Convergence of conjugate gradient methods with a closed-form stepsize formula

Christian Labat [1], Jérôme Idier [2]

November 30, 2005

Abstract. Conjugate gradient methods are efficient to minimize differentiable objective functions in large dimension spaces. However, converging line search strategies are usually not easy to choose, nor to implement. In Refs. 1, 2, Sun and colleagues introduced a simple stepsize formula. However, the associated convergence domain happens to be overrestrictive, since it precludes the optimal stepsize in the convex quadratic case. Here, we identify this stepsize formula with one iteration of Weiszfeld's algorithm in the scalar case. More generally, we propose to make use of a finite number of iterates of such an algorithm to compute the stepsize. In this framework, we establish a new convergence domain, that incorporates the optimal stepsize in the convex quadratic case.

Key Words. Conjugate gradient methods, convergence, stepsize formula, Weiszfeld's method.

1 Introduction

Let us consider the following unconstrained minimization problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \mathcal{J}(\boldsymbol{x}) \tag{1}$$

where J is a differentiable objective function. In the implementation of any conjugate gradient (CG) method, the stepsize is often determined by certain line search conditions such as the Wolfe conditions (Ref. 3). These types of line search involve extensive computation of function values and gradients, which often becomes a significant burden for large-scale problems. Most recently, a simple stepsize formula was proposed by Sun and

[1]PHD Student, Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN, UMR CNRS 6597), Nantes, France.

[2]CNRS Researcher, Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN, UMR CNRS 6597), Nantes, France.

Zhang (Ref. 1) and by Chen and Sun (Ref. 2) for several CG methods. Here, we pursue in the same direction, by proposing a generalized stepsize formula. We also reexamine the convergence conditions, which leads us to a broadened convergence domain for several types of conjugacy.

In this paper, we restrict ourselves to the following family of CG algorithm:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \tag{2}$$

$$\boldsymbol{c}_k = -\boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1} \tag{3}$$

$$\boldsymbol{d}_k = \begin{cases} \boldsymbol{c}_k & \text{if } \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{c}_k \leq 0 \\ -\boldsymbol{c}_k & \text{otherwise} \end{cases} \tag{4}$$

$$\beta_k = \begin{cases} 0, & \text{for } k = 0 \\ \beta_k^{\mu_k,\omega_k}, & \text{for } k \geq 1 \end{cases}$$

where $k \in \mathbb{N}$, $\boldsymbol{g}_k = \nabla \mathcal{J}(\boldsymbol{x}_k)$ and with the following conjugacy formulas:

$$D_k = (1 - \mu_k - \omega_k)\|\boldsymbol{g}_{k-1}\|^2 + \mu_k \boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{y}_{k-1} - \omega_k \boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} \tag{5}$$

$$\beta_k^{\mu_k,\omega_k} = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1} / D_k \tag{6}$$

where $\|\,.\,\|$ is the Euclidean norm, "t" stands for the transpose, $\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$, and D_k depends on parameters $\mu_k \in [0, 1]$ and $\omega_k \in [0, 1 - \mu_k]$. Let us remark that the descent direction $\boldsymbol{d}_k$ is defined such that $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq 0$.

The parametrized expression (6) is taken from Ref. 2. It only covers a subset of a larger family introduced by Dai and Yan in Ref. 4. Three classical versions of nonlinear CG are particular cases of formula (6):

$$\beta_k^{1,0} = \beta_k^{\mathrm{HS}} = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1} / \boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{y}_{k-1} \qquad \text{Hestenes-Stiefel (Ref. 5)}$$

$$\beta_k^{0,0} = \beta_k^{\mathrm{PRP}} = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1} / \|\boldsymbol{g}_{k-1}\|^2 \qquad \text{Polak-Ribière-Polyak (Ref. 6, 7)}$$

$$\beta_k^{0,1} = \beta_k^{\mathrm{LS}} = -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1} / \boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} \qquad \text{Liu-Storey (Ref. 8)}$$

Other important cases are not covered by the present study, such as the Fletcher-Reeves method (Ref. 9), the Conjugate Descent method (Ref. 10) and the Dai-Yuan method (Ref. 11).

On the other hand, we focus on the following stepsize strategy:

$$\alpha_k = \alpha_k^1 = 0 \quad \text{if } \boldsymbol{d}_k = \boldsymbol{0}; \tag{7}$$

otherwise,

$$\begin{cases} \alpha_k^0 = 0 & \text{(8a)} \\ \alpha_k^{i+1} = \alpha_k^i - \theta \boldsymbol{d}_k^{\mathrm{t}} \nabla \mathcal{J}(\boldsymbol{x}_k + \alpha_k^i \boldsymbol{d}_k) / \boldsymbol{d}_k^{\mathrm{t}} \boldsymbol{Q}_k^i \boldsymbol{d}_k, \quad i \in \{0, \ldots, I-1\} & \text{(8b)} \\ \alpha_k = \alpha_k^I & \text{(8c)} \end{cases}$$

where $I \in \mathbb{N} - \{0\}$, $\theta \in \mathbb{R}$ is a parameter, $\{\boldsymbol{Q}_k^i\} \in \mathbb{R}^{n \times n}$ is a series of symmetric, positive definite matrices with a uniformly bounded spectrum and a strictly positive lower bound, *i.e.,* there exist $\nu_1, \nu_2 \in \mathbb{R}$ with $\nu_2 \geq \nu_1 > 0$ such that

$$\nu_1 \|\boldsymbol{v}\|^2 \leq \boldsymbol{v}^{\mathrm{t}} \boldsymbol{Q}_k^i \boldsymbol{v} \leq \nu_2 \|\boldsymbol{v}\|^2, \quad \forall k \in \mathbb{N}, \forall i \in \{0, \dots, I-1\}, \forall \boldsymbol{v} \in \mathbb{R}^n. \tag{9}$$

The fixed number of iterations of (8b) yields a family of CG methods with a *closed-form stepsize formula* (CFSF). It actually stays in the same spirit as the CG methods *without linesearch* introduced in Refs. 1, 2, since no stopping condition is involved in the stepsize scheme to ensure convergence. In the case of a single application of (8b) ($I = 1$), our stepsize formula boils down to

$$\alpha_k = \alpha_k^1 = -\theta \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \boldsymbol{d}_k^{\mathrm{t}} \boldsymbol{Q}_k^0 \boldsymbol{d}_k, \tag{10}$$

which is exactly the formula introduced in Refs. 1, 2. According to (4), note that the latter expression for α_k is nonnegative provided that $\theta > 0$.

To ensure convergence, the condition $\theta \in (0, \nu_1/\mu)$ is introduced in Refs. 1, 2, where μ is a Lipschitz constant (see Assumption 1 below). In Section 5, we show that this condition is overrestrictive, so that the stepsize formula proposed in Refs. 1, 2 produces too small steps. This becomes obvious in the convex quadratic case, since the optimal stepsize $\theta = 1$ does not belong to the interval $(0, \nu_1/\mu) \subset (0, 1)$.

In this paper, we propose relaxed convergence conditions. In particular, the optimal stepsize becomes admissible in the convex quadratic case. The key ingredient we incorporate consists in approximating J by a convex quadratic function from above, which is the basic principle of the Weiszfeld's method (Refs. 12, 13). First of all, we put forward that the stepsize formula proposed in Refs. 1, 2 identifies with one iteration of Weiszfeld's algorithm in the scalar case. More generally, our iterated version (8b) corresponds to a fixed number of the same scalar algorithm. The majorizing convex quadratic approximation framework provides altered convergence conditions compared to the conditions found in Refs. 1, 2: in particular, $\theta \in (0, \nu_1/\mu)$ is replaced by $\theta \in (0, 2)$ for any finite value of I.

The paper is organized as follows. Some preliminary results on the family of CG methods with the closed-form stepsize formula (7)-(8) are given in Section 2. We also introduce the additional assumption of a majorizing convex quadratic function that allow us to make the connection between the closed-form stepsize formula and the scalar Weiszfeld's method. Section 3 gathers some properties concerning the stepsize series generated by (8) useful for the next section. Section 4 includes the main convergence properties of the two-parameter family of CG methods defined by (2)-(8). Finally, discussions on the convex quadratic case and the general case are given in Section 5.

2 Preliminaries

Let N be a neighborhood of the level set $L = \{\boldsymbol{x} \in \mathbb{R}^n | \mathcal{J}(\boldsymbol{x}) \leq \mathcal{J}(\boldsymbol{x}_0)\}$, which is assumed bounded in the sequel. The following assumption is also adopted.

Assumption 1. *Let us assume that* $\mathcal{J} : \mathbb{R}^n \mapsto \mathbb{R}$ *is differentiable on* N*, and that* $\nabla\mathcal{J}$ *is Lipschitz continuous on* N *with the Lipschitz constant* $\mu > 0$*:*

$$\|\nabla\mathcal{J}(\boldsymbol{x}) - \nabla\mathcal{J}(\boldsymbol{x}')\| \leq \mu\|\boldsymbol{x} - \boldsymbol{x}'\|, \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in N.$$

In short, it will be said that $\mathcal{J}$ *is* μ-$\mathcal{LC}^1$.

In the sequel, Assumption 1 will appear to be sufficient for the global convergence of the CG method when $\mu_k = 0$ and $\omega_k \in [0, 1]$, which encompasses the PRP and the LS cases, but not the HS case. Thus, we consider the following stronger assumption for the more general case $\mu_k \in [0, 1]$, $\omega_k \in [0, 1 - \mu_k]$.

Assumption 2. *Let Assumption 1 hold, and let* $\mathcal{J}$ *be strongly convex on* N*: there exists* $\lambda > 0$ *such that*

$$\left[\nabla\mathcal{J}(\boldsymbol{x}) - \nabla\mathcal{J}(\boldsymbol{x}')\right]^{\mathrm{t}} (\boldsymbol{x} - \boldsymbol{x}') \geq \lambda\|\boldsymbol{x} - \boldsymbol{x}'\|^2, \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in N.$$

Note that Assumption 2 implies that L bounded since a strongly convex function has bounded level sets.

Finally, let us introduce convex quadratic majorizing functions through the following assumption.

Assumption 3. *Let*

$$\widehat{\mathcal{J}}_k^i(\boldsymbol{x}', \boldsymbol{x}) = \mathcal{J}(\boldsymbol{x}) + (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}}\, \nabla\mathcal{J}(\boldsymbol{x}) + (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}}\, \boldsymbol{Q}_k^i\, (\boldsymbol{x}' - \boldsymbol{x})/2 \tag{11}$$

where $\{\boldsymbol{Q}_k^i\}$ *is a series of positive definite matrices, be such that*

$$\widehat{\mathcal{J}}_k^i(\boldsymbol{x}', \boldsymbol{x}) \geq \mathcal{J}(\boldsymbol{x}'), \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in N, \tag{12}$$

for all $k \in \mathbb{N}, i \in \{0, \ldots, I - 1\}$.

For sake of notational simplicity, let $f(\alpha) = \mathcal{J}(\boldsymbol{x}_k + \alpha\boldsymbol{d}_k)$. Moreover, the current iteration index k will remain implicit whenever unambiguous: typically, the stepsize α_k^i will be abridged into α^i. Using such compact notations, the stepsize update (8) also reads

$$\begin{cases} \alpha^0 = 0 \\ \alpha^{i+1} = \alpha^i - \theta\dot{f}(\alpha^i)/a_i, \quad i \in \{0, \ldots, I - 1\} \\ \alpha_k = \alpha^I \end{cases} \tag{13}$$

with $\dot{f}(\alpha^i) = \boldsymbol{d}_k^{\mathrm{t}}\nabla\mathcal{J}\left(\boldsymbol{x}_k + \alpha^i\boldsymbol{d}_k\right)$ and $a_i = \boldsymbol{d}_k^{\mathrm{t}}\boldsymbol{Q}_k^i\boldsymbol{d}_k$.

According to (9) we have

$$0 < \nu_1\|\boldsymbol{d}_k\|^2 \leq a_i \leq \nu_2\|\boldsymbol{d}_k\|^2. \tag{14}$$

According to (11), let

$$q_i(\alpha', \alpha) = \widehat{\mathcal{J}}_k^i(\boldsymbol{x}_k + \alpha' \boldsymbol{d}_k, \boldsymbol{x}_k + \alpha \boldsymbol{d}_k) = f(\alpha) + (\alpha' - \alpha)\dot{f}(\alpha) + (\alpha' - \alpha)^2 a_i/2, \tag{15}$$

which is a convex parabola as a function of α'.

Let us rely on a fixed number I of iterations of Weiszfeld's method for the determination of the stepsize. The (scalar) function to minimize is f and, according to Assumption 3, $q_i(\alpha', \alpha)$ is an upper convex quadratic approximation of $f(\alpha')$. Then, the successive iterations of Weiszfeld's method are defined by

$$\begin{aligned} \alpha^{i+1} &= \arg\min_{\alpha'} q_i(\alpha', \alpha^i) \\ &= \alpha^i - \dot{f}(\alpha^i)/a_i \end{aligned} \tag{16}$$

Hence, (13) identifies with a relaxed version of Weiszfeld's method to minimize f. Note that the number of iterations I can be chosen arbitrarily: the convergence results in Section 4 hold regardless of the value of I. This is in contrast with usual line search procedures, where appropriate stopping conditions (*e.g.,* Wolfe conditions) must be checked to ensure convergence.

As already mentioned, the stepsize formula (10) proposed in Refs. 1, 2 formally corresponds to one iteration of the same relaxed Weiszfeld's method. We are now led to a deeper result: the condition $\theta \in (0, \nu_1/\mu)$ stated in Refs. 1, 2 for the convergence of their CG method implies that our Assumption 3 holds. First, let us give an equivalent formulation for (10).

Let $\tilde{\boldsymbol{Q}}_k^0 = \boldsymbol{Q}_k^0/\theta$, so that (10) also reads $\alpha_k = -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \boldsymbol{d}_k^{\mathrm{t}} \tilde{\boldsymbol{Q}}_k^0 \boldsymbol{d}_k$. From (9) and $\theta \in (0, \nu_1/\mu)$, we deduce that

$$\boldsymbol{v}^{\mathrm{t}} \tilde{\boldsymbol{Q}}_k^0 \boldsymbol{v} \geq \mu \|\boldsymbol{v}\|^2, \quad \forall k \in \mathbb{N}, \forall \boldsymbol{v} \in \mathbb{R}^n, \tag{17}$$

i.e., the spectrum of matrices $\tilde{\boldsymbol{Q}}_k^0$ is bounded from below by μ. Such a simple change of notations shows that the constraint $\theta \in (0, \nu_1/\mu)$ stated in Refs. 1, 2 can be translated into a constraint on the matrices $\boldsymbol{Q}_k^i$. The following lemma shows that matrices $\tilde{\boldsymbol{Q}}_k^0$ yield convex quadratic majorizing approximations in the sense of Assumption 3 (provided that N is a convex set).

Lemma 2.1. *Suppose that Assumption 1 holds, and also that the lower bound ν_1 is not smaller than the Lipschitz constant μ. Let us restrict ourselves to the case where N is a convex set. Then Assumption 3 holds, i.e., the function $\widehat{\mathcal{J}}_k^i(\boldsymbol{x}', \boldsymbol{x})$ defined by* (11) *fulfills* (12) *over N.*

Proof. According to the Descent Lemma (Ref. 14, Prop. A.24), we have

$$\mathcal{J}(\boldsymbol{x}') - \mathcal{J}(\boldsymbol{x}) - (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}} \nabla \mathcal{J}(\boldsymbol{x}) \leq \mu \|\boldsymbol{x}' - \boldsymbol{x}\|^2/2 \tag{18}$$

for any $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^n$ if $\mathcal{J}$ is μ-$\mathcal{LC}^1$on $\mathbb{R}^n$. Actually, it is easy to check that (18) still holds for any $\boldsymbol{x}, \boldsymbol{x}' \in N$ if $\mathcal{J}$ is μ-$\mathcal{LC}^1$on N, provided that N is convex.

Since the spectrum of $\{\boldsymbol{Q}_k^i\}$ is bounded from below by $\nu_1 \geq \mu$, we have

$$\mu \|\boldsymbol{x}' - \boldsymbol{x}\|^2 \leq \nu_1 \|\boldsymbol{x}' - \boldsymbol{x}\|^2 \leq (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}} \, \boldsymbol{Q}_k^i \, (\boldsymbol{x}' - \boldsymbol{x})$$

Jointly with (18), the latter yields

$$\mathcal{J}(\boldsymbol{x}') - \mathcal{J}(\boldsymbol{x}) + (\boldsymbol{x} - \boldsymbol{x}')^{\mathrm{t}} \nabla \mathcal{J}(\boldsymbol{x}) \leq (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}} \, \boldsymbol{Q}_k^i \, (\boldsymbol{x}' - \boldsymbol{x}) \, /2,$$

i.e., $\widehat{\mathcal{J}}_k^i(\boldsymbol{x}', \boldsymbol{x}) \geq \mathcal{J}(\boldsymbol{x}')$. □

Lemma 2.1 indicates that Assumption 3 is not a restrictive condition compared to the hypotheses found in Refs. 1, 2. On the contrary, it is a weaker assumption (let alone the fact that Lemma 2.1 only applies when N is a convex set), so that a convergence proof based on Assumption 3 would be of broader applicability. This is the goal reached in Section 4, where $\mathcal{J}$ is not necessarily assumed ν_1-$\mathcal{LC}^1$(and N is not necessarily convex).

3 Properties of the stepsize series

The present section gathers technical results concerning the stepsize series $\alpha^i = \alpha_k^i$ generated by (8), which will be useful to derive the global convergence properties of the next section.

Let us introduce the notation $\Gamma(a, b) = [\min(a, b), \max(a, b)]$ to handle with intervals with unordered endpoints.

Lemma 3.1. *Suppose that Assumption 1 and Assumption 3 hold and that $\theta \in (0, 2)$. Then*

$$\mathcal{J}(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k) \leq \mathcal{J}(\boldsymbol{x}_k^i), \quad \forall \alpha \in \Gamma(\alpha_k^i, \alpha_k^{i+1}) \tag{19}$$

for all $k \geq 0$, $i \in \{0, \ldots, I-1\}$, where $\boldsymbol{x}_k^i = \boldsymbol{x}_k + \alpha_k^i \boldsymbol{d}_k$.

Proof. Let us first assume $\boldsymbol{x}_k^i \in N$, and then let us show that (19) holds, recursively on i and on n.

$\dot{f}(\alpha_k^i)$ exists since $\mathcal{J}$ is differentiable on N. We have $\alpha_k^0 = 0$ and $\dot{f}(0) = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq 0$, but the sign of $\dot{f}(\alpha_k^i) = \boldsymbol{d}_k^{\mathrm{t}} \nabla \mathcal{J}\left(\boldsymbol{x}_k + \alpha_k^i \boldsymbol{d}_k\right)$ is indeterminate for $i > 0$. Let us study each case separately (the index k is omitted in the rest of the proof).

- Suppose $\dot{f}(\alpha^i) = 0$. According to (13), $\alpha^{i+1} = \alpha^i$ so (19) is true.
- Suppose $\dot{f}(\alpha^i) < 0$. According to (13) and $a_i > 0$ we have $\alpha^{i+1} > \alpha^i$. Let us prove (19) by contradiction: suppose, on the contrary, that there exists $\alpha' \in (\alpha^i, \alpha^{i+1}]$ such that

$$f(\alpha') > f(\alpha^i). \tag{20}$$

Let $\ell^i = \{\alpha \in \mathbb{R} | f(\alpha) \leq f(\alpha^i)\}$. Since f is continuous on ℓ^i, according to (20) and $\dot{f}(\alpha^i) < 0$, there exists $\alpha'' \in (\alpha^i, \alpha')$ such that $f(\alpha'') < f(\alpha^i)$. There also exists $\alpha''' \in (\alpha'', \alpha')$ such that

$$f(\alpha''') = f(\alpha^i); \tag{21}$$

in the contrary case, since f is continuous on ℓ^i, the inequality $f(\alpha) < f(\alpha^i)$ would hold for all $\alpha \in (\alpha'', \alpha')$. In particular, we would get

$$\lim_{\substack{\alpha \to \alpha' \\ \alpha < \alpha'}} f(\alpha) < f(\alpha^i), \tag{22}$$

thus $\alpha \in \ell^i$ for all values of $\alpha \in (\alpha'', \alpha')$ arbitrary close to α'. (22) would be incompatible with (20) given the continuity of f on ℓ^i.

Let $q(\alpha) = q_i(\alpha, \alpha^i)$, where q_i is defined by (15). Since $\dot{q}(\alpha^{i+1}) = \dot{f}(\alpha^i)(1 - \theta)$, $\dot{q}(\alpha^i) = \dot{f}(\alpha^i) < 0$ and $\theta \in (0, 2)$, we have $\dot{q}(\alpha^{i+1}) \in (\dot{q}(\alpha^i), -\dot{q}(\alpha^i))$. Because q is a convex parabola and $\alpha''' \in (\alpha'', \alpha') \subset (\alpha^i, \alpha^{i+1})$, we can conclude that $q(\alpha''') < q(\alpha^i) = f(\alpha^i)$. Hence, according to (21), we get $q(\alpha''') < f(\alpha''')$, which contradicts the majorizing character (12) of $\widehat{\mathcal{J}}_k^i$ w.r.t. $\mathcal{J}$ at $\boldsymbol{x}_k + \alpha''' \boldsymbol{d}_k \in N$.

- Suppose $\dot{f}(\alpha^i) > 0$. According to (13) and $a_i > 0$, we have $\alpha^{i+1} < \alpha^i$. We are led back to the previous case if we replace $f(\alpha)$ by $f(-\alpha)$.

Our intermediate conclusion is that (19) holds if $\boldsymbol{x}_k^i \in N$: in particular $\mathcal{J}(\boldsymbol{x}_k^{i+1}) \leq \mathcal{J}(\boldsymbol{x}_k^i)$. Since $\boldsymbol{x}_0 \in N$ and $\boldsymbol{x}_{k+1}^0 = \boldsymbol{x}_k^I = \boldsymbol{x}_k$, we get

$$\mathcal{J}(\boldsymbol{x}_k^i) \leq \ldots \leq \mathcal{J}(\boldsymbol{x}_k^0) = \mathcal{J}(\boldsymbol{x}_{k-1}^I) \leq \ldots \leq \mathcal{J}(\boldsymbol{x}_{k-1}^0) \leq \ldots \leq \mathcal{J}(\boldsymbol{x}_1^0) = \mathcal{J}(\boldsymbol{x}_0)$$

by immediate recursion, which proves that (19) holds for all $k \geq 0$, $i \in \{0, \ldots, I-1\}$. ☐

An immediate consequence of Lemma 3.1 is

$$\boldsymbol{x}_k + \alpha \boldsymbol{d}_k \in N, \quad \forall \alpha \in [0, \alpha_k^i], \tag{23}$$

for all $k \geq 0$, $i \in \{0, \ldots, I-1\}$ since $\boldsymbol{x}_0 \in N$. Thus, according to (12),

$$q_i(\alpha^j, \alpha^i) \geq f(\alpha^j), \quad \forall i, j \in \{0, \ldots, I-1\}. \tag{24}$$

The following three lemmas are specific to the case when $\dot{f}(0) = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k$ does not cancel for the current iteration k, *i.e.*, $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k < 0$. Then $\boldsymbol{d}_k \neq 0$, and the sequence $\{\alpha_k^i\}$ is well defined according to (8b).

Lemma 3.2. *Suppose that Assumption 1 and Assumption 3 hold. Assume also that $\dot{f}(0) < 0$ and $\theta \in (0, 2)$. Then the whole sequence $\{\alpha^i\}$ is strictly positive:*

$$\alpha^i > 0, \quad \forall i \in \{0, \ldots, I-1\}. \tag{25}$$

Proof. According to (24), we have

$$q_i(0,\alpha^i) = f(\alpha^i) - \alpha^i \dot{f}(\alpha^i) + (\alpha^i)^2 a_i/2 \geq f(0).$$

Since $\{f(\alpha^i)\}$ is a nonincreasing sequence according to Lemma 3.1, we deduce that

$$-\alpha^i \dot{f}(\alpha^i) + (\alpha^i)^2 a_i/2 \geq 0$$

so that, according to (13) and $a_i > 0$,

$$\alpha^i(\alpha^{i+1} - 2\delta \dot{f}(\alpha^i)/a_i) \geq 0 \tag{26}$$

with

$$\delta = 1 - \theta/2 \in (0,1)\,. \tag{27}$$

Now, let us show (25) by a recurrence on i. We have $\alpha^1 > 0$ according to (13). Let us assume now that $\alpha^i > 0$ for some i.

- If $\dot{f}(\alpha^i) \leq 0$, then $\alpha^{i+1} > 0$ according to (13).
- If $\dot{f}(\alpha^i) > 0$, then given $\alpha^i > 0$, inequality (26) yields $\alpha^{i+1} > 0$.

□

Lemma 3.3. *Suppose that Assumption 1 and Assumption 3 hold. Assume also that* $\dot{f}(0) < 0$ *and* $\theta \in (0,2)$*. Then, for all* $i \in \mathbb{N} - \{0\}$*,*

$$f(\alpha^i) \leq q_0(\alpha^1, 0), \tag{28}$$

$$c^{\min}\alpha^1 \leq \alpha^i, \tag{29}$$

where

$$c^{\min} = \left(\sqrt{1 + 2\mu\theta\delta/\nu_1} - 1\right)\nu_1/\theta\mu \in (0,1)\,. \tag{30}$$

Proof. The proof of (28) is straightforward: according to (24), we have $q_0(\alpha^1, 0) \geq f(\alpha^1)$. Then (28) holds, because $\{f(\alpha^i)\}$ is a decreasing sequence according to Lemma 3.1.

The derivation of (29) is not so direct. Let g the concave parabola defined by

$$g(\alpha) = f(0) + \alpha \dot{f}(0) - \mu a_0 \alpha^2/2\nu_1. \tag{31}$$

Remark that $g(0) = f(0)$ and that g is decreasing on $\mathbb{R}^+$ since $\dot{g}(0) = \dot{f}(0) < 0$.

Let us first show that

$$g(\alpha^i) \leq f(\alpha^i). \tag{32}$$

Let us consider $\alpha \in [0, \alpha^i]$: $\boldsymbol{x}_k + \alpha \boldsymbol{d}_k \in N$ according to (23). Since $f(\alpha) = \mathcal{J}(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k)$ and Assumption 1 holds, we have

$$|\dot{f}(\alpha) - \dot{f}(0)| = \left|\boldsymbol{d}_k^{\mathrm{t}} \left(\nabla \mathcal{J}(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k) - \nabla \mathcal{J}(\boldsymbol{x}_k)\right)\right| \leq \|\boldsymbol{d}_k\|^2 \mu |\alpha|$$

and according to (14), we get

$$|\dot{f}(\alpha) - \dot{f}(0)| \leq a_0 \mu \alpha / \nu_1.$$

Given $|\dot{f}(\alpha)| \leq |\dot{f}(\alpha) - \dot{f}(0)| + |\dot{f}(0)|$ and $\dot{f}(0) < 0$, we obtain

$$|\dot{f}(\alpha)| \leq a_0 \mu \alpha / \nu_1 - \dot{f}(0). \tag{33}$$

In particular,

$$\dot{f}(0) - a_0 \mu \alpha / \nu_1 \leq \dot{f}(\alpha)$$

or equivalently

$$\dot{g}(\alpha) \leq \dot{f}(\alpha), \quad \forall \alpha \in [0, \alpha^i] \tag{34}$$

according to (31). Since $g(0) = f(0)$, (32) is obtained by integrating (34) between 0 and α^i.

According to (13), (15) and (27), we have

$$\begin{aligned} q_0(\alpha^1, 0) &= f(0) + \alpha^1 \dot{f}(0) + (\alpha^1)^2 a_0 / 2 \\ &= f(0) + \delta \alpha^1 \dot{f}(0). \end{aligned} \tag{35}$$

Let $\alpha^{\min} = c^{\min} \alpha^1$. Combining (31) and (35), it is easy to establish that $g(\alpha^{\min}) = q_0(\alpha^1, 0)$. Moreover, $\{\alpha^i\}$ is positive according to Lemma 3.2. We are now in position to show (29) by contradiction: assume that there exists $i > 0$ such that $0 \leq \alpha^i < \alpha^{\min}$. According to (32) and given that g is decreasing on $\mathbb{R}^+$, we get $f(\alpha^i) \geq g(\alpha^i) > g(\alpha^{\min}) = q_0(\alpha^1, 0)$, which contradicts (28).

Finally, it is obvious that $c^{\min} > 0$. Let us consider the alternate expression

$$c^{\min} = 2\delta / \left(\sqrt{1 + 2\mu\theta\delta/\nu_1} + 1\right),$$

so it becomes also apparent that $c^{\min} < \delta < 1$. □

Lemma 3.4. *Suppose that Assumption 1 and Assumption 3 hold. Assume also that* $\dot{f}(0) < 0$ *and* $\theta \in (0, 2)$. *Then*

$$\alpha^i \leq c_i^{\max} \alpha^1 \tag{36}$$

$\forall i \in \mathbb{N} - \{0\}$, *with*

$$c_i^{\max} = \left(1 + \nu_2 \theta \mu / \nu_1^2\right)^{i-1} (1 + \nu_1 / \theta \mu) - \nu_1 / \theta \mu \geq 1. \tag{37}$$

Proof. It is easy to check that $c_i^{\max}$ is not smaller than 1 for all $i > 0$. Let us show the inequality (36) recursively on i. It is valid for $i = 1$, since $c_1^{\max} = 1$. Now let us suppose that $\alpha^i \leq c_i^{\max}\alpha^1$, and let us prove that $\alpha^{i+1} \leq c_{i+1}^{\max}\alpha^1$.

According to (13), we have $\alpha^{i+1} \leq \alpha^i + |\dot{f}(\alpha^i)|\,\theta/a_i$ and according to (14), we have also $a_i \geq a_0\nu_1/\nu_2$. Thus,

$$\alpha^{i+1} \leq \alpha^i + |\dot{f}(\alpha^i)|\,\theta\nu_2/\nu_1 a_0. \tag{38}$$

On the other hand, (33) implies

$$|\dot{f}(\alpha^i)| \leq a_0\mu\alpha^i/\nu_1 - \dot{f}(0).$$

In combination with the latter inequality and with $\alpha^1 = -\theta\dot{f}(0)/a_0$, (38) yields

$$\alpha^{i+1} \leq \alpha^i(1 + \nu_2\theta\mu/\nu_1^2) + \nu_2\alpha^1/\nu_1,$$

which corresponds to a recursive definition of the series $(c_i^{\max})$ according to

$$c_{i+1}^{\max} = c_i^{\max}(1 + \nu_2\theta\mu/\nu_1^2) + \nu_2/\nu_1.$$

Given $c_1^{\max} = 1$, it can be checked that (37) is the general term of the series. □

Definition 3.1. *The stepsize sequence $\{\alpha_k\}$ satisfies the Armijo condition with $\Omega \in (0, 1)$ if*

$$\mathcal{J}(\boldsymbol{x}_k) - \mathcal{J}(\boldsymbol{x}_{k+1}) + \Omega\alpha_k\boldsymbol{g}_k^{\mathrm{t}}\boldsymbol{d}_k \geq 0, \quad \forall k. \tag{39}$$

Lemma 3.5. *Suppose that Assumption 1 and Assumption 3 hold. Assume also that $\theta \in (0, 2)$. Then the stepsize sequence defined by (8) satisfies the Armijo condition with*

$$\Omega = \Omega_I = \delta/c_I^{\max} \in (0, 1)\,, \tag{40}$$

where δ and $c_I^{\max}$ are defined by (27) and (37), respectively.

Proof. We have $\dot{f}(0) = \boldsymbol{g}_k^{\mathrm{t}}\boldsymbol{d}_k \leq 0$. Let us first examine the particular case $\dot{f}(0) = 0$: according to (8), α_k vanishes, so that (39) holds trivially.

Suppose now $\dot{f}(0) < 0$. According to (35), (28) also reads

$$f(0) - f(\alpha^I) + \delta\dot{f}(0)\alpha^1 \geq 0. \tag{41}$$

Finally, since $\dot{f}(0) < 0$ and $\alpha^1 \geq \alpha^I/c_I^{\max} > 0$ according to (36), (41) implies that

$$f(0) - f(\alpha^I) + \delta\dot{f}(0)\alpha^I/c_I^{\max} \geq 0,$$

which identifies with (39) with $\Omega = \Omega_I$. □

Remark 3.1. *In Lemma 3.5, $\Omega_I = \delta/c_I^{\max}$ does not depend on k, which is an essential point for the fulfillment of the Armijo condition.*

The following theorem sums up the main results that will be useful in the next section.

Theorem 3.1. *Let $\boldsymbol{x}_k$ be defined by (2)-(8) with $\theta \in (0, 2)$, and let Assumption 1 and Assumption 3 hold. Then the Armijo condition (39) is satisfied by the stepsize sequence $\{\alpha_k\}$ with $\Omega = \Omega_I = \delta / c_I^{\max}$, where δ and $c_I^{\max}$ are defined by (27) and (37), respectively. Moreover, we have*

$$0 \leq c^{\min} \alpha_k^1 \leq \alpha_k \leq c_I^{\max} \alpha_k^1, \quad \forall k, \tag{42}$$

where $c^{\min}$ is defined by (30).

Proof. Lemma 3.5 corresponds to the fulfillment of the Armijo condition.

On the other hand, we have $\dot{f}(0) \leq 0$. If $\dot{f}(0) = 0$, then $\alpha_k = 0$, so (42) trivially holds. Otherwise, we have $\dot{f}(0) < 0$, so (42) is a joint consequence of Lemmas 3.3 and 3.4. □

4 Global convergence

The two following lemmas establish results for the whole two-parameter family of conjugacy coefficient $\beta_k = \beta_k^{\mu_k, \omega_k}$. Then, we will draw conclusions for specific CG methods.

Lemma 4.1. *Under the conditions of Theorem 3.1, we have*

$$\sum_{k,\, \boldsymbol{d}_k \neq 0} (\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k)^2 / \|\boldsymbol{d}_k\|^2 < \infty. \tag{43}$$

Proof. According to Theorem 3.1, the Armijo condition (39) is satisfied with $\Omega = \Omega_I$. Given (42) and $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq 0$, we deduce that

$$\mathcal{J}(\boldsymbol{x}_k) - \mathcal{J}(\boldsymbol{x}_{k+1}) \geq -\Omega_I c^{\min} \alpha_k^1 \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k. \tag{44}$$

If $\boldsymbol{d}_k \neq \boldsymbol{0}$, we have

$$\alpha_k^1 = -\theta \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \boldsymbol{d}_k^{\mathrm{t}} \boldsymbol{Q}_k^0 \boldsymbol{d}_k \geq -\theta \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \nu_2 \|\boldsymbol{d}_k\|^2 \tag{45}$$

according to (10) and (9), so that

$$\mathcal{J}(\boldsymbol{x}_k) - \mathcal{J}(\boldsymbol{x}_{k+1}) \geq c_0 (\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k)^2 / \|\boldsymbol{d}_k\|^2 \geq 0 \tag{46}$$

with $c_0 = \Omega_I c^{\min} \theta / \nu_2 > 0$. Given Assumption 1 and the fact that L is bounded, (46) implies $\lim_{k \to \infty} \mathcal{J}(\boldsymbol{x}_k)$ is finite. Finally, we obtain

$$\infty > (\mathcal{J}(\boldsymbol{x}_0) - \lim_{k \to \infty} \mathcal{J}(\boldsymbol{x}_k)) / c_0 \geq \sum_{k,\, \boldsymbol{d}_k \neq 0} (\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k)^2 / \|\boldsymbol{d}_k\|^2.$$

□

Lemma 4.2. *Let $k \in \mathbb{N}$. Under the conditions of Theorem 3.1, we have*

$$|\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k| \leq -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \left(1 + c_I^{\max} \theta \mu / \nu_1\right). \tag{47}$$

Moreover, if Assumption 2 holds, then

$$-\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k \leq -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k (1 - c^{\min} \theta \lambda / \nu_2). \tag{48}$$

Proof. (47) and (48) are trivial assertions if $\boldsymbol{d}_k = \mathbf{0}$. Otherwise, following Refs. 1, 2, let us define

$$\phi_k = \begin{cases} (\boldsymbol{g}_{k+1} - \boldsymbol{g}_k)^{\mathrm{t}} (\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) / \|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|^2 = \boldsymbol{y}_k^{\mathrm{t}} \boldsymbol{d}_k / \alpha_k \|\boldsymbol{d}_k\|^2 & \text{for } \alpha_k \neq 0 \\ 0 & \text{for } \alpha_k = 0. \end{cases} \tag{49}$$

Note that according to (23), $\boldsymbol{x}_k \in N$. If Assumption 1 holds, then $|\phi_k| \leq \mu$ according to Cauchy-Schwartz inequality. If Assumption 2 holds, then $\phi_k \geq \lambda > 0$.

According to (10) and (49), we have

$$\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k + \boldsymbol{y}_k^{\mathrm{t}} \boldsymbol{d}_k = \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k + \alpha_k \phi_k \|\boldsymbol{d}_k\|^2. \tag{50}$$

According to (42), (50), $\mu \geq |\phi_k|$, and $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq 0$, we deduce that

$$|\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k| \leq -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k + \mu c_I^{\max} \alpha_k^1 \|\boldsymbol{d}_k\|^2.$$

According to (10), we have also

$$|\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k| \leq -\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k - \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \mu c_I^{\max} \theta \|\boldsymbol{d}_k\|^2 / \boldsymbol{d}_k^{\mathrm{t}} \boldsymbol{Q}_k^0 \boldsymbol{d}_k.$$

Finally, since $\nu_1 > 0$ is a lower bound for the spectrum of $\boldsymbol{Q}_k^0$, and $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq 0$, we obtain (47).

Let us suppose now that Assumption 2 holds. Given (42) and $\phi_k \geq \lambda > 0$, (50) implies

$$\boldsymbol{g}_{k+1}^{\mathrm{t}} \boldsymbol{d}_k \geq \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k + \lambda c^{\min} \alpha_k^1 \|\boldsymbol{d}_k\|^2$$

and according to (45), we obtain (48). □

Lemma 4.3. *Suppose that Assumption 2 holds, as well as the conditions of Theorem 3.1. Then*

$$D_k \geq (1 - \mu_k - \omega_k) \|\boldsymbol{g}_{k-1}\|^2 - \boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} (\omega_k + \mu_k c^{\min} \theta \lambda / \nu_2) \geq 0, \quad \forall k \in \mathbb{N} - \{0\}. \tag{51}$$

Proof. Since $\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$, (48) also reads

$$\boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{y}_{k-1} \geq -\boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} c^{\min} \theta \lambda / \nu_2, \quad \forall k \in \mathbb{N} - \{0\}.$$

Then, given the expression (5) of D_k and $\boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} \leq 0$, the conclusion is immediate. □

Remark 4.1. *Let us examine the case where the denominator* D_k *of* $\beta_k^{\mu_k,\omega_k}$ *cancels. Here, we assume that the conditions of Theorem 3.1 hold.*

Let us suppose first that Assumption 2 is valid. If D_k *cancels, then (51) implies*

$$(1-\mu_k-\omega_k)\|\boldsymbol{g}_{k-1}\|^2 - (\omega_k + \mu_k c^{\min}\theta\lambda/\nu_2)\boldsymbol{d}_{k-1}^{\mathrm{t}}\boldsymbol{g}_{k-1} = 0.$$

Since the left-hand side is the sum of two nonnegative terms, we obtain

$$\begin{cases} (1-\mu_k-\omega_k)\|\boldsymbol{g}_{k-1}\|^2 = 0, & \text{(52a)} \\ (\omega_k + \mu_k c^{\min}\theta\lambda/\nu_2)\boldsymbol{d}_{k-1}^{\mathrm{t}}\boldsymbol{g}_{k-1} = 0. & \text{(52b)} \end{cases}$$

- *Case 1: If* $\mu_k+\omega_k < 1$, *(52a) boils down to* $\|\boldsymbol{g}_{k-1}\|^2 = 0$, *which means that convergence is reached at iteration* $k-1$. *This case includes the PRP method.*
- *Case 2: If* $\mu_k + \omega_k = 1$, *(52b) implies* $\boldsymbol{d}_{k-1}^{\mathrm{t}}\boldsymbol{g}_{k-1} = 0$, *so that* $\alpha_{k-1} = 0$. *Thus,* $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}$, $\boldsymbol{y}_{k-1} = 0$, *and the numerator of* $\beta_k^{\mu_k,\omega_k}$ *cancels. In this case, we let* $\beta_k^{\mu_k,\omega_k} = 0$, *conventionally. This case includes the HS and the LS method.*

In the situation where Assumption 2 is not necessarily valid, our study only covers the case $\mu_k = 0$*: then* D_k *is the sum of two nonnegative terms, so* $D_k = 0$ *implies that both cancel:*

$$\begin{cases} (1-\omega_k)\|\boldsymbol{g}_{k-1}\|^2 = 0, \\ \omega_k \boldsymbol{d}_{k-1}^{\mathrm{t}}\boldsymbol{g}_{k-1} = 0. \end{cases}$$

- *If* $\omega_k < 1$, *the conclusion is the same as in Case 1. This case includes the PRP method.*
- *If* $\omega_k = 1$, *the conclusion is the same as in Case 2. This case includes the LS method.*

Lemma 4.4. *Under the conditions of Theorem 3.1, we have*

$$\liminf_{k\to\infty} \|\boldsymbol{g}_k\| > 0 \implies \lim_{k\to\infty} \beta_k^{0,\omega_k} = 0.$$

Moreover, if Assumption 2 is valid, then

$$\liminf_{k\to\infty} \|\boldsymbol{g}_k\| > 0 \implies \lim_{k\to\infty} \beta_k^{\mu_k,\omega_k} = 0.$$

Proof. According to (2) and (42), we have

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|^2 = \alpha_k^2\|\boldsymbol{d}_k\|^2 \le (c_I^{\max})^2(\alpha_k^1)^2\|\boldsymbol{d}_k\|^2.$$

Given that (10) holds unless $\boldsymbol{d}_k = \mathbf{0}$, we deduce that

$$\begin{aligned}\sum_k \|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|^2 &\le (c_I^{\max}\theta)^2 \sum_{k,\,\boldsymbol{d}_k\neq 0} \left(\boldsymbol{g}_k^{\mathrm{t}}\boldsymbol{d}_k\right)^2 \|\boldsymbol{d}_k\|^2 / \left(\boldsymbol{d}_k^{\mathrm{t}}\boldsymbol{Q}_k^0\boldsymbol{d}_k\right)^2, \\ &\le (c_I^{\max}\theta/\nu_1)^2 \sum_{k,\,\boldsymbol{d}_k\neq 0} \left(\boldsymbol{g}_k^{\mathrm{t}}\boldsymbol{d}_k\right)^2 / \|\boldsymbol{d}_k\|^2\end{aligned}$$

according to (9). Given (43), we conclude that $\lim_{k\to\infty} \|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|^2 = 0$. Because $\mathcal{J}$ is continuously differentiable and $\|\boldsymbol{g}_k\|$ is bounded according to Assumption 1 and the boundedness of L, we have also $\lim_{k\to\infty} \boldsymbol{y}_{k-1} = 0$ and

$$\lim_{k\to\infty} \boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1} = 0. \tag{54}$$

If $\liminf_{k\to\infty} \|\boldsymbol{g}_k\| > 0$, then there exists $\gamma > 0$ such that

$$\|\boldsymbol{g}_k\| \geq \gamma > 0 \quad \forall k. \tag{55}$$

According to (6), we have

$$|\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| = |\beta_k^{\mu_k,\omega_k}||D_k|. \tag{56}$$

On the one hand, suppose that Assumption 2 is valid.

Firstly, let us consider the iteration indices k such that $\mu_k + \omega_k \in [0, 1/2]$. According to (51) and $\boldsymbol{d}_{k-1}^{\mathrm{t}} \boldsymbol{g}_{k-1} \leq 0$, (56) implies that

$$|\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| \geq |\beta_k^{\mu_k,\omega_k}|(1 - \mu_k - \omega_k)\|\boldsymbol{g}_{k-1}\|^2,$$

which leads to

$$|\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| \geq |\beta_k^{\mu_k,\omega_k}|\gamma^2/2, \tag{57}$$

given (55).

Let us establish a similar result in the more complex case $\mu_k + \omega_k \in (1/2, 1]$. As a preliminary step, let us show that

$$\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k \leq -\gamma^2/2 \tag{58}$$

for all sufficiently large values of k.

According to Remark 4.1, in the case $\boldsymbol{g}_{k-1}^{\mathrm{t}} \boldsymbol{d}_{k-1} = 0$, we have $\beta_k^{\mu_k,\omega_k} = 0$, so $\boldsymbol{d}_k = -\boldsymbol{g}_k$ and (58) is valid according to (55).

Now let us consider the case where $\boldsymbol{g}_{k-1}^{\mathrm{t}} \boldsymbol{d}_{k-1} < 0$. Given (3) and (6), we have

$$\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{c}_k = \boldsymbol{g}_k^{\mathrm{t}}(-\boldsymbol{g}_k + \beta_k^{\mu_k,\omega_k} \boldsymbol{d}_{k-1}) = -\|\boldsymbol{g}_k\|^2 + (\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1})(\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_{k-1})/D_k.$$

According to (47), (51), and (58), we deduce

$$\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{c}_k \leq -\gamma^2 + |\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| \left(1 + c_I^{\max}\theta\mu/\nu_1\right) / (\omega_k + \mu_k c^{\min}\theta\lambda/\nu_2). \tag{59}$$

Given (54), the latter inequality yields $\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{c}_k \leq -\gamma^2/2$ for all sufficiently large k. Because of (4), we can conclude that (58) holds.

Given (51) and (58), (56) implies

$$\begin{aligned} |\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| &\geq |\beta_k^{\mu_k,\omega_k}| \left((1 - \mu_k - \omega_k)\gamma^2 + (\omega_k + \mu_k c^{\min}\theta\lambda/\nu_2)\gamma^2/2\right) \\ &= |\beta_k^{\mu_k,\omega_k}| \left(1 - \omega_k/2 - (1 - c^{\min}\theta\lambda/\nu_2)\mu_k\right) \gamma^2 \end{aligned}$$

for all sufficiently large values of k. Given $\mu_k + \omega_k \in (1/2, 1]$, the latter inequality implies

$$|\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| \geq |\beta_k^{\mu_k,\omega_k}| m\gamma^2, \tag{60}$$

where $m = \min\{1/2, c^{\min}\theta\lambda/\nu_2\}$.

Since $m \leq 1/2$, (60) is implied by (57), so that (60) holds in the whole domain $\mu_k \in [0,1]$, $\omega_k \in [0, 1-\mu_k]$. Finally, (54) and (60) jointly imply $\lim_{k\to\infty} |\beta_k^{\mu_k,\omega_k}| = 0$.

On the other hand, consider the case where Assumption 2 is not necessarily valid. If $\mu_k = 0$, then we have

$$|\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{y}_{k-1}| \geq |\beta_k^{0,\omega_k}| \gamma^2/2.$$

The proof is similar to that of (60), where the two cases to examine are $\omega_k \in [0, 1/2]$ and $\omega_k \in (1/2, 1]$. Finally, according to (54), we have $\lim_{k\to\infty} |\beta_k^{0,\omega_k}| = 0$. □

Remark 4.2. *The proof of Lemma 4.4 is inspired from that of Ref. 2, Lemma 3.2, but we deal with the more general case of the iterated formula (8). Moreover, μ_k and ω_k are possibly varying, while they are constant parameters in Ref. 2.*

Theorem 4.1. *Let $\boldsymbol{x}_k$ be defined by (2)-(8) with $\theta \in (0, 2)$, and let Assumption 1 and Assumption 3 hold.*

Then we have convergence in the sense $\liminf_{k\to\infty} \boldsymbol{g}_k = \boldsymbol{0}$ for the PRP and LS methods, and more generally for $\mu_k = 0$ and $\omega_k \in [0, 1]$.

Moreover, if Assumption 2 holds, then we have also $\liminf_{k\to\infty} \boldsymbol{g}_k = \boldsymbol{0}$ in all cases.

Proof. Assume on the contrary that $\|\boldsymbol{g}_k\| \geq \gamma > 0$ for all k. Since L is bounded, both $\{\boldsymbol{x}_k\}$ and $\{\boldsymbol{g}_k\}$ are bounded.

Let us first suppose that Assumption 2 holds. Since $\liminf_{k\to\infty} \|\boldsymbol{g}_k\| > 0$, by Lemma 4.4 we have $\lim_{k\to\infty} \beta_k^{\mu_k,\omega_k} = 0$.

Since

$$\|\boldsymbol{d}_k\| = \|\boldsymbol{c}_k\| \leq \|\boldsymbol{g}_k\| + |\beta_k^{\mu_k,\omega_k}| \|\boldsymbol{d}_{k-1}\|,$$

we conclude that $\{\|\boldsymbol{d}_k\|\}$ is uniformly bounded for sufficient large k. Thus we have

$$\begin{aligned} |\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k| &= |\boldsymbol{g}_k^{\mathrm{t}}(-\boldsymbol{g}_k + \beta_k^{\mu_k,\omega_k} \boldsymbol{d}_{k-1})| \\ &\geq \|\boldsymbol{g}_k\|^2 - |\beta_k^{\mu_k,\omega_k}| \|\boldsymbol{g}_k\| \|\boldsymbol{d}_{k-1}\| \\ &\geq \|\boldsymbol{g}_k\|^2/2 \end{aligned}$$

for sufficient large k. Then there exists $\epsilon > 0$ so that

$$\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \|\boldsymbol{d}_k\| \|\boldsymbol{g}_k\| \geq \|\boldsymbol{g}_k\|/2\|\boldsymbol{d}_k\| \geq \epsilon$$

for sufficient large k. Finally, we conclude that

$$\sum_{k,\, \boldsymbol{d}_k \neq 0} \|\boldsymbol{g}_k\|^2 \left(\boldsymbol{g}_k^{\mathrm{t}} \boldsymbol{d}_k / \|\boldsymbol{d}_k\| \|\boldsymbol{g}_k\|\right)^2 = \infty.$$

This is a contradiction to Lemma 4.1.

The same proof applies to the case where $\mu_k = 0$, and Assumption 2 is not necessarily valid. □

Remark 4.3. *The proof of Theorem 4.1 is partly inspired from that of Ref. 2, Theorem 3.3. However, our result deals with variable parameters μ_k, ω_k. Moreover, Assumption 2 is not necessary in the case $\mu_k = 0$, which contains the PRP and LS methods.*

5 Discussion

5.1 The convex quadratic case

Let us show that the convergence condition $\theta \in (0, \nu_1/\mu)$ is too restrictive for the stepsize formula proposed in Refs. 1, 2, in the case of a convex quadratic objective function. Let

$$\mathcal{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{t}}\boldsymbol{Q}\boldsymbol{x}/2 - \boldsymbol{b}^{\mathrm{t}}\boldsymbol{x}, \quad \boldsymbol{x} \in \mathbb{R}^n$$

where $\boldsymbol{Q}$ is a symmetric, positive definite matrix. Let ν_1 and ν_2 respectively denote the smallest and largest eigenvalue of $\boldsymbol{Q}$, so that (9) holds.

Now, consider the stepsize formula (10) with $\boldsymbol{Q}_k^0 = \boldsymbol{Q}$. When $\theta = 1$, it yields the optimal stepsize $\alpha_k = \arg\min_\alpha \mathcal{J}(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k)$.

In the convex quadratic case, Theorem 4.1 ensures the convergence for $\theta \in (0, 2)$ and for any fixed $I > 0$. Remark that Assumption 3 is easily checked, since

$$\hat{\mathcal{Q}}(\boldsymbol{x}', \boldsymbol{x}) = \mathcal{Q}(\boldsymbol{x}) + (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}}\, \nabla \mathcal{Q}(\boldsymbol{x}) + (\boldsymbol{x}' - \boldsymbol{x})^{\mathrm{t}}\, \boldsymbol{Q}\, (\boldsymbol{x}' - \boldsymbol{x})/2 = \mathcal{Q}(\boldsymbol{x}').$$

In the case of the optimal stepsize, *i.e.*, $\theta = 1$, the classical linear CG algorithm is covered, since all conjugacy formulas reduce to the FR method.

On the contrary, the convergence domain deduced from Refs. 1, 2 is $\theta \in (0, \nu_1/\nu_2) \subset (0, 1)$, since $\nabla \mathcal{Q}$ is μ-$\mathcal{LC}^1$with μ not larger than ν_2. Hence, the convergence of the classical linear CG algorithm is not recovered. In particular, the condition $\theta \in (0, \nu_1/\nu_2)$ will produce excessively small and inefficient stepsizes when the Hessian matrix $\boldsymbol{Q}$ is ill-conditioned.

5.2 The general case

In the general case, the nontrivial computation of ν_1 and μ is a prerequisite to check the convergence condition $\theta \in (0, \nu_1/\mu)$. In Ref. 2, it is rather proposed to ensure the convergence empirically by choosing an arbitrarily small value of θ. Unfortunately, the resulting algorithm will be hardly competitive, compared to CG methods with a usual line search procedure.

Our convergence results do not share the same drawback, provided that, as a preliminary step, a convex quadratic function has been found to approximate the objective function from above. According to Lemma 2.1, finding such a convex quadratic majorizing function is always possible when N is a convex set.

In practice, case-by-case considerations may provide tighter convex quadratic approximations, that will result in larger stepsizes. This issue is actually not new, since finding a good convex quadratic majorizing function is already a crucial step in the use of Weiszfeld's method Ref. 13. The latter reference provides examples in the field of optimal location (which was Weiszfeld's original concern), and in structural mechanics. Robust regression is another classical area where Weiszfeld's method is widely applied, under the name of *Iterative Reweighting* (Ref. 15). More recently, the latter has also become a standard approach for edge preserving image restoration, under the name of *Half-Quadratic Scheme* (Refs. 16, 17). Edge preserving image restoration is structurally similar to the robust regression problem, but of much larger scale since digital images commonly gather billions of pixels. Direct application of Weiszfeld's method then becomes inefficient or even impracticable, since each iteration amounts to solve a linear system of size n. In Ref. 18, we go further into the details of the image restoration application, to support that our proposed CG method is a natural substitute for Weiszfeld's method.

References

1. J. Sun and J. Zhang, "Global convergence of conjugate gradient methods without line search", *Annals of Operations Research*, vol. 103, pp. 161–173, 2001.

2. X. Chen and J. Sun, "Global convergence of two-parameter family of conjugate gradient methods without line search", *Journal of Computational and Applied Mathematics*, vol. 146, pp. 37–45, 2002.

3. P. Wolfe, "Convergence conditions for ascent methods. II: Some corrections", *SIAM Review*, vol. 13, pp. 185–188, 1971.

4. Y. Dai and Y. Yuan, "A three-parameter family of nonlinear conjugate gradient methods", *Mathematics of Computation*, vol. 70, pp. 1155–1167, 2001.

5. M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear system", *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 409–436, 1952.

6. E. Polak and G. Ribière, "Note sur la convergence de méthodes des directions conjuguées", *Revue française d'Informatique et de Recherche Opérationelle*, vol. 16, pp. 35–43, 1969.

7. B. T. Polyak, "The conjugate gradient method in extreme problems", *USSR Computational Mathematics and Mathematical Physics*, vol. 9, pp. 94–112, 1969.

8. Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: Theory", *Journal of Optimisation Theory and Applications*, vol. 69, pp. 129–137, 1991.

9. R. Fletcher and C. Reeves, "Function minimization by conjugate gradients", *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

10. R. Fletcher, *Practical Methods of Optimization*, Wiley&Sons, New York, 2nd edition, 1987.

11. Y. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property", *SIAM Journal of Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

12. E. Weiszfeld, "Sur le point pour lequel la somme des distances de n points donnés est minimum", *Tôhoku Mathematical Journal*, vol. 43, pp. 355–386, 1937.

13. H. Voss and U. Eckhardt, "Linear convergence of generalized Weiszfeld's method", *Computing*, vol. 25, pp. 243–251, 1980.

14. D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 2nd edition, 1999.

15. P. J. Huber, *Robust Statistics*, John Wiley, New York, NY, 1981.

16. P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging", *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, Feb. 1997.

17. M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms", to appear, IEEE Transactions on Image Processing, 2006.

18. C. Labat and J. Idier, "Preconditioned conjugate gradient without linesearch: a comparison with the half-quadratic approach for edge-preserving image restoration", in *Computational Imaging IV (IS&T/SPIE Symposium on Electronic Imaging)*, San Jose, CA, Jan. 2006.