1}\|^2 \leq \mathbf{g}_{k-1}^T \mathbf{g}_k$, which implies that

$$\frac{\|\mathbf{g}_{k-1}\|}{\|\mathbf{g}_k\|} \leq 1.$$

If $\frac{\|\mathbf{g}_{k-1}\|^2}{\|\mathbf{g}_k\|^2}$ is close to 1, we know that \mathbf{g}_k and \mathbf{g}_{k-1} will incline to be collinear and $\|\mathbf{g}_k\|$ and $\|\mathbf{g}_{k-1}\|$ are approximately equal. Thus, if

$$\frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2} \geq \xi_2 \text{ and } \alpha_{k-1} < \xi_3 \quad (5)$$

hold, where $\xi_2 > 0$ and $\xi_3 > 0$, we obtain that

$$\mathbf{g}_k^T \nabla^2 f(x_k) \mathbf{g}_k \approx \mathbf{g}_{k-1}^T \nabla^2 f(x_k) \mathbf{g}_{k-1} \approx \frac{[(\mathbf{g}(x_k + \alpha_{k-1} \mathbf{g}_{k-1}) - \mathbf{g}(x_k))^T \mathbf{g}_{k-1}]}{\alpha_{k-1}},$$

which also implies a new approximation model:

$$\phi_k^3(\alpha) = f(x_k) - \alpha \|\mathbf{g}_k\|^2 + \frac{1}{2} \alpha^2 \frac{[(\mathbf{g}(x_k + \alpha_{k-1} \mathbf{g}_{k-1}) - \mathbf{g}(x_k))^T \mathbf{g}_{k-1}]}{\alpha_{k-1}}.$$

Similarly, we obtain the approximate optimal stepsize:

$$\alpha_k^{AOS(3)} = \frac{\|\mathbf{g}_k\|^2}{|\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}|} \alpha_{k-1}^2.$$

When $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} = 0$, the stepsize α_k is set according to the stepsize α_{k-1} at the latest iterate. It is well-known that for a quadratic function the stepsize $\alpha_k^{BB_1}$ is exactly equal to the exact stepsize at the latest iterate, that is,

$$\alpha_k^{BB_1} = \frac{\mathbf{s}_{k-1}^T \mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}} = \alpha_{k-1}^{SD}.$$

Moreover, it also has been shown that if $\alpha_k^{BB_1}$ or α_k^{SD} is reused in a cyclic fashion, then the convergence rate is accelerated. It seems that the stepsize α_{k-1} may provide some important information for the current stepsize. Therefore, when $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} = 0$, we set the stepsize to

$$\alpha_k = \delta \alpha_{k-1}, \quad (6)$$

where $\delta > 0$. Besides, if $\rho_k = 0$, the stepsize is also computed by (6).

Therefore, when $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \leq 0$, the stepsize can be stated as follow:

$$\alpha_k = \begin{cases} \frac{\|\mathbf{g}_k\|^2}{\rho_k}, & \text{if } \rho_k \neq 0 \text{ and } \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2} < \xi_2, \\ \frac{\|\mathbf{g}_k\|^2 \alpha_{k-1}^2}{|\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}|}, & \text{if (5) hold and } \mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \neq 0, \\ \delta \alpha_{k-1}, & \text{otherwise,} \end{cases} \quad (7)$$

where ρ_k is given by (4).

3.Gradient Method with Approximate Optimal Stepsize

We present an efficient gradient method with approximately optimal stepsize for unconstrained optimization. Although GLL line search [5] was firstly incorporated into the BB method, it is observed by numerical experiments that the nonmonotone line search (Zhang-Hager line search) [14] is more preferable for BB-like methods. As a result, we adopt Zhang-Hager line search.

Algorithm 1 (GM_AOS)

Step 0 Initialization: Let $x_0 \in R^n$, $\varepsilon > 0$, $Q_0 = 1$,

$C_0 = f_0$, α_0^0 , λ_{\min} , λ_{\max} , $\sigma > 0$, ξ_1 , ξ_2 , ξ_3 , τ_k , δ , and set $k := 0$;

Step 1 If $\|\mathbf{g}_k\|_{\infty} \leq \varepsilon$, stop;

Step 2 If $k = 0$, set $\alpha = \alpha_0^0$, go to **Step 3**. If $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} > 0$, compute α_k by (3). Otherwise compute α_k by (7). Set $\alpha_k^0 = \max\{\min\{\alpha_k, \lambda_{\max}\}, \lambda_{\min}\}$ and $\alpha = \alpha_k^0$.

Step 3 Zhang-Hager line search. If

$$f(x_k - \alpha \mathbf{g}_k) \leq C_k - \sigma \alpha \|\mathbf{g}_k\|^2,$$

then go to **Step 4**. Otherwise, update α by [15]

$$\alpha = \begin{cases} \bar{\alpha}, & \text{if } \alpha > 0.1\alpha_k^0 \text{ and } \bar{\alpha} \in [0.1\alpha_k^0, 0.9\alpha], \\ 0.5\alpha, & \text{otherwise,} \end{cases} \quad \text{where } \bar{\alpha} \text{ is the trial stepsize obtained by a}$$

quadratic

interpolation at x_k and $x_k - \alpha \mathbf{g}_k$, go to **Step 3**.

Step 4 Set $\eta_k = 1$ and update Q_{k+1}, C_{k+1} by the following ways:

$Q_{k+1} = \eta_k Q_k + 1, C_{k+1} = (\eta_k Q_k C_k + f(x_{k+1})) / Q_{k+1}$. **Step 5** Set $\alpha_k = \alpha$, $x_{k+1} = x_k - \alpha_k \mathbf{g}_k$, $k = k + 1$,

and go to **Step 1**.

Now we turn to the global convergence of the GM_AOS under the following assumption:

A1 f is bounded on the set $\Omega = \{x \mid f(x) \leq f(x_0)\}$;

A2 f is bounded below on R^n ;

A3 \mathbf{g} is Lipschitz continuous, i.e. there exists a positive constant L such that

$$\|\mathbf{g}(x) - \mathbf{g}(y)\| \leq L \|x - y\|, \forall x, y \in R^n.$$

Lemma 3.1 Let $A_k = \frac{1}{k+1} \sum_{i=0}^k f_i$, we have

$$f_k \leq C_k \leq A_k.$$

Proof The proof can be found in Lemma 1.1 of [14].

Since $d_k = -g_k$, it is obvious that $g_k^T d_k = -\|g_k\|^2$ and $\|d_k\|^2 = \|g_k\|^2$ for all $k \geq 0$. By Theorem 2.2 of [14], we obtain the convergence of the GM_AOS.

Theorem 3.1 Suppose that assumptions A1, A2 and A3 hold. Then, we have $\liminf_{k \rightarrow \infty} \|g_k\| = 0$.

4. Numerical Results

In this section we do some numerical experiments to examine the effectiveness of GM_AOS. The standard test function set includes 80 unconstrained problems which are mainly from [16]. The BB method, the SBB4 method CG_DESCENT (5.3) and CGOPT are also chosen to be compared with GM_AOS, and the codes of CGOPT and CG_DESCENT (5.3) can be downloaded from http://coa.amss.ac.cn/wordpress/?page_id=21 and <http://users.clas.ufl.edu/hager/papers/Software>, respectively.

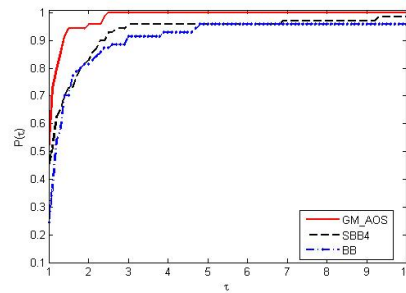


Figure 1. Performance profile based on N_{iter}

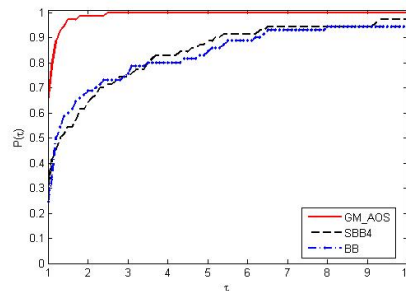


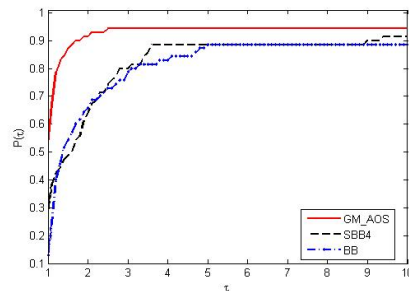
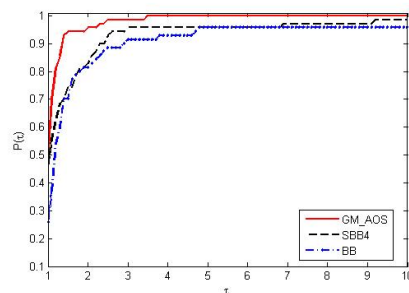
Figure 2. Performance profile based on N_f

All codes are run on PC with 3.20 GHz CPU processor, 4 GB RAM memory and Windows 7 operation system. We choose the following parameters for GM_AOS:

$n = 10000$, $\varepsilon = 10^{-6}$, $\sigma = 10^{-4}$, $\lambda_{min} = 10^{-30}$, $\lambda_{max} = 10^{30}$, $\xi_1 = 10^{-4}$, $\xi_2 = 10^{-1}$, $\xi_3 = 0.85$, $\tau_k = \min(0.1\alpha_{k-1}, 0.01)$. The BB method and the SBB4 method also use Zhang-Hager line search.

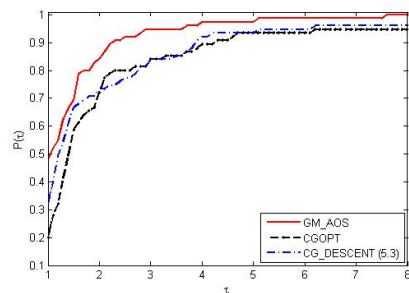
All test methods are stopped if the number of iterations exceeds 50000 or $\|g_k\|_{\infty} \leq 10^{-6}$ is satisfied, and GM_AOS, the SBB4 method and the BB method are also stopped if the number of function evaluations exceeds 80000. CG_DESCENT (5.3) and CGOPT use the default parameter values in their codes expect the above stopping conditions.

The performance profile introduced by Dolan and Moré [17] is used to display the performance of the test methods.

Figure 3. Performance profile based on N_g Figure 4. Performance profile based on T_{cpu}

We divide the numerical experiments into two groups. In Figs. 1-6, N_{iter} , N_f , N_g and T_{cpu} represent the number of iterations, the number of function evaluations, the number of gradient evaluations and CPU time (s), respectively.

In the first group of the numerical experiments, we compare GM_AOS with the BB method and the SBB4 method. GM_AOS successfully solves 79 problems, while the BB method and SBB4 method successfully solve 73 problems. As shown in Fig. 1, GM_AOS is superior to the SBB4 method and the BB method relative to the number of iterations. We see from Fig.2 that GM_AOS requires much less function evaluations than the SBB4 method and the BB method. Fig. 3 shows that

Figure 5. Performance profile based on N_{tot}

GM_AOS requires less gradient evaluations than the

BB method and the SBB4 method. In Fig. 4, we observe that GM_AOS is faster than the BB method and the SBB4 method. It indicates that GM_AOS outperforms the SBB4 method and the BB method.

In the second group of the numerical experiments, we compare GM_AOS with CGOPT and CG_DESCENT (5.3). GM_AOS successfully solves 79 problems, while CGOPT and CG_DESCENT (5.3) successfully solve 79 and 74 problems, respectively. As shown in Fig. 5, GM_AOS has some advantage over CG_DESCENT (5.3) and CGOPT relative to the total number $N_{tot} : N_{tot} = N_f + 3N_g$. In Fig. 6 we observe that GM_AOS is slightly faster than CG_DESCENT

(5.3) and CGOPT. It indicated that GM_AOS is competitive to CG_DESCENT (5.3) and CGOPT for the given test problem set.

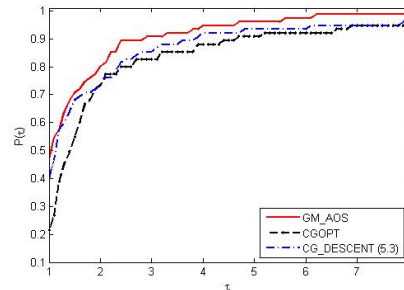


Figure 6. Performance profile based on T_{cpu}

5. Conclusion

In this paper we construct some approximate models to generate approximately optimal stepsize and present an efficient gradient method for unconstrained optimization. The numerical results indicate that GM_AOS is not only superior to the BB method and the SBB4 method but also is competitive to CGOPT and CG_DESCENT (5.3) for the given test problems set. We think that gradient methods with approximately optimal stepsizes should be paid more attention, and the gradient methods with approximately optimal stepsizes are able to become a strong candidate for large scale unconstrained optimization.

Acknowledgements

We would like to thank Dr. Huang yakui for his careful reading and helpful comments, and thank Professors Hager, W.W. and Zhang, H. C. for their C code of CG_DESCENT (5.3). This research is supported by National Science Foundation of China (No.11461021), Scientific Research Foundation of Guangxi Education Department (No.2013YB236), Scientific Research Project of Hezhou University (Nos.2014YBZK06, hzxyjxtd201502, 2016HZXYSX03).

References

- [1] J. Barzilai, J. M. Borwein. Two-point step size gradient methods. IMA J. Numer. Anal. **8**(1), 141-148 (1988).
- [2] M. Raydan. On the Barzilai and Borwein choice of steplength for the gradient method. IMA J. Numer. Anal. **13**(3), 321-326 (1993).
- [3] Y. H. Dai, L. Z. Liao. R-linear convergence of the Barzilai and Borwein gradient method. IMA J. Numer. Anal. **22**(1), 1-10 (2002).
- [4] M. Raydan. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM J. Optim. **7**(1), 26-33 (1997).
- [5] L. Grippo, S. Lucidi. A Nonmonotone Line Search Technique for Newton's Method. SIAM J. Numer. Anal. **23**(4), 707-716 (1986).
- [6] Y. H. Dai, Yuan, J. Y. Y. X. Yuan. Modified two-point stepsize gradient methods for unconstrained optimization. Comput. Optim. Appl. **22**(1), 103-109 (2002).
- [7] M. Miladinovic, P. Stanimirovic, S. Miljkovic. Scalar correction method for solving large scale unconstrained minimization problems. J. Optim. Theory Appl. **151**, 304-320 (2011).
- [8] H. W. Liu, Z. X. Liu, X. L. Dong. A new adaptive Barzilai and Borwein method for unconstrained optimization. Optim. Lett. **12**(4), 845-873 (2018).
- [9] F. Biglari, M. Solimanpur. Scaling on the spectral gradient method. J. Optim. Theory Appl. **158**(2), 626-635 (2013).
- [10] Z. X. Liu, H. W. Liu. An efficient gradient method with approximate optimal stepsize for large-scale unconstrained optimization. Numerical Algorithms, **78**(1), 21-39, 2018.

- [11] W. W. Hager, Zhang, H. C. A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J. Optim.* **16(1)**, 170-192(2005).
- [12] Y. H. Dai, C. X. Kou. A nonlinear conjugate gradient algorithm with an optimal property and an improved Wolfe line search. *SIAM J. Optim.* **23(1)**, 296-320 (2013).
- [13] Y. X. Yuan, W. Y. Sun. Theory and methods of optimization. Science Press of China, Beijing (1999)
- [14] H. C. Zhang, W. W. Hager, A nonmonotone line search technique and its application to Unconstrained Optimization. *SIAM J. Optim.* **14(4)**, 1043-1056 (2004).
- [15] E. G. Birgin. J. M. Martinez. M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM J. Optim.* **10**, 1196-1211 (2000).
- [16] N. Andrei. An unconstrained optimization test functions collection. *Adv. Model. Optim.* **10**, 147-161 (2008).
- [17] E. D. Dolan, J. J. More. Benchmarking optimization software with performance profiles. *Math. Program.* **91**, 201-213(2002).