TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC SƯ PHẠM TP HỒ CHÍ MINH
Tập 21, Số 3 (2024): 446-459
Website: https://journal.hcmue.edu.vn
ISSN: 2734-9918

HO CHI MINH CITY UNIVERSITY OF EDUCATION
JOURNAL OF SCIENCE
Vol. 21, No. 3 (2024): 446-459
https://doi.org/10.54607/hcmue.js.21.3.3927(2024)

Research Article

CONVERGENCE AND CONVERGENCE RATES OF DAMPED NEWTON METHODS

Dang Ngoc Do Quyen
Ho Chi Minh City University of Education, Vietnam
Corresponding author: Dang Ngoc Do Quyen – Email: 4501101091@student.hcmue.edu.vn
Received: August 29, 2023; Revised: October 05, 2023; Accepted: October 18, 2023

ABSTRACT

In this paper, we study the convergence and convergence rates of damped Newton algorithms for solving unconstrained optimization problems with twice continuously differentiable objective functions. Under the assumption of the positive definiteness of the Hessian matrix of the objective function on an open set containing the level set corresponding to the value of the objective function at the starting point, we prove that the sequence generated by the damped Newton algorithm belongs to that open set, and the corresponding sequence of objective function values is monotonically decreasing. If the sequence has a limit point, that limit point is a locally strong minimum of the objective function, and the iterative sequence superlinearly globally converges to this minimizer. Furthermore, if the Hessian matrix of the objective function is Lipschitz continuous, the iterative sequence achieves the quadratic convergence rate.

Keywords: convergence rates; damped Newton algorithm; global convergence; positive-definiteness; quadratic; superlinear

Mathematics Subject Classification (2020) 49M15, 65K10, 41A25

1. Introduction

Consider the unconstrained optimization problem of the type

$$\text{minimize } f(x) \text{ subject to } x \in \mathbb{R}^n, \tag{1.1}$$

with a twice continuously differentiable (C^2-smooth) cost function $f : \mathbb{R}^n \to \mathbb{R}$. Newton's methods are the most effective methods to tackle such problems (1.1). Given a starting point $x_0 \in \mathbb{R}^n$, Newton algorithms generate the iterative sequences in the form of

$$x_{k+1} := x_k + t_k d_k \quad \text{for all } k \in \mathbb{N} := \{1, 2, \ldots\},$$

where $t_k \geq 0$ is a *step size* and $d_k \neq 0$ is a *Newton direction* at the k^{th} iteration (Beck, 2014; Boyd & Vandenberghe, 2004; Ben-Tal & Nemirovski, 1987; Bertsekas, 1999; Dennis &

Cite this article as: Dang Ngoc Do Quyen (2024). Convergence and its rates of damped Newton methods. *Ho Chi Minh City University of Education Journal of Science, 21*(3), 446-459.

Schnabel, 1987; James, 2014; Izmailov & Solodov, 2014; Nesterov, 2004; Nocedal & Wright, 1999; h, 1987; Ruszczyński, 2006). When t_k is chosen by the backtracking line search technique, such method is called the *damped Newton method* or *guarded Newton method*, to distinguish it from the *pure Newton method*, which uses a fixed step size $t = 1$.

To the best of our knowledge, there have been numerous types of research relevant to the damped Newton method (Beck, 2014; Boyd & Vandenberghe, 2004; James, 2014; Nesterov, 2004; Polyak, 1987). The literature review shows that they just derive the local results under the global assumptions. James (2014) claims that the uniformly positive definiteness of all Hessian matrices is required on the entire space $\mathbb{R}^n$. However, the results are just that all accumulation points of the iterative sequences generated by the damped Newton method are stationary points of the objective function, i.e. James (2014) has not confirmed the convergence of those iterative sequences yet. Although Nesterov (2004) only used Hessian matrices to be uniformly positive-definite at the optimal solution, it requires an additional assumption of Lipschitz continuity of all Hessian matrices over $\mathbb{R}^n$ and also gets only the local convergence of the damped Newton algorithm. The books by Boyd and Vandenberghe (2004) and Polyak (1987) have already shown the global convergence of the iterative sequences. However, the results from these studies were obtained under some very strong assumptions that the objective function must be strongly convex and its corresponding Hessian matrices are obliged to be Lipschitz continuous on the whole space $\mathbb{R}^n$.

Motivated by these works, we attempt to set some weaker assumptions than those in previous studies (Beck, 2014; Boyd & Vandenberghe, 2004; James, 2014; Nesterov, 2004; Polyak, 1987), and also achieve the global convergence and convergence rates of the damped Newton algorithm. Specifically, we verify that if there exists an open set containing the level set corresponding to the objective function value at the initial point such that the Hessian matrix of the objective function is positive-definite over that set, the damped Newton algorithm generates a sequence belonging to that open set. Additionally, the sequence of objective function values corresponding to this sequence is monotonically decreasing. If the sequence possesses a limit point, that point is a locally strong minimum of the objective function, and the iterative sequence converges superlinearly to this minimizer on a global scale. Furthermore, the iterative sequence achieves a quadratic rate of convergence if the Lipschitz continuity of the Hessian matrix of the objective function is guaranteed.

The rest of the paper is organized as follows. In the next section, we introduce some basic notions of locally strongly convex functions, strong local minimizers, the rates of convergence of the iterative sequences, and clarify essential lemmas for the main results. Section 3 presents the main results of this paper which are the global convergence and the convergence rates of the damped Newton algorithm.

2. Preliminaries

Let $(\mathbb{R}^n, \|\cdot\|)$ be an Euclid space. The open and closed balls with center $\bar{x} \in \mathbb{R}^n$ and radius δ are denoted by $B(\bar{x}, \delta)$ and $\bar{B}(\bar{x}, \delta)$, respectively. We recall the notions of the locally strongly convex functions and the strong local minimizers, which are used throughout this paper.

Definition 2.1. **(Locally strongly convex functions).** The function $f : \mathbb{R}^n \to \mathbb{R}$ is called to be **locally strongly convex around** $\bar{x} \in \mathbb{R}^n$ with modulus $\alpha > 0$ if there exists $\delta > 0$ such that

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) - \frac{\alpha}{2} \lambda (1-\lambda) \|x - y\|^2,$$

for all $x, y \in B(\bar{x}, \delta)$ and $\lambda \in [0,1]$.

Definition 2.2. A point $\bar{x} \in \mathbb{R}^n$ is called **a strong local minimizer** of $f : \mathbb{R}^n \to \mathbb{R}$ with modulus $\alpha > 0$ if there exists $\delta > 0$ such that

$$f(x) \geq f(\bar{x}) + \frac{\alpha}{2} \|x - \bar{x}\|^2, \quad \forall x \in B(\bar{x}, \delta).$$

For the class C^2-smooth functions, the two aforementioned definitions are equivalent.

Lemma 2.3. **(The characterization of strong local minimizers).** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2-smooth function and $\bar{x} \in \mathbb{R}^n$ such that $\nabla f(\bar{x}) = 0$. Then f is locally strongly convex around $\bar{x}$ if and only if $\bar{x}$ is a strong local minimizer of f with the same modulus.

Remark 2.4. Suppose that f is locally strongly convex around $\bar{x}$ with respect to $B(\bar{x}, \delta)$. If $\tilde{x} \in B(\bar{x}, \delta)$ is a strong local minimizer of f, then $\tilde{x} = \bar{x}$.

To guarantee the gradient mapping being Lipschitz continuous, we provide a necessary and sufficient condition in the lemma below.

Lemma 2.5. Let $C \subset \mathbb{R}^n$ be a nonempty compact set and $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function. Then ∇f is Lipschitz continuous on C, i.e., there exists $L > 0$ such that

$$\|\nabla f(y) - \nabla f(x)\| \leq L \|y - x\|, \quad \forall x, y \in C.$$

The next lemma provides some estimates of the values and the gradient mappings of a C^2-smooth function around its strong local minimizer.

Lemma 2.6. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2-smooth function and $\bar{x} \in \mathbb{R}^n$ be its strong local minimizer with modulus $\alpha > 0$. Then there exists $\delta > 0$ such that

$$\frac{\alpha}{2} \|x - \bar{x}\|^2 \leq |f(x) - f(\bar{x})| \leq \frac{L}{2} \|x - \bar{x}\|^2, \quad \forall x \in B(\bar{x}, \delta), \tag{2.1}$$

$$\alpha\|x-\bar{x}\| \le \|\nabla f(x)\| \le L\|x-\bar{x}\|, \quad \forall x \in B(\bar{x},\delta), \tag{2.2}$$

where L is the Lipschitz constant of ∇f on $\bar{B}(\bar{x},\delta)$.

Proof. Since $\bar{x}$ is a strong local minimizer of f with modulus α, it follows from Lemma 2.3 that f is locally strongly convex with modulus α. Thus $\nabla f(\bar{x})=0$ and there exists $\delta>0$ such that f is strongly convex on $B(\bar{x},\delta)$ and

$$f(x)-f(\bar{x}) \ge \frac{\alpha}{2}\|x-\bar{x}\|^2, \quad \forall x \in B(\bar{x},\delta). \tag{2.3}$$

Since f is twice continuously differentiable on $\bar{B}(\bar{x},\delta)$, it follows from Lemma 2.5 that ∇f is Lipschitz continuous on $\bar{B}(\bar{x},\delta)$ with some constant $L>0$. Applying the descent lemma (see Lemma A.11 in (Izmailov & Solodov, 2014)), we obtain

$$\left|f(x)-f(\bar{x})\right| = \left|f(x)-f(\bar{x})-\langle\nabla f(\bar{x}),x-\bar{x}\rangle\right| \le \frac{L}{2}\|x-\bar{x}\|^2, \quad \forall x \in B(\bar{x},\delta). \tag{2.4}$$

Combining (2.3) and (2.4), we get (2.1). Due to (2.3) and the first-order characterizations of strong convexity Theorem 5.24 in (Beck, 2017) together with Cauchy-Schwarz inequality, we have

$$\|\nabla f(x)-\nabla f(\bar{x})\|.\|x-\bar{x}\| \ge \langle\nabla f(x)-\nabla f(\bar{x}),x-\bar{x}\rangle \ge \alpha\|x-\bar{x}\|^2, \quad \forall x \in B(\bar{x},\delta). \tag{2.5}$$

Since $\nabla f(\bar{x})=0$ and ∇f is Lipschitz continuous on $\bar{B}(\bar{x},\delta)$ with modulus $L>0$, we get

$$\|\nabla f(x)\| = \|\nabla f(x)-\nabla f(\bar{x})\| \le L\|x-\bar{x}\|, \quad \forall x \in B(\bar{x},\delta). \tag{2.6}$$

Combining (2.5) and (2.6), we achieve (2.2). □

Next, we consider some notable rates of convergence.

Definition 2.7. **(Rates of convergence).** Let $\{x_k\} \subset \mathbb{R}^n$ be a sequence of vectors converging to $\bar{x}$ as $k \to \infty$ with $x_k \ne \bar{x}$ for all $k \in \mathbb{N}$. The convergence rate is said to be (at least)

(i) **superlinear** if we have

$$\lim_{k\to\infty}\frac{\|x_{k+1}-\bar{x}\|}{\|x_k-\bar{x}\|}=0.$$

(ii) **quadratic** if there exists $\beta>0$ such that

$$\frac{\|x_{k+1}-\bar{x}\|}{\|x_k-\bar{x}\|^2} \le \beta$$

whenever k is sufficiently large.

Lemma 2.6 allows us to verify the rates of convergence of two sequences $\{f(x_k)\}$ and $\{\nabla f(x_k)\}$ based on the rates of convergence of $\{x_k\}$.

Lemma 2.8. Let $f:\mathbb{R}^n \to \mathbb{R}$ be a C^2-smooth function and $\overline{x} \in \mathbb{R}^n$ be its strong local minimizer. Suppose that the sequence $\{x_k\}$ converges to $\overline{x}$. Then the two following assertions hold

(i) If $\{x_k\}$ converges superlinearly to $\overline{x}$, the sequences $\{f(x_k)\}$ and $\{\nabla f(x_k)\}$ converge superlinearly to $f(\overline{x})$ and 0, respectively.

(ii) If $\{x_k\}$ converges quadratically to $\overline{x}$, the sequences $\{f(x_k)\}$ and $\{\nabla f(x_k)\}$ converge quadratically to $f(\overline{x})$ and 0, respectively.

Proof. (i) By the convergence $x_k \to \overline{x}$ we have that $x_k \in B(\overline{x},\delta)$ for all k sufficiently large. Following from (2.1), we can deduce that

$$\frac{\left|f(x_{k+1})-f(\overline{x})\right|}{\left|f(x_k)-f(\overline{x})\right|} \le \frac{L}{\alpha}\frac{\|x_{k+1}-\overline{x}\|^2}{\|x_k-\overline{x}\|^2}, \quad \forall k \text{ is sufficiently large}.$$

The superlinear convergence of $\{x_k\}$ to $\overline{x}$ implies that $\lim\limits_{k\to\infty}\frac{\|x_{k+1}-\overline{x}\|}{\|x_k-\overline{x}\|}=0$. Therefore

$$\lim_{k\to\infty}\frac{\left|f(x_{k+1})-f(\overline{x})\right|}{\left|f(x_k)-f(\overline{x})\right|}=0,$$

and hence the sequence $\{f(x_k)\}$ converges superlinearly to $f(\overline{x})$.

Next, we prove the superlinear convergence of $\{\nabla f(x_k)\}$ to 0. Inequalities (2.2) give us

$$\frac{\|\nabla f(x_{k+1})\|}{\|\nabla f(x_k)\|} \le \frac{L}{\alpha}\frac{\|x_{k+1}-\overline{x}\|}{\|x_k-\overline{x}\|}, \quad \forall k \text{ is sufficiently large}.$$

Combining with $\lim\limits_{k\to\infty}\frac{\|x_{k+1}-\overline{x}\|}{\|x_k-\overline{x}\|}=0$, we obtain

$$\lim_{k\to\infty}\frac{\|\nabla f(x_{k+1})\|}{\|\nabla f(x_k)\|}=0,$$

and hence the sequence $\{\nabla f(x_k)\}$ converges superlinearly to 0.

(ii) Inequalities (2.1) bring us

$$\frac{\left|f(x_{k+1})-f(\overline{x})\right|}{\left|f(x_k)-f(\overline{x})\right|^2} \le \frac{2L}{\alpha^2}\frac{\|x_{k+1}-\overline{x}\|^2}{\|x_k-\overline{x}\|^4}, \quad \forall k \text{ is sufficiently large}.$$

The quadratic convergence of $\{x_k\}$ to $\bar{x}$ implies that there exists $\beta > 0$ such that $\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|^2} \le \beta$ whenever k is sufficiently large. Thus

$$\frac{|f(x_{k+1}) - f(\bar{x})|}{|f(x_k) - f(\bar{x})|^2} \le \frac{2L}{\alpha^2}\beta^2$$

for all large $k \in \mathbb{N}$, and hence the sequence $\{f(x_k)\}$ converges quadratically to $f(\bar{x})$.

Next, we indicate that the sequence $\{\nabla f(x_k)\}$ converges quadratically to 0. According to (2.2), we have

$$\frac{\|\nabla f(x_{k+1})\|}{\|\nabla f(x_k)\|^2} \le \frac{L}{\alpha^2}\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|^2}, \quad \forall k \text{ is sufficiently large}.$$

The quadratic convergence of $\{x_k\}$ to $\bar{x}$ ensures that

$$\frac{\|\nabla f(x_{k+1})\|}{\|\nabla f(x_k)\|^2} \le \frac{L}{\alpha^2}\beta$$

for all large $k \in \mathbb{N}$, which verifies that the sequence $\{\nabla f(x_k)\}$ converges quadratically to 0. □

Now, we have enough necessary conditions to present and verify our main results.

3. Main results

We first recall the damped Newton algorithm (Beck, 2014; Boyd & Vandenberghe, 2004; Nesterov, 2004; Polyak, 1987; James, 2014) for solving (1.1).

***Algorithm 3.1.* (damped Newton algorithm).**

Input: $x_0 \in \mathbb{R}^n$, $\sigma \in \left(0, \frac{1}{2}\right)$, $\beta \in (0,1)$

1: **for** $k = 0,1,\ldots$ **do**

2: If $\|\nabla f(x_k)\| = 0$, stop; otherwise go to the next step

3: Choose $d_k \in \mathbb{R}^n$ such that $\nabla f(x_k) + \nabla^2 f(x_k) d_k = 0$

4: Set $t_k = 1$

5: **while** $f(x_k + t_k d_k) > f(x_k) + \sigma t_k \langle \nabla f(x_k), d_k \rangle$ **do**

6: set $t_k = \beta t_k$

7: **end while**

8: Set $x_{k+1} = x_k + t_k d_k$

9: **end for**

Remark 3.2. In Algorithm 3.1, if at k^{th} iteration we ensure that $\nabla^2 f(x_k) \succ 0$, $d_k = -\left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$ is a descent direction of f at x_k, which means that $\left\langle \nabla f(x_k), d_k \right\rangle < 0$. Based on Lemma 4.3 in Beck (2014), the backtracking line search procedure in step 5 terminates and an appropriate t_k is found.

Next, we establish the global convergence of Algorithm 3.1.

Theorem 3.3. **(The global convergence of damped Newton algorithm).** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2-smooth function on an open set containing Ω and let $x_0 \in \mathbb{R}^n$ be an arbitrary point such that $\nabla^2 f(x) \succ 0$ for every x in the level set*

$$\Omega := \operatorname{Lev}\left(f, f(x_0)\right) = \left\{x \in \mathbb{R}^n : f(x) \le f(x_0)\right\}.$$

Then Algorithm 3.1 with the initial point x_0 either stops after finitely many iterations or produces a sequence $\{x_k\} \subset \Omega$ such that the corresponding sequence $\{f(x_k)\}$ is monotonically decreasing. In addition, if the iterative sequence $\{x_k\}$ has a limit point $\bar{x}$, $\{x_k\}$ converges to $\bar{x}$, and $\bar{x}$ is a strong local minimizer of f.

Proof. The proof is split into the three following claims.

Claim 1: *Algorithm 3.1 either stops after finitely many iterations or produces a sequence $\{x_k\} \subset \Omega$ such that the corresponding sequence $\{f(x_k)\}$ is monotonically decreasing and $\left\langle \nabla f(x_k), d_k \right\rangle < 0$ for all $k \in \mathbb{N}$, where $\{d_k\}$ is a sequence generated in Step 3 of Algorithm 3.1.*

Indeed, if there exists $k_0 \in \mathbb{N}$ such that $\nabla f\left(x_{k_0}\right) = 0$ then Algorithm 3.1 stops at the k_0^{th} iteration. Thus we only consider Algorithm 3.1 generating the iterative sequence $\{x_k\}$ satisfied $\nabla f(x_k) \neq 0$ for all $k \in \mathbb{N}$. Then $x_k \neq \bar{x}$ for all $k \in \mathbb{N}$. Obviously $x_0 \in \Omega$, it follows that $\nabla^2 f(x_0) \succ 0$, which means that $\left\langle \nabla^2 f(x_0)u, u \right\rangle > 0,\ \forall u \in \mathbb{R}^n \setminus \left\{0_{\mathbb{R}^n}\right\}$, and hence $d_0 = -\left(\nabla^2 f(x_0)\right)^{-1} \nabla f(x_0) \neq 0_{\mathbb{R}^n}$. Thus $\left\langle \nabla^2 f(x_0) d_0, d_0 \right\rangle > 0$, which implies that $\left\langle \nabla f(x_0), d_0 \right\rangle = \left\langle -\nabla^2 f(x_0) d_0, d_0 \right\rangle < 0$. Therefore

$$f(x_1) = f(x_0 + t_0 d_0) \le f(x_0) + \sigma t_0 \left\langle \nabla f(x_0), d_0 \right\rangle < f(x_0),$$

where the first inequality is the existing condition of the backtracking line search technique, and hence $x_1 \in \Omega$. Using the inductive method and arguing similarly for the cases of $k = 2, 3, \ldots,$ we obtain

$$f\left(x_{k+1}\right) < f\left(x_k\right) \text{ for all } k \in N \tag{3.1}$$

and the sequence $\{x_k\} \subset \Omega$. This implies that $\{f(x_k)\}$ is monotonically decreasing and $\langle \nabla f(x_k), d_k \rangle < 0$ for all $k \in \mathbb{N}$.

Claim 2: *If $\bar{x}$ is a limit point of $\{x_k\}$, $\bar{x}$ is a strong local minimizer of f.*

Suppose that $\{x_k\}$ has a limit point $\bar{x}$. Since the set Ω is closed and $\{x_k\} \subset \Omega$, we get $\bar{x} \in \Omega$, and hence $\nabla^2 f(\bar{x}) \succ 0$. It follows from Proposition 4.6 in Chieu et al. (2017) that there exist positive numbers α and δ such that

$$\left\langle \nabla^2 f(x)u, u \right\rangle \geq \alpha \|u\|^2 \text{ for all } x \in B(\bar{x}, \delta), \text{ and } u \in \mathbb{R}^n. \tag{3.2}$$

Let $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$ converging to $\bar{x}$ and $\{t_{k_j}\}$ be a corresponding sequence of positive numbers generated in Algorithm 3.1.

- **Claim 2a:** *The sequence $\{t_{k_j}\}$ is bounded below by a positive number γ and we have*

$$f\left(x_{k_j}\right) - f\left(x_{k_j+1}\right) \geq \sigma\gamma\alpha \left\|d_{k_j}\right\|^2 \text{ for sufficiently large } j \in \mathbb{N}. \tag{3.3}$$

Suppose on the contrary that $\{t_{k_j}\}$ is not bounded below by a positive number. Then there exists a subsequence of $\{t_{k_j}\}$ that converges to 0. Assume without loss of generality that $t_{k_j} \to 0$ as $j \to \infty$. Since $x_{k_j} \to \bar{x}$, we have that $x_{k_j} \in B(\bar{x}, \delta)$ for all $j \in \mathbb{N}$ sufficiently large. Substituting $u = d_{k_j} = -\left(\nabla^2 f\left(x_{k_j}\right)\right)^{-1} \nabla f\left(x_{k_j}\right)$ and $x = x_{k_j}$ into (3.2), we get

$$\left\langle -\nabla f\left(x_{k_j}\right), d_{k_j} \right\rangle \geq \alpha \left\|d_{k_j}\right\|^2, \text{ for sufficiently large } j \in \mathbb{N}. \tag{3.4}$$

Applying the Cauchy-Schwarz inequality, we obtain

$$\left\|\nabla f\left(x_{k_j}\right)\right\| . \left\|d_{k_j}\right\| \geq \left|\left\langle \nabla f\left(x_{k_j}\right), d_{k_j} \right\rangle\right| \geq \left\langle -\nabla f\left(x_{k_j}\right), d_{k_j} \right\rangle, \quad \forall j \in \mathbb{N}. \tag{3.5}$$

Combining (3.4) and (3.5) verifies that

$$\frac{\left\|\nabla f\left(x_{k_j}\right)\right\|}{\alpha} \geq \left\|d_{k_j}\right\|, \text{ for sufficiently large } j \in \mathbb{N}.$$

Since $\{x_{k_j}\}$ converges to $\bar{x}$ and ∇f is continuous, the sequence $\left\{\nabla f\left(x_{k_j}\right)\right\}$ converges to $\nabla f(\bar{x})$, which implies that the sequence $\{d_{k_j}\}$ is bounded. Since $x_{k_j} \to \bar{x}$, $t_{k_j} \to 0$ and $\{d_{k_j}\}$ is bounded, we get $x_{k_j} + \beta^{-1} t_{k_j} d_{k_j} \to \bar{x}$ as $j \to \infty$, and hence $x_{k_j} + \beta^{-1} t_{k_j} d_{k_j} \in B(\bar{x}, \delta)$ whenever j is sufficiently large. Since f is a twice continuously differentiable function on

$\overline{B}(\overline{x},\delta)$, it follows from Lemma 2.5 that ∇f is Lipschitz continuous on $\overline{B}(\overline{x},\delta)$ with a constant $L>0$. Then the descent lemma (see Lemma A.11 in Izmailov and Solodov in 2014) gives us

$$f\left(x_{k_j}+\beta^{-1}t_{k_j}d_{k_j}\right)\le f\left(x_{k_j}\right)+\beta^{-1}t_{k_j}\left\langle\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle+\frac{L\beta^{-2}t_{k_j}^2}{2}\left\|d_{k_j}\right\|^2 \tag{3.6}$$

for sufficiently large $j\in\mathbb{N}$. According to Lemma 4.3 in Beck (2014), the backtracking line search in Step 5 of Algorithm 3.1 brings us t_{k_j} satisfied

$$f\left(x_{k_j}+\beta^{-1}t_{k_j}d_{k_j}\right)> f\left(x_{k_j}\right)+\sigma\beta^{-1}t_{k_j}\left\langle\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle, \text{for sufficiently large } j\in\mathbb{N}. \tag{3.7}$$

Combining (3.4), (3.6), and (3.7), for all $j\in\mathbb{N}$ is sufficiently large, we have

$$\sigma\beta^{-1}t_{k_j}\left\langle\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle<\beta^{-1}t_{k_j}\left\langle\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle+\frac{L\beta^{-2}t_{k_j}^2}{2\alpha}\left\langle\nabla f\left(x_{k_j}\right),-d_{k_j}\right\rangle. \tag{3.8}$$

Dividing both sides of (3.8) by $\beta^{-1}t_{k_j}\left\langle\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle<0$, we get

$$\sigma>1-\frac{l}{2\alpha\beta}t_{k_j},\quad \text{for sufficiently large } j\in\mathbb{N}.$$

Taking $j\to\infty$, we obtain $\sigma\ge 1$, which is a contradiction to the choice of $\sigma<1$. Thus $\left\{t_{k_j}\right\}$ is bounded below by $\gamma>0$. Moreover, using the exit condition of backtracking line search and the estimate in (3.4) allows us to indicate that the below inequalities hold whenever j is sufficiently large

$$f\left(x_{k_j}\right)-f\left(x_{k_j+1}\right)\ge\sigma t_{k_j}\left\langle-\nabla f\left(x_{k_j}\right),d_{k_j}\right\rangle\ge\sigma\gamma\alpha\left\|d_{k_j}\right\|^2.$$

The proof of Claim 2a is completed.

- **Claim 2b:** *$\overline{x}$ is a strong local minimizer of f.*

Since $\left\{f\left(x_k\right)\right\}$ is monotonically decreasing and $f\left(\overline{x}\right)$ is a limit point of $\left\{f\left(x_k\right)\right\}$, the sequence $\left\{f\left(x_k\right)\right\}$ must converge to $f\left(\overline{x}\right)$. Letting $j\to\infty$ in (3.3), we obtain $\left\|d_{k_j}\right\|\to 0$. Since ∇f is Lipschitz continuous on $B(\overline{x},\delta)$ with modulus $L>0$, it follows from Theorem 2.1.6 in Nesterov (2004) that $\left\|\nabla^2 f\left(x\right)\right\|\le L$ for all $x\in B(\overline{x},\delta)$. Since $x_{k_j}\in B(\overline{x},\delta)$ for all large j, with $d_{k_j}=-\left(\nabla^2 f\left(x_{k_j}\right)\right)^{-1}\nabla f\left(x_{k_j}\right)$ we obtain

$$\left\|\nabla f\left(x_{k_j}\right)\right\|=\left\|\nabla^2 f\left(x_{k_j}\right)d_{k_j}\right\|\le\left\|\nabla^2 f\left(x_{k_j}\right)\right\|\left\|d_{k_j}\right\|\le L\left\|d_{k_j}\right\|,$$

for sufficiently large $j\in\mathbb{N}$. Passing to the limit as $j\to\infty$ in this inequality tells us that $\nabla f\left(\overline{x}\right)=0$. Following from (3.2) and combining with Theorem 4.3.1 in Hiriart-Urruty and Lemaréchal (2004) give us f is α-strongly convex on $B(\overline{x},\delta)$. The first-order characterizations of strong convexity (Theorem 5.24 in Beck (2017)) and $\nabla f(\overline{x})=0$ bring us

$$f(x)\geq f(\overline{x})+\frac{\alpha}{2}\|x-\overline{x}\|^2,\ \ \forall x\in B(\overline{x},\delta).$$

This verifies that $\overline{x}$ is a strong local minimizer.

Claim 3: *The iterative sequence $\{x_k\}$ converges to $\overline{x}$.*

- **Claim 3.1:** *The sequence $\{x_k\}$ has no other limit point other than $\overline{x}$ in $B(\overline{x},\delta)$.*

Suppose that there exists $\tilde{x}\in B(\overline{x},\delta)$ such that $\tilde{x}$ is a limit point of $\{x_k\}$. We previously proved that all limit points of the sequence $\{x_k\}$ are the strongly local minimizes of f. Thus, we can deduce that $\tilde{x}$ is also a strong local minimizer of f on $B(\overline{x},\delta)$. Remark 2.4 gives us $\tilde{x}=\overline{x}.$ The proof of Claim 3.1 is completed.

- **Claim 3.2:** *The sequence $\{x_k\}$ converges to $\overline{x}$.*

Supposing that $\{x_{k_j}\}$ is an arbitrary subsequence of $\{x_k\}$ with $x_{k_j}\to\overline{x}$ as $j\to\infty$. We have $x_{k_j+1}=x_{k_j}+t_{k_j}d_{k_j}$ combined with (3.3), we obtain

$$\left\|x_{k_j+1}-x_{k_j}\right\|^2=t_{k_j}^2\left\|d_{k_j}\right\|^2\leq\left\|d_{k_j}\right\|^2\leq\frac{f(x_{k_j})-f(x_{k_j+1})}{\sigma\gamma\alpha},\ \text{ for sufficiently large } j\in\mathbb{N}.$$

The convergence of $\{f(x_k)\}$ to $f(\overline{x})$ gives us $\lim_{j\to\infty}\left\|x_{k_j+1}-x_{k_j}\right\|=0.$ Then Proposition 8.3.10 in Facchinei and Pang (2003) gives us that the sequence $\{x_k\}$ converges to $\overline{x}.$ □

Remark 3.4. In Theorem 3.3, we obtained some better results than those in Corollary 6.2.3 by James (2014). In terms of assumptions, Corollary 6.2.3 requires the existence of $\beta>0$ such that $\nabla^2 f(x)\succcurlyeq\beta I$ on Ω, we just need $\nabla^2 f(x)\succ 0$ on Ω. Thus our assumptions are much weaker than those in Corollary 6.2.3. Despite using weaker assumptions, we still achieve some stronger results. Besides proving all limit points are stationary points of f, we additionally clarify that $\{x_k\}$ converges to $\overline{x}$ and $\overline{x}$ is a strong local minimizer of f.

Theorem 3.5. **(Convergence rates of damped Newton algorithm).** *In the setting of Theorem 3.3 and $\overline{x}$ as a limit point of the iterative sequence $\{x_k\}$ generated by Algorithm 3.1, the following statements hold*

(a)The sequence $\{x_k\}$, the value sequence $\{f(x_k)\}$ and the gradient sequence $\{\nabla f(x_k)\}$ converge superlinearly to $\overline{x}$, $f(\overline{x})$ and 0, respectively.

(b)Suppose in addition that $\nabla^2 f$ *is Lipschitz continuous with some constant* $M > 0$*, then all convergence rates in (a) are quadratic.*

Proof. Due to Theorem 3.3, $\{x_k\}$ converges to $\bar{x}$ and $\nabla f(\bar{x}) = 0$. Since $-\nabla f(x_k) = \nabla^2 f(x_k) d_k$, we get

$$\nabla^2 f(x_k)(x_k + d_k - \bar{x}) = -\nabla f(x_k) - \nabla^2 f(x_k)(-x_k + \bar{x}), \quad \forall k \in \mathbb{N}. \tag{3.9}$$

Substituting $u = x_k + d_k - \bar{x}$ and $x = x_k$ into (3.2) and using Cauchy-Schwarz inequality together with (3.9), we obtain

$$\|x_k + d_k - \bar{x}\| \le \frac{1}{\alpha}\|\nabla f(x_k) + \nabla^2 f(x_k)(-x_k + \bar{x})\| \text{ for sufficiently large } k \in \mathbb{N}. \tag{3.10}$$

Since ∇f is differentiable at $\bar{x}$ and $\nabla f(\bar{x}) = 0$, it follows from Lemma 5.5 in Pham et al. (2022) that

$$\|\nabla f(x_k) + \nabla^2 f(x_k)(-x_k + \bar{x})\| = \|\nabla f(x_k) - \nabla f(\bar{x}) + \nabla^2 f(x_k)(-x_k + \bar{x})\| = o(\|x_k - \bar{x}\|).$$

Combining this with (3.10), we have

$$\|x_k + d_k - \bar{x}\| = o(\|x_k - \bar{x}\|). \tag{3.11}$$

Substituting $u = d_k = -\left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$ and $x = x_k$ into (3.2), we obtain

$$\langle -\nabla f(x_k), d_k \rangle \ge \alpha \|d_k\|^2 \text{ for all large } k \in \mathbb{N}. \tag{3.12}$$

We have already proved that the sequence $\{x_k\}$ converges to $\bar{x}$, $\nabla f(\bar{x}) = 0$, the sequence $\{d_k\}$ satisfies (3.11) and (3.12). Therefore, following from Proposition 8.3.18 in Facchinei and Pang (2003), we get

$$f(x_k + d_k) \le f(x_k) + \sigma \langle \nabla f(x_k), d_k \rangle \text{ for all large } k,$$

which means that all t_k chosen by backtracking line search technique always equals 1 whenever k is sufficiently large. Then we have

$$\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} = \frac{\|x_k + d_k - \bar{x}\|}{\|x_k - \bar{x}\|} \text{ whenever } k \text{ is sufficiently large} \tag{3.13}$$

a) Combining (3.11) and (3.13), we obtain $\lim\limits_{k\to\infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} = 0$, which means that the sequence $\{x_k\}$ converges superlinearly to $\bar{x}$. Following from Lemma 2.8, two sequences $\{f(x_k)\}$ and $\{\nabla f(x_k)\}$ converge superlinearly to $f(\bar{x})$ and 0, respectively.

b) Since f is twice continuously differentiable and $\nabla^2 f$ is Lipschitz continuous with some constant $M > 0$ on $\bar{B}(\bar{x}, \delta)$, it follows from the descent lemma (see Lemma A.11 in Izmailov and Solodov (2014)) that

$$\left\|\nabla f(x_k) - \nabla f(\bar{x}) - \nabla^2 f(\bar{x})(x_k - \bar{x})\right\| \le \frac{M}{2}\|x_k - \bar{x}\|^2 \text{ for sufficiently large } k. \quad (3.14)$$

The inclusion $x_k \in \bar{B}(\bar{x}, \delta)$ for all k sufficiently large and the Lipschitz continuity of $\nabla^2 f$ on $\bar{B}(\bar{x}, \delta)$ ensure that

$$\left\|\nabla^2 f(x_k) - \nabla^2 f(\bar{x})\right\| \le M\|x_k - \bar{x}\| \text{ for sufficiently large } k. \quad (3.15)$$

Using the Cauchy–Schwartz inequality, $\|AB\| \le \|A\|.\|B\|$ and combining (3.14) with (3.15), we obtain

$$\begin{aligned}
&\left\|\nabla f(x_k) - \nabla f(\bar{x}) - \nabla^2 f(x_k)(x_k - \bar{x})\right\| \\
&\le \left\|\nabla f(x_k) - \nabla f(\bar{x}) - \nabla^2 f(\bar{x})(x_k - \bar{x})\right\| + \left\|\nabla^2 f(x_k) - \nabla^2 f(\bar{x})\right\|.\|x_k - \bar{x}\| \\
&\le \frac{M}{2}\|x_k - \bar{x}\|^2 + M\|x_k - \bar{x}\|^2 = \frac{3M}{2}\|x_k - \bar{x}\|^2, \quad \forall k \in \mathbb{N} \text{ is sufficiently large.}
\end{aligned}$$

Combining this with $\nabla f(\bar{x}) = 0$ and (3.10) gives us

$$\|x_k + d_k - \bar{x}\| \le \frac{3M}{2\alpha}\|x_k - \bar{x}\|^2 \text{ for sufficiently large } k. \quad (3.16)$$

(3.13) and (3.16) ensure that the sequence $\{x_k\}$ converges quadratically to $\bar{x}$. Following Lemma 2.8, two sequences $\{f(x_k)\}$ and $\{\nabla f(x_k)\}$ converge quadratically to $f(\bar{x})$ and 0, respectively. □

❖ ***Conflict of Interest:*** *Author have no conflict of interest to declare.*

❖ ***Acknowledgement:*** *This research is funded by Ho Chi Minh City University of Education Foundation for Science and Technology under grant number CS.2022.19.20TĐ.*

REFERENCES

Beck, A. (2014). *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM. https://doi.org/10.1137/1.9781611973655

Beck, A. (2017). *First-Order Methods in Optimization*. SIAM. https://doi.org/10.1137/1.9781611974997

Ben-Tal, A., & Nemirovski, A. (1987). *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM. https://doi.org/10.1137/1.9780898718829

Bertsekas, D. P. (1999). *Nonlinear programming*. Athena Scientific. ISBN: 978-1-886529-05-2

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804441

Chieu, N. H., Lee, G. M., & Yen, N. D. (2017). Second-order subdifferentials and optimality conditions for C^1-smooth optimization problems. *Applied Analysis and Optimization*, *1*(3), 461-476.

Dennis, J. E., & Schnabel, R. B. (1987). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Pennsylvania. SIAM. https://doi.org/10.1137/1.9781611971200

Facchinei, F., & Pang, J.-S. (2003). *Finite-dimensional Variational Inequalities and Complementarity Problems, volume II*. Springer. https://doi.org/10.1007/b97544

Hiriart-Urruty, J.-B., & Lemaréchal, C. (2004). *Fundamentals of Convex Analysis*. Springer. https://doi.org/10.1007/978-3-642-56468-0

Izmailov, A., & Solodov, M. (2014). *Newton-Type Methods for Optimization and Variational Problems*. Springer. https://doi.org/10.1007/978-3-319-04247-3

James, V. B. (2014). *Nonlinear Optimization*. University of Washington.

Mordukhovich, B. S., & Nam, N. M. (2014). *An Easy Path to Convex Analysis and Applications*. Morgan & Claypool Publishers. https://doi.org/10.1007/978-3-031-02406-1

Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization: A basic course*. Springer US. https://doi.org/10.1007/978-1-4419-8853-9

Nocedal, J., & Wright, S. (1999). *Numerical optimization*. Springer. https://doi.org/10.1007/978-0-387-40065-5

Pham, D. K., Mordukhovich, B. S., & Vo, T. P.(2022). A generalized Newton method for subgradient systems. *Mathematics of Operations Research, 48*(4). https://doi.org/10.1287/moor.2022.1320

Polyak, B. T. (1987). *Introduction to Optimization*. Optimization Software.

Rockafellar, R. T., & Wets, R.-B. (1998). *Variational Analysis*. Springer. https://doi.org/10.1007/978-3-642-02431-3

Ruszczyński, A. (2006). *Nonlinear Optimization*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4hcj

SỰ HỘI TỤ VÀ TỐC ĐỘ HỘI TỤ CỦA CÁC PHƯƠNG PHÁP DAMPED NEWTON

Đặng Ngọc Đỗ Quyên

Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam

Tác giả liên hệ: Đặng Ngọc Đỗ Quyên – Email: 4501101091@student.hcmue.edu.vn

Ngày nhận bài: 29-8-2023; ngày nhận bài sửa: 05-10-2023; ngày duyệt đăng: 18-10-2023

TÓM TẮT

Trong bài báo này, chúng tôi nghiên cứu sự hội tụ và tốc độ hội tụ của các thuật toán damped Newton để giải các bài toán tối ưu không ràng buộc với các hàm mục tiêu khả vi liên tục cấp hai. Dưới giả thiết về tính xác định dương của ma trận Hessian của hàm mục tiêu trên một tập mở chứa tập mức ứng với giá trị hàm mục tiêu tại điểm khởi động, chúng tôi chứng minh dãy lặp sinh bởi thuật toán damped Newton sẽ nằm trong tập mở đó và dãy giá trị hàm tương ứng là đơn điệu giảm. Nếu dãy lặp có điểm tụ thì điểm tụ sẽ là điểm cực tiểu mạnh của hàm mục tiêu, và dãy lặp hội tụ toàn cục siêu tuyến tính về điểm cực tiểu này. Hơn nữa, nếu ma trận Hessian liên tục Lipschitz, dãy lặp đạt được tốc độ hội tụ bậc hai.

Từ khóa: các tốc độ hội tụ; thuật toán damped Newton; sự hội tụ toàn cục; tính xác định dương; bậc hai; siêu tuyến tính