A Positive Barzilai–Borwein-Like Stepsize and an Extension for Symmetric Linear Systems

Yu-Hong Dai, Mehiddin Al-Baali, and Xiaoqi Yang

Abstract The Barzilai and Borwein (BB) gradient method has achieved a lot of attention since it performs much more better than the classical steepest descent method. In this paper, we analyze a positive BB-like gradient stepsize and discuss its possible uses. Specifically, we present an analysis of the positive stepsize for two-dimensional strictly convex quadratic functions and prove the R-superlinear convergence under some assumption. Meanwhile, we extend BB-like methods for solving symmetric linear systems and find that a variant of the positive stepsize is very useful in the context. Some useful discussions on the positive stepsize are also given.

Keywords Unconstrained optimization • Barzilai and Borwein gradient method • Quadratic function • R-superlinear convergence • Condition number

1 Introduction

Consider the unconstrained quadratic optimization problem,

$$\min f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x}, \quad (1)$$

Y.-H. Dai (✉)
State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences, No. 55, ZhongGuanCun Donglu, P.O. Box 2719, Beijing 100190, P.R. China
e-mail: dyh@lsec.cc.ac.cn

M. Al-Baali
Department of Mathematics and Statistics, Sultan Qaboos University, Muscat, Oman
e-mail: albaali@squ.edu.om

X. Yang
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
e-mail: mayangxq@polyu.edu.hk

M. Al-Baali et al. (eds.), *Numerical Analysis and Optimization*, Springer Proceedings in Mathematics & Statistics 134, DOI 10.1007/978-3-319-17689-5_3

where $A \in R^{n\times n}$ is a real symmetric positive definite matrix and $\mathbf{b} \in R^n$. The (negative) gradient method for solving (1) takes the negative gradient as its search direction and updates the solution approximation iteratively by

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{g}_k, \tag{2}$$

where $\mathbf{g}_k = \nabla f(\mathbf{x}_k)$ and α_k is some stepsize. Denote $\mathbf{s}_{k-1} = \mathbf{x}_k - \mathbf{x}_{k-1}$ and $\mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$. Since the matrix $B_k = \alpha_k^{-1} I$, where I is the identity matrix, can be regarded as an approximation to the Hessian of f at $\mathbf{x}_k$, Barzilai and Borwein [2] choose the stepsize α_k such that B_k has a certain quasi-Newton property:

$$B_k = \arg \min_{B=\alpha^{-1}I} \|B\mathbf{s}_{k-1} - \mathbf{y}_{k-1}\|, \tag{3}$$

where $\|\cdot\|$ means the two norm, yielding the long stepsize

$$\alpha_k^{BB1} = \frac{\mathbf{s}_{k-1}^T \mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}. \tag{4}$$

An alternative way is to approximate the inverse Hessian by the matrix $H_k = \alpha_k I$ and solve

$$H_k = \arg \min_{H=\alpha I} \|\mathbf{s}_{k-1} - H\mathbf{y}_{k-1}\|, \tag{5}$$

which gives the short stepsize

$$\alpha_k^{BB2} = \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}}. \tag{6}$$

Comparing with the steepest descent (SD) method, which was due to Cauchy [4], the Barzilai–Borwein (BB) method often requires less computational work and speeds up the convergence greatly. Due to its simplicity and efficiency, the BB method has been extended or generalized in many occasions or applications. For example, Raydan [17] designed an efficient global Barzilai and Borwein algorithm for unconstrained optimization by incorporating the nonmonotone line search by Grippo et al. [15]. In the context of neural network, Dai and Liao [12] considered the one-delay method, that consists in the model

$$\frac{d\mathbf{x}(t)}{dt} = -P\nabla f(\mathbf{x}(t)), \quad t \geq 0, \tag{7}$$

where

$$P = I + \frac{\mathbf{s}\mathbf{s}^T}{\mathbf{s}^T \mathbf{y}}. \tag{8}$$

Here, $\mathbf{s} = \mathbf{x}(t - \Delta t) - \mathbf{x}(t - 2\Delta t)$, $\mathbf{y} = \nabla f(\mathbf{x}(t - \Delta t)) - \nabla f(\mathbf{x}(t - 2\Delta t))$, and Δt is the time delay. One advantage of the above model is that, if some modification is made so that the denominator in (8) is greater than zero, each eigenvalue of P will be not less than one, which makes the model not slower than the gradient neural network. The algorithm of Raydan was further generalized by Birgin et al. (2000) for minimizing a differentiable function on a closed convex set, yielding an efficient projected gradient methods. Efficient projected algorithms based on BB-like methods have also been designed (see Serafini et al. [18] and Dai and Fletcher [10]) for special quadratic programs arising from training support vector machine. The BB method has also received much attention in finding sparse approximation solutions to large underdetermined linear systems of equations from signal/image processing and statics (for example, see Wright et al. [20]).

Several attention have been paid to theoretical properties of the BB method in spite of the potential difficulties due to its heavy nonmonotone behavior. These analyses proceed in the unconstrained quadratic case (this is also the case in this paper). Specifically, Barzilai and Borwein [2] present an interesting R-superlinear convergence result for their method when the dimension is only two. For the general n-dimensional strong convex quadratic function, the BB method is also convergent (see Raydan 1993) and the convergence rate is R-linear (see Dai and Liao 2002). Further analysis on the asymptotic behavior of BB-like methods can be found in [8, 9].

One disadvantage of the BB stepsize, however, is that it may become negative for non-convex objective functions. In this case, one remedy used in [17] is to restrict the BB stepsize into some interval like $[10^{-30}, 10^{30}]$. The setting of such interval seems very artificial. The main purpose of this paper is to analyze the following positive stepsize

$$\alpha_k = \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|}. \tag{9}$$

This stepsize is exactly the geometrical mean of the long BB stepsize and the short BB stepsize. Here we should remark that the stepsize (9) has been noticed by the authors for several times (see (4.28) in [7], an unpublished preprint [9] therein, Dai and Yang [13], or Cheng and Dai [5]). This stepsize has also been noticed by Al-Baali [1]. Vrahatis et al. [19] directly replaced the Lipschitz constant L in the constant stepsize $\frac{1}{2L}$ by the estimate $\frac{\|\mathbf{y}_{k-1}\|}{\|\mathbf{s}_{k-1}\|}$, yielding a stepsize similar but not identical to (9). Nevertheless, there is no any theoretical analysis for the stepsize (9) yet.

For simplicity, we refer to the gradient method (2) with the stepsize formula (9) as method (9). In the quadratic case, since $\mathbf{s}_{k-1} = -\alpha_{k-1}\mathbf{g}_{k-1}$ and $\mathbf{y}_{k-1} = A\mathbf{s}_{k-1}$, an equivalent expression of formula (9) is

$$\alpha_k = \frac{\|\mathbf{g}_{k-1}\|}{\|A\mathbf{g}_{k-1}\|}. \tag{10}$$

Therefore formula (10) can be regarded with the one-retard extension of the stepsize considered in [13],

$$\alpha_k^{DY} = \frac{\|\mathbf{g}_k\|}{\|A\mathbf{g}_k\|}. \tag{11}$$

Interestingly enough, for the gradient method with the stepsize formula (11), it was shown in [13] that the stepsize (11) will eventually tend to the stepsize that minimizes the modulus $\|I - \alpha A\|$ (this stepsize is called the optimal stepsize in [14]). More exactly,

$$\liminf_{k\to\infty} \alpha_k^{DY} = \frac{2}{\lambda_1 + \lambda_n}, \tag{12}$$

where λ_1 and λ_n are the minimal and maximal eigenvalues of the matrix A, respectively. Simultaneously, the eigenvectors corresponding to λ_1 and λ_n can be recovered from

$$\frac{\mathbf{g}_k}{\|\mathbf{g}_k\|} + \frac{\mathbf{g}_{k+1}}{\|\mathbf{g}_{k+1}\|} \quad \text{and} \quad \frac{\mathbf{g}_k}{\|\mathbf{g}_k\|} - \frac{\mathbf{g}_{k+1}}{\|\mathbf{g}_{k+1}\|},$$

respectively.

Though simple, the two-dimensional analysis has a special meaning to the BB method. As was just mentioned, the BB method is significantly faster than the SD method in practical computations, but there is still lack of theoretical evidences that explain why the BB method is better than the SD method in the n-dimensional case. Nevertheless, the notorious zigzagging phenomenon of the SD method is well known to us; namely, the search directions in the SD method usually tend to two orthogonal directions when applied to any-dimensional quadratic functions. Unlike the SD method, however, the BB method will not produce zigzags due to its R-superlinear convergence in the two-dimensional case. This explains to some extent the efficiency of the BB method over the SD method. In this paper, we shall also analyze the convergence properties of method (9) for two-dimensional quadratic functions.

The rest of this paper is organized as follows. In the next section, we devote ourselves into the analysis of method (9) in the two-dimensional case. After giving some basic analysis in Section 2.1, we will establish the R-superlinear convergence of method (9) under some assumptions in Section 2.2. Then we make some discussions in Section 2.3. In the third section, we provide the use of the BB-like methods for solving symmetric linear systems. A typical numerical example is presented in Section 3.1, which shows that BB-like gradient methods are still very useful for solving symmetric systems. Specifically, we will see that formula (9) has a stronger ability to approximate the eigenvalues (except the signs) of a symmetric (but not necessarily positive definite) matrix A than the BB stepsizes, since formula (53) is more efficient. Some related discussions on the topic are made in Section 3.2. Finally, concluding remarks are given in the last section.

2 Analysis of Method (9) for Solving (1)

2.1 Some Basic Analysis on Method (9)

We focus on method (9) for minimizing the quadratic function (1) with $n = 2$. In this case, since the method is invariant under translations and rotations, we assume without loss of generality that

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \quad \mathbf{b} = \mathbf{0}, \tag{13}$$

where $\lambda \geq 1$. Assume that $\mathbf{x}_1$ and $\mathbf{x}_2$ are given with

$$g_1^{(i)} \neq 0, \qquad g_2^{(i)} \neq 0, \quad \text{for } i = 1 \text{ and } 2. \tag{14}$$

To analyze $\|\mathbf{g}_k\|$ for all $k \geq 3$, we denote $\mathbf{g}_k = (g_k^{(1)}, g_k^{(2)})^T$ and define

$$q_k = \frac{\left(g_k^{(1)}\right)^2}{\left(g_k^{(2)}\right)^2}. \tag{15}$$

Then it follows that

$$\|\mathbf{g}_k\|^2 = \left(g_k^{(2)}\right)^2 \left(1 + q_k\right),$$

$$\alpha_k = \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|} = \frac{\|\mathbf{g}_{k-1}\|}{\|A\mathbf{g}_{k-1}\|} = \frac{\sqrt{1 + q_{k-1}}}{\sqrt{\lambda^2 + q_{k-1}}}.$$

Noticing that $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{g}_k$ and $\mathbf{g}_k = A\mathbf{x}_k$, we have that

$$\mathbf{g}_{k+1} = (I - \alpha_k A)\mathbf{g}_k. \tag{16}$$

Writing the above relation in componentwise form,

$$\begin{pmatrix} g_{k+1}^{(1)} \\ g_{k+1}^{(2)} \end{pmatrix} = \left(\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} - \frac{\sqrt{1 + q_{k-1}}}{\sqrt{\lambda^2 + q_{k-1}}} \begin{bmatrix} 1 & \\ & \lambda \end{bmatrix} \right) \begin{pmatrix} g_k^{(1)} \\ g_k^{(2)} \end{pmatrix}$$

$$= \begin{bmatrix} \frac{\sqrt{\lambda^2+q_{k-1}} - \sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} & \\ & \frac{\sqrt{\lambda^2+q_{k-1}} - \lambda\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} \end{bmatrix} \begin{pmatrix} g_k^{(1)} \\ g_k^{(2)} \end{pmatrix}.$$

Therefore we get for all $k \geq 2$,

$$\begin{cases} \left(g_{k+1}^{(1)}\right)^2 = \dfrac{(\sqrt{\lambda^2+q_{k-1}}-\sqrt{1+q_{k-1}})^2}{\lambda^2+q_{k-1}}\left(g_k^{(1)}\right)^2, \\ \left(g_{k+1}^{(2)}\right)^2 = \dfrac{(\sqrt{\lambda^2+q_{k-1}}-\lambda\sqrt{1+q_{k-1}})^2}{\lambda^2+q_{k-1}}\left(g_k^{(2)}\right)^2. \end{cases} \tag{17}$$

In the case that $\lambda = 1$, which means that the object function has sphere contours, the method will take a unit stepsize $\alpha_2 = 1$ and give the exact solution at the third iteration. If $g_2^{(1)} = 0$ but $g_2^{(2)} \neq 0$, we have that $q_2 = 0$ and hence by (17) that $g_k^{(1)} = 0$ for $k \geq 3$ and $g_4^{(2)} = 0$, which means that the method gives the exact solution in at most four iterations. This is also true if $g_2^{(2)} = 0$ but $g_2^{(1)} \neq 0$ due to symmetry of the first and second components. If $g_1^{(1)} = 0$ but $g_1^{(2)} \neq 0$, we have that $q_1 = 0$ and $g_3^{(2)} = 0$. Then by considering $\mathbf{x}_2$ and $\mathbf{x}_3$ as two starting points, we must have $\mathbf{g}_k = 0$ for some $k \leq 5$. The symmetry works for the case that $g_1^{(2)} = 0$ but $g_1^{(1)} \neq 0$. Thus, similarly to the analysis for the BB method in [8], we may assume that $\lambda > 1$ and the assumption (14) holds, for otherwise the method has the finite termination property. On the other hand, if (14) holds, then we will have $g_k^{(1)} \neq 0$ and $g_k^{(2)} \neq 0$ for all $k \geq 1$ and hence q_k is always well defined.

Now, substituting (17) into the definition of q_{k+1}, we can obtain the following recurrence relation

$$\begin{aligned} q_{k+1} &= \left(\frac{\sqrt{\lambda^2+q_{k-1}}-\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}-\lambda\sqrt{1+q_{k-1}}}\right)^2 q_k \\ &= \left(\frac{(\sqrt{\lambda^2+q_{k-1}}-\sqrt{1+q_{k-1}})(\sqrt{\lambda^2+q_{k-1}}+\lambda\sqrt{1+q_{k-1}})}{(\lambda^2-1)q_{k-1}}\right)^2 q_k \\ &= \left(\frac{\lambda-q_{k-1}+\sqrt{\tau(q_{k-1})}}{\lambda+1}\right)^2 \frac{q_k}{q_{k-1}^2}, \end{aligned} \tag{18}$$

where τ is the following quadratic function

$$\tau(w) = (1+w)(\lambda^2+w), \quad \text{where } w \geq 0. \tag{19}$$

To proceed with our analysis, we denote $M_k = \log q_k$ and

$$h(w) = \frac{\lambda - w + \sqrt{\tau(w)}}{\lambda+1}, \quad \text{where } w \geq 0. \tag{20}$$

It follows from the recurrence relation (18) that

$$M_{k+1} = M_k - 2M_{k-1} + 2\log h(q_{k-1}). \tag{21}$$

2.2 R-Superlinear Convergence of Method (9)

Lemma 1.2.1 *Assume that $\lambda > 1$. The function $h(w)$ in (20) is monotonically increasing for $w \in [0, +\infty)$. Further, we have that*

$$h(w) \in \left[\frac{2\lambda}{\lambda+1}, \frac{\lambda+1}{2}\right), \quad \textit{for any } w \geq 0. \tag{22}$$

Proof. By the definition of τ in (19), we have that

$$(\tau')^2 - 4\tau = (\lambda^2 - 1)^2. \tag{23}$$

Then by direct calculations, we get

$$\begin{aligned} h'(w) &= \frac{-1 + \frac{1}{2}\tau^{-\frac{1}{2}}\tau'}{\lambda+1} \\ &= \frac{(\tau')^2 - 4\tau}{2(\lambda+1)\tau^{\frac{1}{2}}(\tau' + 2\tau^{\frac{1}{2}})} \\ &= \frac{(\lambda^2-1)^2}{2(\lambda+1)(\tau^{\frac{1}{2}}\tau' + 2\tau)}. \end{aligned} \tag{24}$$

Thus we see that $h'(w) > 0$ for $w \geq 0$, which indicates that $h(w)$ is monotonically increasing. Noticing that

$$h(0) = \frac{2\lambda}{\lambda+1} \quad \text{and} \quad \lim_{w\to\infty} h(w) = \frac{\lambda+1}{2},$$

we know that (22) holds. This completes our proof. □

Lemma 1.2.2 *Assume that $\lambda > 1$. Consider the function*

$$\psi(w) = \frac{wh'(w)}{h(w)}, \quad \textit{where } w \geq 0, \tag{25}$$

where $h(w)$ is given in (20). Then $\psi(w) \geq 0$ for all $w \geq 0$. Further, it reaches its maximal value at $w_{\max} = \lambda$ and

$$\psi_{\max} := \psi(w_{\max}) = \frac{1}{2} - \frac{\sqrt{\lambda}}{\lambda+1}. \tag{26}$$

Proof. The nonnegativity of $\psi(w)$ over $[0, +\infty)$ is obvious due to Lemma 1.2.1. To analyze the maximal value of $\psi(w)$ for $w > 0$, by setting $\psi'(w) = 0$ and noting that

$h'(w) \neq 0$ for $\lambda \neq 1$, we can get that

$$\frac{1}{w} - \frac{h'(w)}{h(w)} = -\frac{h''(w)}{h'(w)}. \tag{27}$$

Direct calculations show that

$$\begin{aligned}\frac{1}{w} - \frac{h'(w)}{h(w)} &= \frac{1}{w} - \frac{-1 + \frac{1}{2}\tau^{-\frac{1}{2}}\tau'}{\lambda - w + \tau^{\frac{1}{2}}} = \frac{\lambda + \tau^{\frac{1}{2}} - \frac{1}{2}w\tau^{-\frac{1}{2}}\tau'}{w(\lambda - w + \tau^{\frac{1}{2}})} \\ &= \frac{\lambda\tau + \tau^{\frac{1}{2}}(\tau - \frac{1}{2}w\tau')}{w\tau(\lambda - w + \tau^{\frac{1}{2}})} = \frac{2\lambda\tau + \tau^{\frac{1}{2}}\left[2\lambda^2 + (\lambda^2+1)w\right]}{2w\tau(\lambda - w + \tau^{\frac{1}{2}})}. \end{aligned} \tag{28}$$

On the other hand, noticing that $\tau'' = 2$, we have by (23) and direct calculations that

$$(\lambda+1)h''(w) = -\frac{1}{4}\tau^{-\frac{3}{2}}\left[(\tau')^2 - 2\tau\tau''\right] = -\frac{1}{4}(\lambda^2-1)^2\tau^{-\frac{3}{2}}.$$

The above relation indicates that $h(w)$ is a concave function. It follows from this relation and (24) that

$$-\frac{h''(w)}{h'(w)} = \tau^{-\frac{1}{2}}(1 + \frac{1}{2}\tau^{-\frac{1}{2}}\tau'). \tag{29}$$

Substituting (28) and (29) into the Equation (27) and noticing that $\tau' = 1 + \lambda^2 + 2w$, we can get

$$2\lambda\tau + \tau^{\frac{1}{2}}\left[2\lambda^2 + (\lambda^2+1)w\right] = w\left(\lambda - w + \tau^{\frac{1}{2}}\right)\left(1 + \lambda^2 + 2w + 2\tau^{\frac{1}{2}}\right). \tag{30}$$

The relation (30) is equivalent to

$$(\lambda - w)\left[-w(1+\lambda^2+2w) + 2\lambda\tau^{\frac{1}{2}} + 2\tau\right] = 0.$$

Substituting $\tau = (1+w)(\lambda^2+w)$ into the above relation yields

$$(\lambda - w)\left[2\lambda^2 + (1+\lambda^2)w + 2\lambda\tau^{\frac{1}{2}}\right] = 0. \tag{31}$$

Thus, to meet (27), which is equivalent to (31), we must have that $w = \lambda$. Since $\psi(0) = 0$ and $\psi(w) > 0$ for $w > 0$, we know that $\psi(w)$ must reach its maximal value at its unique stationary point $w = \lambda$. Therefore $w_{\max} = \lambda$. Noticing that

$$h(\lambda) = \sqrt{\lambda} \quad \text{and} \quad h'(\lambda) = \frac{(\sqrt{\lambda}-1)^2}{2(\lambda+1)\sqrt{\lambda}},$$

we can deduce (26). □

In addition to the function $\psi(w)$ in (25), we consider the function

$$\phi(w) = \begin{cases} \dfrac{\log h(w) - \log h(1)}{\log w}, & \text{if } w > 0 \text{ but } w \neq 1; \\ \dfrac{h'(1)}{h(1)}, & \text{if } w = 1. \end{cases} \tag{32}$$

Lemma 1.2.3 *For the function $\phi(w)$ defined in (32), we have that*

$$0 < \phi(w) \leq \psi_{\max}, \quad \textit{for all } w > 0,$$

where $\psi_{\max}$ is given by (26).

Proof. It is obvious that ϕ is continuous in $(0, +\infty)$ and continuously differentiable over $(0,1) \cup (1,+\infty)$. Due to Lemma 1.2.1, we can also see that $\phi(w)$ tends to zero when w tends to zero or when w tends to $+\infty$. Further, by setting the derivative of $\phi(w)$ to be zero, we know that the optimal w^* that maximize $\phi(w)$ over $(0,1) \cup (1,+\infty)$ must satisfy

$$\phi(w^*) = \frac{w^* h'(w^*)}{h(w^*)}.$$

Consequently, by Lemma 1.2.2, we have for $w > 0$,

$$0 < \phi(w) \leq \max\{\phi(w^*), \phi(1)\} = \max\{\psi(w^*), \psi(1)\} \leq \psi_{\max}.$$

This completes our proof. □

Now, noticing the relation (21) and using the definition of ϕ, we can get that

$$M_{k+1} = M_k - 2(1 - \phi(q_{k-1}))M_{k-1} + 2\log h(1). \tag{33}$$

By Lemma 1.2.3, we know that the coefficient of M_{k-1} belongs to the interval

$$\left(-2, -1 - \frac{2\sqrt{\lambda}}{\lambda+1}\right]. \tag{34}$$

This interval, however, cannot enable us to find some suitable parameter γ such that the sequence of $|M_k + \gamma M_{k-1}|$ is monotonically increasing with k. To do so, we have to strengthen the upper bound of $\phi(w)$ in Lemma 1.2.3. Meanwhile, we still need some suitable assumption on the initial value of M_1 and M_2 similarly to Lemma 1.2.4. Based on this reason, we directly work with the recursive relation (21). Pick γ to be any root of the equation $\gamma^2-\gamma+2=0$; namely,

$$\gamma = \frac{1 \pm \sqrt{7}i}{2}, \tag{35}$$

where i is the imaginary unit (sometimes i is also used as an index, but it is easy for the reader to tell). We have the following lemma.

Lemma 1.2.4 *Consider the sequence $\{M_k\}$ satisfying (21). Denote $\xi_k = M_k + (\gamma-1)M_{k-1}$, where γ is given in (35). If*

$$|\xi_2| > 2\log\frac{\lambda+1}{2}, \tag{36}$$

there exist some positive constants c_1 and c_2 such that

$$|\xi_k| \geq c_1 2^{k-2} + c_2, \quad \textit{for all } k \geq 2. \tag{37}$$

Proof. It follows from the definition of ξ_k, the relation (21) and the choice of γ that

$$\xi_{k+1} = \gamma M_k - 2M_{k-1} + 2\log h(q_{k-1}) = \gamma\xi_k + 2\log h(q_{k-1}).$$

Noticing that $|\gamma| = 2$ and by Lemma 1.2.1, $|\log h(q_{k-1})| < \log\frac{\lambda+1}{2}$, we have from the above relation that

$$|\xi_{k+1}| \geq 2|\xi_k| - c_2, \tag{38}$$

where $c_2 = 2\log\frac{\lambda+1}{2}$. The relation (38) is equivalent to

$$|\xi_{k+1}| - c_2 \geq 2\left(|\xi_k| - c_2\right). \tag{39}$$

Denoting $c_1 = |\xi_2| - c_2$, which is strictly greater than zero by assumption, we can know from the repeated use of (39) that (37) holds. □

Notice that $|\gamma - 1| = 2$ and hence

$$|\xi_k| \leq |M_k| + 2|M_{k-1}| \leq 3\max\{|M_k|, |M_{k-1}|\}.$$

This with (37) gives that

$$\max\{|M_k|, |M_{k-1}|\} \geq \frac{1}{3}\left(c_1 2^{k-2} + c_2\right), \quad \text{for all } k \geq 2. \tag{40}$$

Lemma 1.2.5 *Consider the sequence* $\{M_k\}$ *satisfying (21). Under the same assumption in Lemma 1.2.4, we have for all* $k \geq 2$ *that*

$$\max_{-1\leq i\leq 3} M_{k+i} \geq \frac{1}{3}c_1 2^{k-2} - 4\log\frac{\lambda+1}{2} \tag{41}$$

and

$$\min_{-1\leq i\leq 3} M_{k+i} \leq -\frac{1}{3}c_1 2^{k-2} + 4\log\frac{\lambda+1}{2}. \tag{42}$$

Proof. It follows from the recursive relation (21) that

$$M_{k+2} = -M_k - 2M_{k-1} + 2\log h(q_k) + 2\log h(q_{k-1}). \tag{43}$$

We focus on the relation (40). If there exists some $i = 0$ or 1 such that

$$M_{k-i} \geq \frac{1}{3}\left(c_1 2^{k-2} + c_2\right),$$

then it is obvious that (41) holds. Otherwise, we must have that

$$M_{k-i} \leq -\frac{1}{3}\left(c_1 2^{k-2} + c_2\right)$$

holds for some $i = 0$ or 1. In this case, noticing Lemma 1.2.1, we can see from (21) and (43) (with $k-1$ replaced with $k-i$) that the following relation

$$M_{k-i+j} \geq \frac{2}{3}\left(c_1 2^{k-2} + c_2\right) - 4\log\frac{\lambda+1}{2}$$

holds for $j = 2$ or 3. As a matter of fact, we can use the relation (21) if $M_{k-i+1} \geq 0$ or the relation (43) otherwise. Therefore (41) must be true. The proof of (42) is similar. □

The above lemma indicates that there must exist two subsequences of $\{M_k\}$ which tend to $+\infty$ and $-\infty$, respectively, at a geometrical rate. Then we are able to show that both the components of the gradient tend to zero R-superlinearly and hence the whole gradient norm is R-superlinearly convergent.

Theorem 1.2.6 *Consider method (9). Assume that (14) and (36) hold. Then the sequence of the gradient norm* $\{\|\mathbf{g}_k\|\}$ *converges to zero and the convergence is* R*-superlinear.*

Proof. First, noticing that $\alpha_k \in (\lambda^{-1}, 1)$ for any k, we know from (16) that

$$|g_{k+1}^{(i)}| \leq (\lambda-1)|g_k^{(i)}| \tag{44}$$

holds for $i = 1$ and 2 and all $k \geq 1$. Let us focus on the second component of $\mathbf{g}_k$. By the second relation in (17), it is not difficult to prove that

$$\begin{aligned} |g_{k+1}^{(2)}| &\leq \frac{(\lambda^2-1)q_{k-1}}{\sqrt{\lambda^2+q_{k-1}}(\sqrt{\lambda^2+q_{k-1}}+\lambda\sqrt{1+q_{k-1}})}|g_k^{(2)}| \\ &\leq \frac{(\lambda^2-1)q_{k-1}}{2\lambda^2}|g_k^{(2)}| \\ &< (\lambda-1)q_{k-1}|g_k^{(2)}|. \end{aligned} \tag{45}$$

Combining (44) and (45), we can get that

$$|g_{k+5}^{(2)}| \leq (\lambda-1)^5 \left(\min_{-1\leq i\leq 3} q_{k-1}\right) |g_k^{(2)}|,$$

which, with $M_k = \log q_k$ and the relation (42), yields

$$|g_{k+5}^{(2)}| \leq (\lambda-1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 4\log\frac{\lambda+1}{2}\right) |g_k^{(2)}|.$$

Similarly, we can build

$$|g_{k+5}^{(1)}| \leq \frac{1}{2}(\lambda+1)(\lambda-1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 4\log\frac{\lambda+1}{2}\right) |g_k^{(1)}|.$$

Thus we can obtain for all k,

$$||\mathbf{g}_{k+5}|| \leq \frac{1}{2}(\lambda+1)(\lambda-1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 4\log\frac{\lambda+1}{2}\right) ||\mathbf{g}_k||. \tag{46}$$

Therefore we can see that $\|\mathbf{g}_k\|$ converges to zero and the convergence is R-superlinear. □

2.3 Some Discussions

Comparing the above two-dimensional analysis for method (9) with those for the BB method, we can see that the analysis in this paper is more difficult. The current analysis requires an assumption on the initial points, that is (36), so that we can prove the divergence of a subsequence of $\{M_k\}$ (see Lemma 1.2.4). Then we are able to show that there are two subsequences of $\{M_k\}$ which tend to $+\infty$ and $-\infty$, respectively (see Lemma 1.2.5). A direct implication of this result is that there are two subsequences of $\{\alpha_k\}$ which converges to the two eigenvalues of the matrix A. Finally, we can establish the R-superlinear convergence for method (9) in the two-dimensional case. Although our numerical observations show that the assumption (36) is not necessary, we do not know yet whether this assumption can be removed or not.

Since by Lemma 1.2.1, the last term in the recursive relation (21) is bounded above and below, we may think that the properties of the sequence $\{M_k\}$ are similar to the one satisfying the linear recursion relation $M_{k+1} = M_k - 2M_{k-1}$. The latter is exactly what Barzilai and Borwein [2] obtained for the BB method. Therefore, we might feel that method (9) itself performs not better than the BB method. An illustrative example is as following. Consider the 1000-dimensional example

$$A = diag(1:1000), \qquad \mathbf{b} = zeros(1000,1). \tag{47}$$

Here and below *diag* and *zeros* are standard matlab languages. The starting point and the stopping criterion are

$$\mathbf{x}_1 = ones(1000), \qquad \|\mathbf{g}_k\| \le 10^{-12}, \tag{48}$$

respectively. It was found that, to reach the stopping criterion, the BB1 method, the BB2 method, and method (9) require 590, 697, and 1,139, iterations, respectively.

Nevertheless, when applied the BB method for general nonconvex optimization, it is possible that $\mathbf{s}_{k-1}^T\mathbf{y}_{k-1} < 0$, in which case some truncations are often done. For example, by projecting the BB stepsizes onto the interval like $[10^{-30}, 10^{30}]$. With the help of the stepsize (9), we may now consider, for example, the following possibilities

$$\bar{\alpha}_k^{BB1} = \max\left\{\frac{\mathbf{s}_{k-1}^T\mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T\mathbf{y}_{k-1}}, \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|}\right\} \tag{49}$$

and

$$\bar{\alpha}_k^{BB2} = \max\left\{\frac{\mathbf{s}_{k-1}^T\mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T\mathbf{y}_{k-1}}, \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|}\right\}. \tag{50}$$

It is easy to see that if $\mathbf{s}_{k-1}^T\mathbf{y}_{k-1} > 0$, $\bar{\alpha}_k^{BB1} = \alpha_k^{BB1}$ and $\bar{\alpha}_k^{BB2}$ reduces to the stepsize (9). However, if $\mathbf{s}_{k-1}^T\mathbf{y}_{k-1} \le 0$, the stepsize (9) will be used instead. Theoretically, by the analysis in [7], it is not difficult to see that all the stepsizes (9), (49), and (50) possess the so-called Property (A) and hence the corresponding gradient methods are R-linearly convergent for any-dimensional strictly convex quadratic functions. More numerical experiments are still required to test the efficiency of the proposed variants.

3 Solving Symmetric Linear Systems

This section aims to expose another good property of the stepsize (9). More exactly, if the Hessian matrix A is only symmetric, but not necessarily positive definite, we will find that formula (9) has stronger ability to approximate the eigenvalues (except the signs) of A than the formulae (4) and (6).

3.1 A Typical Numerical Example

In this section, we shall consider the symmetric linear system

$$A\mathbf{x} = \mathbf{b}, \tag{51}$$

where $A \in R^{n\times n}$ is assumed to be symmetric and invertible and $\mathbf{b} \in R^n$. It is obvious that if A is symmetric positive definite, the unconstrained quadratic optimization problem (1) is equivalent to the linear system (51). In this subsection, however, we only assume A to be symmetric, but not necessarily positive definite. As BB-like gradient methods have achieved great success in various aspects, there seem not many studies on the methods for solving symmetric linear systems.

For easy illustration, for any dimension n, we define the n-dimensional vector $\mathbf{v}$ with $v(i) = (-1)^i i$ and consider the following example

$$A = diag(\mathbf{v}), \qquad \mathbf{b} = zeros(n,1).$$

Here again, *diag* and *zeros* are standard matlab languages. In the context of linear systems, we define $\mathbf{g}_k = A\mathbf{x}_k - \mathbf{b}$, which is exactly the derivative of the quadratic function in (1). The starting point and the stopping criterion are

$$\mathbf{x}_1 = ones(n) \quad \text{and} \quad \|\mathbf{g}_k\| \le 10^{-6}, \tag{52}$$

respectively. In practical computations, we consider the following five values of n: $n = 10, 20, 30, 40, 50$.

Firstly, we tried the naive use of the classical steepest descent method, that is, the method (2) with the stepsize

$$\alpha_k^{SD} = \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{g}_k^T A\mathbf{g}_k},$$

and found that the norm $\|\mathbf{g}_k\|$ goes to infinity at a fast rate and cannot converge at all.

Secondly, we tested the two choices, (4) and (6), of the Barzilai–Borwein methods. They are denoted by BB1 and BB2, respectively. In this case, the steepest descent stepsize is used for the first iteration. In Table 1, we listed the number of

Table 1 Comparing different methods for symmetric linear systems

	n				
Method	10	20	30	40	50
BB1	1,117	2,806	2,568	2,948	4,685
BB2	238	499	1,138	2,104	2,345
(53)	147	426	607	687	847

iterations required by each method for each problem. It is remarkable to see that both BB1 and BB2 can provide a solution satisfying the stopping criterion in (52). Further, unlike the unconstrained optimization, where it is believed that BB1 is preferable to BB2, the BB2 method requires significantly fewer iterations than the BB1 method does.

Now, we think of how to make use of the stepsize (9) for solving the symmetric linear system (51). Due to its equivalent definition (10) of the stepsize, it is easy to see that α_k^2 is an approximation to some inverse eigenvalue of the matrix A^2. To decrease the components of the residual vector $\mathbf{g}_k$ corresponding to the negative eigenvalues of A, we need to design a mechanism how to choose the sign of the stepsize α_k. An easy way is to consider the function

$$sign(a) = \begin{cases} 1, & \text{if } a \geq 0; \\ -1, & \text{otherwise.} \end{cases}$$

Then we calculate the stepsize in the following way

$$\alpha_k = sign(\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}) \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|}. \tag{53}$$

In other words, the stepsize (53) aims to approximate the inverse eigenvalue of the matrix A based on the sign of the inner product $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}$. If $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}$ is greater than or equal to zero, it tends to estimate the inverse of the positive eigenvalues; otherwise, if $\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}$ is less than zero, it goes to approximate the inverse of the negative eigenvalues. The iterations required by the method (53) are denoted in Table 1 in the row "(53)". Again, the steepest descent stepsize is used for the first iteration. From the table, we can see that the new method performs much efficient than the BB1 and BB2 methods.

3.2 Some Discussions

It is obvious that more numerical experiments with symmetric linear systems are needed to check the efficiency of the new method. Nevertheless, the above example is typical, which explains that the new method performs much better than the BB1 and BB2 methods. The example also provides some reason of directly using the BB1 and BB2 stepsizes in the context of optimization, instead of truncating them to be some tiny positive numbers like $\alpha_{\min} = 10^{-30}$ as mentioned in the introduction. The first author once found that the direct use of the negative BB stepsizes can reduce the number of iterations.

As there have been a lot of Barzilai–Borwein-like gradient methods in the context of optimization, we do not know yet whether there exists more efficient stepsizes in the gradient method for solving symmetric linear systems. Another issue is the application of the new stepsize in nonlinear systems. Cruz et al. [6] built an efficient gradient algorithm for nonlinear systems based on the BB stepsizes. Can we improve their gradient algorithms by using our new stepsize?

Finally, an important theoretical question related to the new method (or the BB1 and BB2 methods) is, does the new method converge for general symmetric linear systems? Although our numerical experiments show that the answer might be yes, it seems very difficult for us to provide a proof.

4 Concluding Remarks

In this paper, we have analyzed a positive BB-like gradient stepsize and discussed its possible uses. We provide an analysis of the positive stepsize for two-dimensional strictly convex quadratic functions and prove the R-superlinear convergence under the assumption (36). It is not known yet whether the assumption (36) can be removed or not. At the same time, we have extended BB-like methods for solving symmetric linear systems and found that a variant of the positive stepsize, that is (53), is very useful in the context. More numerical experiments are required to examine the efficiency of the stepsize (53) for symmetric linear systems. The convergence of BB-like methods in this context is also not known to us in theory.

From the discussions in Sections 2.3 and 3.2, we have seen two possibilities to deal with the case where the BB stepsizes are negative. The first is by truncation [for example, see (49) and (50)]. The second is still to use the BB stepsize even when negative values of the BB stepsize have been detected due to the successful numerical example in Section 3.1. On the whole, the proposition of the positive stepsize (9) might provide much room in finding more efficient and reasonable BB-like gradient methods.

Acknowledgements The authors are grateful to Dr. Bo Jiang for checking an early version of this manuscript and to Ms. Liaoyuan Zeng for her editing of this paper. They also thank an anonymous referee for his/her useful suggestions and comments.

References

1. Al-Baali, M.: On alternate steps for gradient methods. Talk at 22-nd Biennial Conference on Numerical Analysis, University of Dundee, Scotland, 26–29 June 2007
2. Barzilai, J., Borwein, J.M.: Two-point step size gradient methods. IMA J. Numer. Anal. **8**, 141–148 (1988)

3. Birgin, E.G., Martíinez, J.M., Raydan, M.: Nonmonotone spectral projected gradient methods on convex sets. SIAM J. Optim. **10**, 1196–1211 (2000)
4. Cauchy, A.: Méthode générale pour la résolution des systèms d'équations simultanées. Comput. Rend. Sci. Paris **25**, 536–538 (1847)
5. Cheng, M., Dai, Y.H.: Adaptive nonmonotone spectral residual method for large-scale nonlinear systems. Pac. J. Optim. **8**, 15–25 (2012)
6. Cruz, W.L., Martínez, J.M., Raydan, M.: Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. Math. Comput. **75**, 1429–1448 (2006)
7. Dai, Y.H.: Alternate step gradient method. Optimization **52**, 395–415 (2003)
8. Dai, Y.H.: A new analysis on the Barzilai-Borwein gradient method. J. Oper. Res. Soc. China **1**, 187–198 (2013)
9. Dai, Y.H., Fletcher, R.: On the asymptotic behaviour of some new gradient methods. Math. Program. Ser. A **103**, 541–559 (2005)
10. Dai, Y.H., Fletcher, R.: New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. Math. Program. Ser. A **106**, 403–421 (2006)
11. Dai, Y.H., Liao, L.Z.: R-linear convergence of the Barzilai and Borwein gradient method. IMA J. Numer. Anal. **26**, 1–10 (2002)
12. Dai, Y.H., Liao, L.Z.: A new first-order neural network for unconstrained nonconvex optimization. Research Report, Academy of Mathematics and Systems Science, Chinese Academy of Sciences (1999)
13. Dai, Y.H., Yang, X.Q.: A new gradient method with an optimal stepsize property. Comput. Optim. Appl. **33**, 73–88 (2006)
14. Elman, H.C., Golub, G.H.: Inexact and preconditioning Uzawa algorithms for saddle point problems. SIAM J. Numer. Anal. **36**, 1645–1661 (1994)
15. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
16. Raydan, M.: On the Barzilai and Borwein choice of steplength for the gradient method. IMA J. Numer. Anal. **13**, 321–326 (1993)
17. Raydan, J.: The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM J. Optim. **7**, 26–33 (1997)
18. Serafini, T., Zanghirati, G., Zanni, L.: Gradient projection methods for quadratic programs and applications in training support vector machines. Optim. Methods Softw. **20**, 347–372 (2005)
19. Vrahatis, M.N., Androulakis, G.S., Lambrinos, J.N., Magoulas, G.D.: A class of gradient unconstrained minimization algorithms with adaptive stepsize. J. Comput. Appl. Math. **114**, 367–386 (2000)
20. Wright, S.J., Nowak, R.D., Figueiredo, M.A.T.: Sparse reconstruction by separable approximation. IEEE Trans. Signal Process. **57**, 2479–2493 (2009)