Applied Mathematics and Computation 491 (2025) 129219

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Full Length Article

Convergence analysis of a regularized Newton method with generalized regularization terms for unconstrained convex optimization problems

Yuya Yamakawa *, Nobuo Yamashita

Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

A R T I C L E   I N F O

Keywords:
Unconstrained convex optimization
Regularized Newton method
Generalized regularization
Global $\mathcal{O}(k^{-2})$ convergence
Superlinear convergence
Local convergence

A B S T R A C T

This paper presents a regularized Newton method (RNM) with generalized regularization terms for unconstrained convex optimization problems. The generalized regularization includes quadratic, cubic, and elastic net regularizations as special cases. Therefore, the proposed method serves as a general framework that includes not only the classical and cubic RNMs but also a novel RNM with elastic net regularization. We show that the proposed RNM has the global $\mathcal{O}(k^{-2})$ and local superlinear convergence, which are the same as those of the cubic RNM.

1. Introduction

We consider the following unconstrained convex optimization problem:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x), \tag{1}$$

where the function f is twice continuously differentiable and convex on $\mathbb{R}^n$. Moreover, we assume that problem (1) has a minimizer x^* and define $f^* := f(x^*)$.

Newton's method is one of the most well-known and basic iterative methods for solving unconstrained convex optimization problems. Each iteration computes a search direction d_k, which is a solution of the following subproblem:

$$\underset{d\in\mathbb{R}^n}{\text{minimize}} \quad \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle,$$

and updates the current point x_k as $x_{k+1} := x_k + t_k d_k$, where $t_k > 0$ denotes a step size. It converges rapidly thanks to the use of the second-order information, that is, $\nabla^2 f(x_k)$, of the objective function. However, it requires that $\nabla^2 f(x_k)$ is nonsingular at each iteration. Even if the Hessian is nonsingular, the convergence rate may be reduced to linear when the Hessian is close to singular. Several variants of Newton's method have been proposed to overcome these drawbacks including regularized Newton methods (RNMs) [1–5], cubic RNMs [6–8], and so forth [9–18].

RNMs can be considered modifications of Newton's method because they improve the subproblem of Newton's method such that it can be solved even if the Hessian matrix is singular. More precisely, RNMs iteratively solve the following subproblem to find a search direction d_k:

* Corresponding author.
E-mail address: yuya@i.kyoto-u.ac.jp (Y. Yamakawa).

https://doi.org/10.1016/j.amc.2024.129219
Received 11 July 2024; Received in revised form 22 November 2024; Accepted 23 November 2024

Available online 2 December 2024

$$\underset{d\in\mathbb{R}^n}{\text{minimize}} \quad \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle + \frac{\mu_k}{2}\|d\|^2,$$

where $\mu_k > 0$ denotes a parameter. Since the objective function is strongly convex, the subproblem has a unique optimum. Nesterov and Polyak [6] proposed an RNM with the cubic regularization $\frac{\mu_k}{6}\|d\|^3$. The proposed method is called the cubic RNM and iteratively solves the following subproblem:

$$\underset{d\in\mathbb{R}^n}{\text{minimize}} \quad \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle + \frac{\mu_k}{6}\|d\|^3.$$

For classical and cubic RNMs, global $\mathcal{O}(k^{-2})$ and local superlinear convergence were proven in [2,4,6,5].

For least squares problems, Ariizumi, Yamakawa, and Yamashita [19] recently proposed a Levenberg-Marquardt method (LMM) equipped with a generalized regularization term and showed its global and local superlinear convergence. Although a subproblem of the classical LMM has a quadratic regularization term $\frac{\mu_k}{2}\|d\|^2$, they generalized the regularization term such that another regularization can be adopted, such as the L^1 and elastic-net regularization. Moreover, they reported numerical experiments in which their LMM with the elastic-net regularization worked well for certain examples, owing to the sparsity of the search direction.

Inspired by Ariizumi, Yamakawa, and Yamashita [19], we propose a generalized RNM (GRNM) for solving problem (1). We adopt new regularization terms provided as $\frac{\mu_k}{p}\|d\|_2^p + \rho_k\|d\|_1$, where $p \in (1,3]$ is a pre-fixed parameter. With the addition of the new regularization terms, the GRNM includes classical and cubic RNMs as well as novel RNMs with other regularizations, such as the elastic net and so forth. More precisely, if $p=2$ and $\rho_k = 0$, the GRNM is reduced to the classical RNM; if $p=3$ and $\rho_k=0$, it is equivalent to the cubic RNM. Moreover, if $p=2$ and $\rho_k>0$, it can be regarded as a novel RNM with the elastic-net regularization.

The contributions of this study are as follows. This study provides

(i) the generalized RNM stated above;
(ii) sufficient conditions of p, μ_k, and ρ_k for which the GRNM has the global $\mathcal{O}(k^{-2})$ convergence;
(iii) local superlinear convergence under the local error bound condition.

Hence, these contributions include classical and cubic RNMs as special cases and provide a framework for new RNMs such as the elastic-net RNM.

The remainder of this paper is organized as follows. Section 2 provides a general proposition that plays an important role in the analysis of global $\mathcal{O}(k^{-2})$ convergence. Section 3 describes the proposed method. Section 4 presents global $\mathcal{O}(k^{-2})$ convergence of the proposed method. Section 5 proves local and superlinear convergence. In section 6, we conduct numerical experiments to confirm the performance of the proposed method by using several regularization terms. Finally, concluding remarks are presented in Section 7.

Throughout the paper, we use the following mathematical notation. Let $\mathbb{N}$ be the set of natural numbers (positive integers). For $p\in\mathbb{N}$ and $q\in\mathbb{N}$, the set of real matrices with p rows and q columns is denoted by $\mathbb{R}^{p\times q}$. Note that $\mathbb{R}^{p\times 1}$ is equal to the set of p-dimensional real vectors, that is, $\mathbb{R}^{p\times 1} = \mathbb{R}^p$, and note that $\mathbb{R}^1$ represents the set of real numbers, namely, $\mathbb{R}^1 = \mathbb{R}$. For any $w\in\mathbb{R}^p$, the transposition of w is represented as $w^\top \in \mathbb{R}^{1\times p}$. For $u\in\mathbb{R}^p$ and $v\in\mathbb{R}^p$, the inner product of u and v is defined by $\langle u,v\rangle := u^\top v$. We denote by I the identity matrix, where these dimensions are defined by the context. For each $w\in\mathbb{R}^p$, the Euclidean and L^1 norms of w are respectively defined by $\|w\| := \sqrt{\langle w,w\rangle}$ and $\|w\|_1 := |[w]_1| + |[w]_2| + \cdots + |[w]_p|$, where $[w]_j$ represents the j-th element of w. For $W\in\mathbb{R}^{p\times q}$, we denote by $\|W\|$ the operator norm of W, that is, $\|W\| := \sup\{\|Wu\|; \|u\|\le 1\}$. Let φ be a function from $\mathbb{R}^p$ to $\mathbb{R}$. The gradient of φ at $w\in\mathbb{R}^p$ is represented as $\nabla\varphi(w)$. The Hessian of φ at $w\in\mathbb{R}^p$ is denoted by $\nabla^2\varphi(w)$. For a convex function $\phi: \mathbb{R}^p\to\mathbb{R}$, we denote by $\partial\phi(w)$ the subdifferential of ϕ at w. For $\eta\in\mathbb{R}^p$ and $r>0$, we define $B(\eta,r) := \{\mu\in\mathbb{R}^p; \|\mu-\eta\|\le r\}$. For infinite sequences $\{a_k\}\subset\mathbb{R}$ and $\{b_k\}\subset\mathbb{R}$, we write $a_k = \mathcal{O}(b_k)$ $(k\to\infty)$ if there exist $c>0$ and $n\in\mathbb{N}$ such that $|a_k|\le c|b_k|$ for all $k\ge n$.

2. Preliminaries

This section presents a general proposition that provides sufficient conditions under which arbitrary sequences generated by optimization methods have global $\mathcal{O}(k^{-2})$ convergence. This proposition plays a critical role in Section 4. The proof of the proposition is inspired by the technique presented in [5, Theorem 1].

Proposition 1. *Suppose that $f: \mathbb{R}^n\to\mathbb{R}$ is a continuously differentiable convex function. Let $\{x_k\}\subset\mathbb{R}^n$ be an infinite sequence. Suppose also that*

(i) *$f(x_{k+1}) \le f(x_k)$ for all $k\in\mathbb{N}\cup\{0\}$;*
(ii) *there exists $\gamma>0$ such that $\|\nabla f(x_{k+1})\| \le \gamma\|\nabla f(x_k)\|$ for all $k\in\mathbb{N}\cup\{0\}$;*
(iii) *there exists $\delta>0$ such that $f(x_k)-f^* \le \delta\|\nabla f(x_k)\|$ for all $k\in\mathbb{N}\cup\{0\}$;*
(iv) *there exist $\theta\in(0,1)$, $\nu>0$, and $\ell\in\mathbb{N}$ such that $\theta^k \le \nu k^{-2}$ and $(\gamma\theta)^{\frac{k}{2}} \le \nu k^{-2}$ for all $k\ge\ell$, where γ is given in* (ii).

Let $\mathcal{I}(\theta) := \{i\in\mathbb{N}\cup\{0\}; \theta\|\nabla f(x_i)\| \le \|\nabla f(x_{i+1})\|\}$. Suppose also that

(v) *there exists $\tau>0$ such that $f(x_{k+1}) - f(x_k) \le -\tau(f(x_k)-f^*)^{\frac{3}{2}}$ for all $k\in\mathcal{I}(\theta)$.*

Then, one of the following statements holds:

(a) *If* $|\mathcal{I}(\theta)| < \infty$ *holds, then*

$$f(x_k) - f^* \le \frac{\theta^{-(\hat{i}+1)} \nu\delta \|\nabla f(x_{\hat{i}+1})\|}{k^2} \quad \forall k \ge \max\{\ell, \hat{i}+2\},$$

where $\hat{i}$ *is the largest element of* $\mathcal{I}(\theta)$.

(b) *If* $|\mathcal{I}(\theta)| = \infty$ *holds, then*

$$f(x_k) - f^* \le \max\left\{ \frac{36\tau^{-2}}{(k+4)^2}, \frac{\nu\delta\|\nabla f(x_0)\|}{k^2} \right\} \quad \forall k \ge \ell.$$

Proof. Let $i_\ell \in \mathbb{N}$ be the ℓ-th smallest element of $\mathcal{I}(\theta)$, that is, $\mathcal{I}(\theta) = \{i_1, i_2, \dots, \hat{i}\}$ with $i_\ell < i_{\ell+1}$. Moreover, regarding assumption (ii), we suppose $\gamma \ge 1$ without loss of generality.

To begin with, we consider case (a), that is, $|\mathcal{I}(\theta)| < \infty$ is satisfied. Let $\mathcal{J}(\theta) := \{i \in \mathbb{N} \cup \{0\}; \theta\|\nabla f(x_i)\| > \|\nabla f(x_{i+1})\|\}$. Note that $\hat{i} = i_{|\mathcal{I}(\theta)|}$. We can easily observe that

$$k \in \mathcal{J}(\theta) \quad \forall k > \hat{i}. \tag{2}$$

For every $k \ge \hat{i}+2$, it follows from (2) that $j \in \mathcal{J}(\theta)$ for any $j \in \{\hat{i}+1, \hat{i}+2, \dots, k-1\}$, that is,

$$\|\nabla f(x_k)\| < \theta\|\nabla f(x_{k-1})\| < \cdots < \theta^{k-(\hat{i}+1)}\|\nabla f(x_{\hat{i}+1})\|. \tag{3}$$

Let us take $k \ge \max\{\ell, \hat{i}+2\}$ arbitrarily. By assumption (iii) and (3), we obtain

$$f(x_k) - f^* \le \delta\|\nabla f(x_k)\| < \theta^{k-(\hat{i}+1)}\delta\|\nabla f(x_{\hat{i}+1})\|. \tag{4}$$

Recall that $\theta^k \le \nu k^{-2}$ from assumption (iv). Thus, the desired inequality is derived from (4).

Next, we discuss the case where $|\mathcal{I}(\theta)| = \infty$ holds. Let $k \in \mathbb{N} \cup \{0\}$ and define $\psi_k := \tau^2(f(x_{i_k}) - f^*)$. Combining assumption (i) and $i_{k+1} \ge i_k + 1$ yields

$$\psi_{k+1} = \tau^2(f(x_{i_{k+1}}) - f^*) \le \tau^2(f(x_{i_k+1}) - f^*). \tag{5}$$

Since assumption (v) implies that $f(x_{i_k+1}) - f(x_{i_k}) \le -\tau(f(x_{i_k}) - f^*)^{\frac{3}{2}}$,

$$\tau^2(f(x_{i_k+1}) - f^*) \le \tau^2(f(x_{i_k}) - f^*) - \tau^3(f(x_{i_k}) - f^*)^{\frac{3}{2}} = \psi_k - \psi_k^{\frac{3}{2}}. \tag{6}$$

Exploiting (5) and (6) derives $\psi_{k+1} \le \psi_k - \psi_k^{\frac{3}{2}} \le \psi_k - \frac{2}{3}\psi_k^{\frac{3}{2}}$. It then follows from [5, Proposition 1] that $\psi_k \le 9(k+2)^{-2}$, that is,

$$f(x_{i_k}) - f^* \le \frac{9\tau^{-2}}{(k+2)^2} \quad \forall k \in \mathbb{N} \cup \{0\}. \tag{7}$$

Let $\mathcal{I}_k := \{i \in \mathcal{I}(\theta); i \le k\}$. In the following, we assume $k \ge \ell$. There are two possible cases: Case (1) $|\mathcal{I}_k| \ge \frac{k}{2}$ and Case (2) $|\mathcal{I}_k| < \frac{k}{2}$.

Case (1): The largest element of $\mathcal{I}_k$ can be represented by $i_{|\mathcal{I}_k|}$, and thus $i_{|\mathcal{I}_k|} \le k$. This fact, assumption (i), and (7) imply that

$$f(x_k) - f^* \le f(x_{i_{|\mathcal{I}_k|}}) - f^* \le \frac{9\tau^{-2}}{(|\mathcal{I}_k|+2)^2} \le \frac{36\tau^{-2}}{(k+4)^2}.$$

Case (2): From assumption (ii), each $j \in \{0, 1, \dots, k-1\}$ satisfies

$$\begin{aligned} &\|\nabla f(x_{j+1})\| \le \gamma\|\nabla f(x_j)\| \quad \text{if } j \in \mathcal{I}_{k-1}, \\ &\|\nabla f(x_{j+1})\| < \theta\|\nabla f(x_j)\| \quad \text{if } j \notin \mathcal{I}_{k-1}. \end{aligned} \tag{8}$$

Combining assumption (iii) and (8) derives

$$f(x_k) - f^* \le \delta\|\nabla f(x_k)\| \le \delta\gamma^{|\mathcal{I}_{k-1}|}\theta^{k-|\mathcal{I}_{k-1}|}\|\nabla f(x_0)\|. \tag{9}$$

Note that $\gamma \ge 1$, $\theta \in (0,1)$, and $(\gamma\theta)^{\frac{k}{2}} \le \nu k^{-2}$ hold from assumption (iv) and $k \ge \ell$. It then follows from $|\mathcal{I}_{k-1}| \le |\mathcal{I}_k| < \frac{k}{2}$ that

$$\gamma^{|\mathcal{I}_{k-1}|}\theta^{k-|\mathcal{I}_{k-1}|} \le (\gamma\theta)^{\frac{k}{2}} \le \frac{\nu}{k^2}. \tag{10}$$

We have from (9) and (10) that

$$f(x_k) - f^* \leq \frac{\nu\delta\|\nabla f(x_0)\|}{k^2}.$$

Cases (1) and (2) guarantee that the desired inequality holds when $|\mathcal{I}(\theta)| = \infty$. Therefore, the assertion is proven. □

Remark 1. We discuss sufficient conditions for assumptions (i)–(iv) of Proposition 1. Assumptions (i) and (ii) would be satisfied for any sequence generated by descent methods. Note that γ of item (ii) is allowed to be greater than or equal to 1. Assumption (iii) is satisfied when $\{x_k\}$ is bounded. Assumption (iv) holds if $\theta \in (0, \gamma^{-1})$ because it implies $\theta^k = \mathcal{O}(k^{-2})$ and $(\gamma\theta)^{\frac{k}{2}} = \mathcal{O}(k^{-2})$ as $k \to \infty$. From these discussions, we can see that assumption (v) is the key to global $\mathcal{O}(k^{-2})$ convergence.

3. An RNM with generalized regularization terms

In this paper, we consider an RNM with generalized regularization terms that iteratively solves the following subproblem:

$$\underset{d\in\mathbb{R}^n}{\text{minimize}} \quad \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle + \frac{\mu_k}{p}\|d\|^p + \rho_k\|d\|_1, \tag{11}$$

where $\mu_k > 0$ and $\rho_k \geq 0$ are parameters, and $p \in (1, 3]$ is a pre-fixed constant. The proposed method obtains a search direction d_k by solving subproblem (11), and the point x_k is updated as $x_{k+1} := x_k + d_k$.

Now, we denote by φ_k the objective function of subproblem (11), namely,

$$\varphi_k(d) := \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle + \frac{\mu_k}{p}\|d\|^p + \rho_k\|d\|_1.$$

Since the objective function φ_k has generalized regularization terms $\frac{\mu_k}{p}\|d\|^p$ and $\rho_k\|d\|_1$, we call the proposed method a generalized RNM (GRNM).

Remark 2. The GRNM includes the quadratic, cubic, and elastic net regularization as special cases. Moreover, it includes a novel regularization in addition to the aforementioned regularization.

Remark 3. The proposed GRNM can adopt the L^1 regularization term, that is, $\mu_k = 0$ and $\rho_k > 0$. However, subproblem (11) with $\mu_k = 0$ might have no global optimum when $\nabla^2 f(x_k)$ is not positive definite. Conversely, if ρ_k is sufficiently large, the solution becomes 0. We provide sufficient conditions under which (11) has nonzero solutions.

Lemma 1. *Let $x_k \in \mathbb{R}^n$, $\mu_k > 0$, and $\rho_k \geq 0$ be given. If $\rho_k < \|\nabla f(x_k)\|_\infty$, then problem (11) has a unique global optimum $d_k \neq 0$ that satisfies*

$$\nabla f(x_k) + (\nabla^2 f(x_k) + \mu_k\|d_k\|^{p-2}I)d_k + \rho_k\eta_k = 0$$

for some $\eta_k \in \partial\|d_k\|_1$. Moreover, d_k is the descent direction of f at x_k, that is, $\langle \nabla f(x_k), d_k\rangle < 0$.

Proof. First, we show the solvability of (11). Recall that φ_k is closed, proper, and coercive. Hence, by using [20, Proposition 3.2.1], problem (11) has a global optimum d_k. The uniqueness of d_k is derived from the strict convexity of φ_k.

Hereafter, we show that $d_k \neq 0$ is satisfied. We assume to the contrary that $d_k = 0$ holds. As $d_k = 0$ satisfies the first-order optimality condition of (11), there exists $\eta_k \in \partial\|d_k\|_1$ such that $\nabla f(x_k) + \rho_k\eta_k = 0$. It then follows from $\rho_k < \|\nabla f(x_k)\|_\infty$ that $\|\nabla f(x_k)\|_\infty = \rho_k < \|\nabla f(x_k)\|_\infty$. This result contradicts, that is, $d_k \neq 0$.

Finally, the first-order optimality condition of (11) leads to the desired equality, and it yields

$$\langle \nabla f(x_k), d_k\rangle = -\langle (\nabla^2 f(x_k) + \mu_k\|d_k\|^{p-2}I)d_k, d_k\rangle - \rho_k\|d_k\|_1 < 0,$$

where note that $\langle d_k, \eta_k\rangle = \|d_k\|_1$ and $d_k \neq 0$. This completes the proof. □

Remark 4. By utilizing the line search strategy or an appropriate choice of μ_k and ρ_k, we can prove the global convergence of Algorithm 1. However, because $\mathcal{O}(k^{-2})$ convergence implies global convergence, we omit discussions on the line search.

We provide a formal description of the proposed method in Algorithm 1.

Algorithm 1 (GRNM).

1: Choose $p \in (1, 3]$, $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$, and set $k := 0$.
2: If $\|\nabla f(x_k)\| \leq \varepsilon$, then stop.
3: Set parameters $\mu_k > 0$ and $\rho_k \geq 0$, and find a global optimum d_k of (11).
4: Set $x_{k+1} := x_k + d_k$.
5: Set $k \leftarrow k + 1$, and go to Line 2.

Remark 5. We briefly discuss how to solve subproblem (11) for specific choices of regularization parameters. For the setting $p = 2$, $\mu_k > 0$, and $\rho_k = 0$, the solution d_k can be obtained by solving $(\nabla^2 f(x_k) + \mu_k I)d_k = -\nabla f(x_k)$. Regarding the setting $p = 3$, $\mu_k > 0$, and $\rho_k = 0$, the subproblem is nonconvex minimization with respect to a cubic function. However, it is known that it can be reduced to a convex one-dimensional optimization problem [6]. Moreover, if $p = 2$, $\mu_k > 0$, and $\rho_k > 0$, then (11) is formulated as the L^1–L^2 optimization problem and this regularization is called the elastic net. In particular, it is possible that problem (1) can be quickly solved when the search direction d_k is sparse for each $k \in \mathbb{N}$. For the other cases, such as $p \in (1,3)\backslash\{2\}$ and $\rho_k > 0$, it is generally difficult to solve the subproblem efficiently. However, it may be possible to construct efficient methods by exploiting specific structures of f if we restrict (1) to some special cases.

4. Global $\mathcal{O}(k^{-2})$ convergence of Algorithm 1

This section shows that Algorithm 1 globally converges with the $\mathcal{O}(k^{-2})$ rate. From now on, we denote by $x^* \in \mathbb{R}^n$ an optimal solution of problem (1), and use the following notation: $f^* := f(x^*)$ and $\mathcal{S} := \{x \in \mathbb{R}^n; f(x) \le f(x_0)\}$.

We will show global $\mathcal{O}(k^{-2})$ convergence of Algorithm 1 by showing that assumptions (i)–(v) in Proposition 1 hold for a sequence $\{x_k\}$ generated by Algorithm 1.

In the subsequent argument, we suppose that $\varepsilon = 0$ and Algorithm 1 generates an infinite sequence $\{x_k\}$ satisfying $\nabla f(x_k) \ne 0$ for each $k \in \mathbb{N} \cup \{0\}$. Moreover, we make the following assumptions.

(A1) There exists $L > 0$ such that for any $x, y \in \mathbb{R}^n$,

$$\|\nabla f(x) - \nabla f(y) - \nabla^2 f(y)(x-y)\| \le L\|x-y\|^2,$$
$$|f(x) - f(y) - \langle \nabla f(y), x-y\rangle - \frac{1}{2}\langle \nabla^2 f(y)(x-y), x-y\rangle| \le \frac{L}{3}\|x-y\|^3.$$

(A2) The parameters μ_k and ρ_k are set as follows: For all $k \in \mathbb{N} \cup \{0\}$,

$$\mu_k := c_1^{\frac{p-1}{2}}\|\nabla f(x_k)\|^{\frac{3-p}{2}},\ \rho_k := \min\left\{\frac{q}{\sqrt{n}}\|\nabla f(x_k)\|, c_2\|\nabla f(x_k)\|^{\frac{p+1}{2}}\right\},$$

where $c_1 \ge L$, $c_2 \in (0,1)$, and $q \ge 0$.

(A3) There exists $R > 0$ such that $\mathcal{S} \subset B(0,R)$.

Note that subproblem (11) has a global optimum $d_k \ne 0$ because (A2) satisfies the condition of Lemma 1. Note also that several basic properties of linear algebra derive

$$\left\|(\nabla^2 f(x_k) + \mu_k\|d_k\|^{p-2}I)^{-1}\right\| \le \mu_k^{-1}\|d_k\|^{2-p}, \tag{12}$$
$$\left\|(\nabla^2 f(x_k) + \mu_k\|d_k\|^{p-2}I)^{-1}\nabla^2 f(x_k)\right\| \le 1. \tag{13}$$

We now show that assumptions (i) and (ii) in Proposition 1 hold.

Lemma 2. *Suppose that* (A1) *and* (A2) *are satisfied. Suppose also that* $3 - (1+q)^{\frac{3-p}{p-1}} > 0$ *where q is a constant in* (A2). *For any $k \in \mathbb{N} \cup \{0\}$, the following inequalities hold:*

(a) $f(x_{k+1}) - f(x_k) \le -\dfrac{3 - (1+q)^{\frac{3-p}{p-1}}}{3}\mu_k\|d_k\|^p < 0,$

(b) $\|\nabla f(x_{k+1})\| \le \left(1 + (1+q)^{\frac{3-p}{p-1}}\right)\mu_k\|d_k\|^{p-1} + \min\left\{q\|\nabla f(x_k)\|, \sqrt{n}c_2\|\nabla f(x_k)\|^{\frac{p+1}{2}}\right\},$

(c) $\|\nabla f(x_{k+1})\| \le \left(1 + 2q + (1+q)^{\frac{2}{p-1}}\right)\|\nabla f(x_k)\|.$

Proof. It follows from Lemma 1 and (12) that

$$\|d_k\| \le \left\|(\nabla^2 f(x_k) + \mu_k\|d_k\|^{p-2}I)^{-1}(\nabla f(x_k) + \rho_k\eta_k)\right\| \le \mu_k^{-1}\|d_k\|^{2-p}(\|\nabla f(x_k)\| + \sqrt{n}\rho_k) \le \mu_k^{-1}\|d_k\|^{2-p}(1+q)\|\nabla f(x_k)\|,$$

that is,

$$\|d_k\| \le \frac{(1+q)^{\frac{1}{p-1}}}{\sqrt{c_1}}\sqrt{\|\nabla f(x_k)\|}. \tag{14}$$

By (14) and the first equality of (A2), we have

$$L\|d_k\|^2 \le c_1\|d_k\|^{3-p}\cdot\|d_k\|^{p-1} \le (1+q)^{\frac{3-p}{p-1}}\cdot c_1^{\frac{p-1}{2}}\|\nabla f(x_k)\|^{\frac{3-p}{2}}\cdot\|d_k\|^{p-1} = (1+q)^{\frac{3-p}{p-1}}\mu_k\|d_k\|^{p-1}. \tag{15}$$

Recall that $x_{k+1} = x_k + d_k$ and $\langle \eta_k, d_k \rangle = \|d_k\|_1$. Combining the second inequality of (A1), Lemma 1, and (15) yields

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\le \langle \nabla f(x_k) + \nabla^2 f(x_k) d_k, d_k \rangle - \frac{1}{2} \langle \nabla^2 f(x_k) d_k, d_k \rangle + \frac{L}{3} \|d_k\|^3 \\ &\le -\mu_k \|d_k\|^p - \rho_k \|d_k\|_1 + \frac{1}{3} \|d_k\| \cdot L \|d_k\|^2 \\ &\le -\frac{3 - (1+q)^{\frac{3-p}{p-1}}}{3} \mu_k \|d_k\|^p. \end{aligned}$$

From $x_{k+1} = x_k + d_k$ and Lemma 1, we obtain

$$\nabla f(x_{k+1}) = \nabla f(x_k + d_k) - \nabla f(x_k) - (\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I) d_k - \rho_k \eta_k.$$

Subsequently, exploiting the first inequality of (A1), the second equality of (A2), and (15) derives

$$\begin{aligned} \|\nabla f(x_{k+1})\| &\le \|\nabla f(x_k + d_k) - \nabla f(x_k) - \nabla^2 f(x_k) d_k\| + \mu_k \|d_k\|^{p-1} + \sqrt{n} \rho_k \\ &\le L \|d_k\|^2 + \mu_k \|d_k\|^{p-1} + \min\left\{ q \|\nabla f(x_k)\|, \sqrt{n} c_2 \|\nabla f(x_k)\|^{\frac{p+1}{2}} \right\} \\ &\le \left(1 + (1+q)^{\frac{3-p}{p-1}} \right) \mu_k \|d_k\|^{p-1} + \min\left\{ q \|\nabla f(x_k)\|, \sqrt{n} c_2 \|\nabla f(x_k)\|^{\frac{p+1}{2}} \right\}, \end{aligned} \tag{16}$$

namely, item (b) is verified. Meanwhile, from (14) and the first equality of (A2), we have $\mu_k \|d_k\|^{p-1} \le (1+q)\|\nabla f(x_k)\|$. Utilizing this result, (16), and $\min\{q\|\nabla f(x_k)\|, \sqrt{n} c_2 \|\nabla f(x_k)\|^{\frac{p+1}{2}}\} \le q\|\nabla f(x_k)\|$ means

$$\|\nabla f(x_{k+1})\| \le \left(1 + 2q + (1+q)^{\frac{2}{p-1}} \right) \|\nabla f(x_k)\|.$$

Therefore, the desired inequalities are obtained. □

Now, we provide the global $\mathcal{O}(k^{-2})$ convergence property of Algorithm 1 by showing assumptions (iii)-(v) in Proposition 1.

Theorem 1. *Suppose that* (A1)–(A3) *hold. Moreover, suppose that the following assumptions* (A4)–(A6) *hold:*

(A4) $3 - (1+q)^{\frac{3-p}{p-1}} > 0$;
(A5) $1 + 2q + (1+q)^{\frac{2}{p-1}} \ge 1$;
(A6) *there exist* $\theta \in (q, 1)$, $\nu > 0$, *and* $\ell \in \mathbb{N}$ *such that*

$$\theta^k \le \nu k^{-2}, \quad \left(\left(1 + 2q + (1+q)^{\frac{2}{p-1}} \right) \theta \right)^{\frac{k}{2}} \le \nu k^{-2} \quad \forall k \ge \ell.$$

Let $\mathcal{I}(\theta) := \{ i \in \mathbb{N} \cup \{0\}; \theta \|\nabla f(x_i)\| \le \|\nabla f(x_{i+1})\| \}$. *Let* D *and* τ *be defined as*

$$D := R + \|x^*\|, \quad \tau := \frac{(\theta - q)^{\frac{p}{p-1}} \left(3 - (1+q)^{\frac{3-p}{p-1}} \right)}{3\sqrt{c_1} D^{\frac{3}{2}} \left(1 + (1+q)^{\frac{3-p}{p-1}} \right)^{\frac{p}{p-1}}}.$$

Then, a sequence $\{x_k\}$ *generated by Algorithm 1 satisfies one of the following statements:*

(a) *If* $|\mathcal{I}(\theta)| < \infty$, *then*

$$f(x_k) - f^* \le \frac{\theta^{-(\hat{i}+1)} \nu D \|\nabla f(x_{\hat{i}+1})\|}{k^2} \quad \forall k \ge \max\{\ell, \hat{i} + 2\},$$

where $\hat{i}$ *is the largest element of* $\mathcal{I}(\theta)$.

(b) *If* $|\mathcal{I}(\theta)| = \infty$, *then*

$$f(x_k) - f^* \le \max\left\{ \frac{36\tau^{-2}}{(k+4)^2}, \frac{\nu D \|\nabla f(x_0)\|}{k^2} \right\} \quad \forall k \ge \ell.$$

Proof. If items (i)–(v) in Proposition 1 hold, then the desired result can be obtained. Item (i) directly follows from (a) of Lemma 2. Let us define $r := 1 + 2q + (1+q)^{\frac{2}{p-1}}$. Recall that (A5) ensures $r \ge 1$. Thus, item (ii) holds from (c) of Lemma 2. The definition of r and (A6) imply that item (iv) is satisfied. Thus, it is sufficient to show items (iii) and (v).

Since $\{x_k\} \subset S$ holds, we have from (A3) that $\|x_k\| \le R$ for $k \in \mathbb{N} \cup \{0\}$. Let us take $k \in \mathbb{N} \cup \{0\}$ arbitrarily. Then, it is clear that $\|x_k - x^*\| \le D$, and hence the convexity of f yields $f(x_k) - f^* \le \langle \nabla f(x_k), x_k - x^* \rangle \le D\|\nabla f(x_k)\|$. This fact implies that item (iii) holds, and

$$\left(\frac{f(x_k) - f^*}{D}\right)^{\frac{3}{2}} \le \|\nabla f(x_k)\|^{\frac{3}{2}}. \tag{17}$$

Now, we take arbitrary $k \in \mathcal{I}(\theta)$. The definition of μ_k in (A2) and (b) of Lemma 2 derive

$$\theta\|\nabla f(x_k)\| \le \left(1 + (1+q)^{\frac{3-p}{p-1}}\right) c_1^{\frac{p-1}{2}} \|\nabla f(x_k)\|^{\frac{3-p}{2}} \|d_k\|^{p-1} + q\|\nabla f(x_k)\|,$$

which implies

$$\frac{(\theta - q)^{\frac{p}{p-1}}}{c_1^{\frac{p}{2}}\left(1 + (1+q)^{\frac{3-p}{p-1}}\right)^{\frac{p}{p-1}}} \|\nabla f(x_k)\|^{\frac{p}{2}} \le \|d_k\|^p.$$

Multiplying both sides of this inequality by $\mu_k = c_1^{\frac{p-1}{2}} \|\nabla f(x_k)\|^{\frac{3-p}{2}}$ yields

$$\frac{(\theta - q)^{\frac{p}{p-1}}}{\sqrt{c_1}\left(1 + (1+q)^{\frac{3-p}{p-1}}\right)^{\frac{p}{p-1}}} \|\nabla f(x_k)\|^{\frac{3}{2}} \le \mu_k\|d_k\|^p. \tag{18}$$

Using item (a) of Lemma 2 and (18), we obtain

$$f(x_{k+1}) - f(x_k) \le -\frac{(\theta - q)^{\frac{p}{p-1}}\left(3 - (1+q)^{\frac{3-p}{p-1}}\right)}{3\sqrt{c_1}\left(1 + (1+q)^{\frac{3-p}{p-1}}\right)^{\frac{p}{p-1}}} \|\nabla f(x_k)\|^{\frac{3}{2}}. \tag{19}$$

Moreover, from (17) and (19),

$$f(x_{k+1}) - f(x_k) \le -\tau(f(x_k) - f^*)^{\frac{3}{2}} \qquad \forall k \in \mathcal{I}(\theta).$$

Therefore, we can verify that item (v) of Proposition 1 holds. $\square$

From Theorem 1, we have to indicate the existence of q and θ satisfying (A4)–(A6) to show global $\mathcal{O}(k^{-2})$ convergence of Algorithm 1. Although the existence of these parameters cannot be ensured for all $p > 1$, we can show their existence for specific $p \in (1, 3]$. Two examples of these concrete parameters are presented.

Example 1.

$$p \in (1, 3], \quad q := 0, \quad \theta := \frac{3}{8}.$$

Example 2.

$$p \in (1, 3], \quad q := \min\left\{\frac{1}{10}(2^{\frac{p-1}{3-p}} - 1), \frac{1}{20}2^{\frac{3-p}{p-1}}\right\}, \quad \theta := \frac{1}{5}.$$

The parameters p, q, and θ described in Examples 1 and 2 satisfy conditions (A4)–(A6). Note that these examples are valid for any $p \in (1, 3]$. For details, see Appendix A.

Remark 6. When $(p, q, \theta) = (2, 0, 4^{-1})$, we can easily verify that $\tau = (96D^{3/2}\sqrt{c_1})^{-1}$. This value coincides with that of Mishchenko [5], implying that the proposed method is a generalization of [5].

Remark 7. Parameter θ described in Theorem 1 is only required for the proof and is unrelated to problem (1) and Algorithm 1. Hence, it should be selected to provide a good coefficient regarding the convergence rate. Since the coefficients are determined by

$$\frac{1}{\theta^{(\hat{i}+1)}}, \quad \left[\frac{3\sqrt{c_1}D^{\frac{3}{2}}\left(1 + (1+q)^{\frac{3-p}{p-1}}\right)^{\frac{p}{p-1}}}{(\theta - q)^{\frac{p}{p-1}}\left(3 - (1+q)^{\frac{3-p}{p-1}}\right)}\right]^2,$$

we should take θ as large as possible.

5. Local superlinear convergence of Algorithm 1

This section aims to show local and superlinear convergence of Algorithm 1. Throughout this section, the set of optimal solutions is denoted by $X^* \subset \mathbb{R}^n$. We first make an additional assumption.

(A7) There exist $r_1 > 0$ and $m_1 > 0$ such that $\mathrm{dist}(x, X^*) \le m_1 \|\nabla f(x)\|$ for each $x \in B(x^*, r_1)$.

For a given point $x \in \mathbb{R}^n$, let $\hat{x}$ be a point satisfying

$$\hat{x} \in X^*, \quad \|\hat{x} - x\| = \mathrm{dist}(x, X^*).$$

Some important inequalities for local convergence are as follows:

Lemma 3. *Suppose that* (A1), (A2), and (A7) *hold. Then, there exist* $r_2 > 0$*,* $m_2 \ge m_3 > 0$*, and* $m_4 > 0$ *such that*

(a) $\|d_k\| \le m_2 \mathrm{dist}(x_k, X^*)$ *and* $\|d_k\| \ge m_3 \mathrm{dist}(x_k, X^*)$ *for* $x_k \in B(x^*, r_2)$*;*
(b) $\mathrm{dist}(x_{k+1}, X^*) \le m_4 \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}$ *for* $x_k, x_{k+1} \in B(x^*, r_2)$*.*

Proof. To begin with, we define u_1 and u_2 as

$$u_1 := \sup\left\{\|\nabla^2 f(z)\|; z \in B(x^*, r_1)\right\}, \quad u_2 := u_1 + \frac{c_1 r_1}{2},$$

respectively, and will show the following inequality:

$$\|\nabla f(x)\| \le u_2 \mathrm{dist}(x, X^*) \quad \forall x \in B(x^*, r_1). \tag{20}$$

We have $\|\hat{x} - x\| \le \|x - x^*\| \le r_1$. Hence, the first inequality of (A1) guarantees $\|\nabla f(\hat{x}) - \nabla f(x) - \nabla^2 f(x)(\hat{x} - x)\| \le \frac{c_1}{2}\mathrm{dist}(x, X^*)^2$. It then follows from $\nabla f(\hat{x}) = 0$ and $\mathrm{dist}(x, X^*) \le \|x - x^*\| \le r_1$ that $\|\nabla f(x)\| - u_1 \mathrm{dist}(x, X^*) \le \frac{c_1 r_1}{2}\mathrm{dist}(x, X^*)$. Thus, inequality (20) holds.

We show item (a). Define r_2 as follows:

$$r_2 := \min\left\{ r_1, \left(2\sqrt{n} c_2 u_2^{\frac{p-1}{2}}\right)^{-\frac{2}{p-1}} \right\}.$$

Let $x_k \in B(x^*, r_2)$. The definition of $\mathrm{dist}(x_k, X^*)$ and inequality (20) ensure

$$\mathrm{dist}(x_k, X^*) \le \|x_k - x^*\| \le r_2, \tag{21}$$

$$\|\nabla f(x_k)\| \le u_2 \mathrm{dist}(x_k, X^*) \le u_2 \|x_k - x^*\| \le u_2 r_2. \tag{22}$$

Lemma 1 and (12) lead to $\|d_k\| \le \|(\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I)^{-1} \nabla f(x_k)\| + \frac{\sqrt{n}\rho_k}{\mu_k \|d_k\|^{p-2}}$. It then follows from (A2) and (20) that

$$\|d_k\| \le \left\|(\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I)^{-1} \nabla f(x_k)\right\| + \frac{\sqrt{n} c_2 u_2^{\frac{p+1}{2}}}{\mu_k \|d_k\|^{p-2}} \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}. \tag{23}$$

Now, we notice that $\nabla f(x_k) = -(\nabla f(\hat{x}_k) - \nabla f(x_k) - \nabla^2 f(x_k)(\hat{x}_k - x_k)) - \nabla^2 f(x_k)(\hat{x}_k - x_k)$ holds from $\nabla f(\hat{x}_k) = 0$. Then, combining (A1), (12), and (13) implies

$$\begin{aligned}
&\left\|(\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I)^{-1} \nabla f(x_k)\right\| \\
&\le \left\|(\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I)^{-1}\right\| \|\nabla f(\hat{x}_k) - \nabla f(x_k) - \nabla^2 f(x_k)(\hat{x}_k - x_k)\| + \left\|(\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I)^{-1} \nabla^2 f(x_k)\right\| \|\hat{x}_k - x_k\| \\
&\le \frac{c_1}{2\mu_k \|d_k\|^{p-2}} \mathrm{dist}(x_k, X^*)^2 + \mathrm{dist}(x_k, X^*).
\end{aligned} \tag{24}$$

By (21), (23) and (24), we get $\|d_k\| \le \frac{u_3}{\mu_k \|d_k\|^{p-2}} \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}} + \mathrm{dist}(x_k, X^*)$, where $u_3 := \frac{1}{2} c_1 r_2^{\frac{3-p}{2}} + \sqrt{n} c_2 u_2^{\frac{p+1}{2}}$. This inequality can be reformulated as $\mu_k \|d_k\|^{p-1} \le 2 \max\{u_3 \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}, \mu_k \|d_k\|^{p-2} \mathrm{dist}(x_k, X^*)\}$. There are two possible cases: (i) $\mu_k \|d_k\|^{p-1} \le 2u_3 \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}$; (ii) $\mu_k \|d_k\|^{p-1} \le 2\mu_k \|d_k\|^{p-2} \mathrm{dist}(x_k, X^*)$. In case (i), utilizing (A2) and (A7) derives

$$\frac{c_1^{\frac{p-1}{2}} \mathrm{dist}(x_k, X^*)^{\frac{3-p}{2}}}{m_1^{\frac{3-p}{2}}} \|d_k\|^{p-1} \le \mu_k \|d_k\|^{p-1} \le 2u_3 \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}.$$

Hence, we have $\|d_k\| \le c_1^{-\frac{1}{2}}(2u_3 m_1^{\frac{3-p}{2}})^{\frac{1}{p-1}}\mathrm{dist}(x_k, X^*)$. Meanwhile, case (ii) leads to $\|d_k\| \le 2\mathrm{dist}(x_k, X^*)$. Thus, there exists $m_2 > 0$ such that $\|d_k\| \le m_2 \mathrm{dist}(x_k, X^*)$.

Now, Lemma 1, (A2), and (22) yield

$$\|d_k\| \ge \frac{\|\nabla f(x_k) + \rho_k \eta_k\|}{\|\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I\|} \ge \frac{1 - \sqrt{n} c_2 (u_2 r_2)^{\frac{p-1}{2}}}{u_1 + \mu_k \|d_k\|^{p-2}} \|\nabla f(x_k)\| \ge \frac{1}{2(u_1 + \mu_k \|d_k\|^{p-2})} \|\nabla f(x_k)\|, \tag{25}$$

where the last inequality follows from the definition of r_2. Exploiting (A2) and $\|d_k\| \le m_2 \mathrm{dist}(x_k, X^*)$ derives

$$\mu_k \|d_k\|^{p-2} \le c_1^{\frac{p-1}{2}} \|\nabla f(x_k)\|^{\frac{3-p}{2}} m_2^{p-2} \mathrm{dist}(x_k, X^*)^{p-2}.$$

It then follows from (21) and (22) that

$$\mu_k \|d_k\|^{p-2} \le c_1^{\frac{p-1}{2}} u_2^{\frac{3-p}{2}} m_2^{p-2} r_2^{\frac{p-1}{2}}. \tag{26}$$

By (25), (26), and (A7), there exists $m_3 > 0$ satisfying $\|d_k\| \ge m_3 \mathrm{dist}(x_k, X^*)$.

Next, we prove item (b). Let $x_{k+1} = x_k + d_k \in B(x^*, r_2)$. Lemma 1 implies

$$\begin{aligned}\nabla f(x_{k+1}) &= (\nabla f(x_k + d_k) - \nabla f(x_k) - \nabla^2 f(x_k) d_k) - (\mu_k \|d_k\|^{p-2} d_k + \rho_k \eta_k) + (\nabla f(x_k) + (\nabla^2 f(x_k) + \mu_k \|d_k\|^{p-2} I) d_k + \rho_k \eta_k) \\ &= (\nabla f(x_k + d_k) - \nabla f(x_k) - \nabla^2 f(x_k) d_k) - (\mu_k \|d_k\|^{p-2} d_k + \rho_k \eta_k).\end{aligned}$$

Then, we have from (A1), (A2), and (A7) that

$$\begin{aligned}\mathrm{dist}(x_{k+1}, X^*) &\le m_1 \|\nabla f(x_k + d_k)\| \\ &\le m_1 \|\nabla f(x_k + d_k) - \nabla f(x_k) - \nabla^2 f(x_k) d_k\| + m_1 \mu_k \|d_k\|^{p-1} + \sqrt{n} m_1 \rho_k \\ &\le \frac{c_1 m_1}{2} \|d_k\|^2 + c_1^{\frac{p+1}{2}} m_1 \|\nabla f(x_k)\|^{\frac{3-p}{2}} \|d_k\|^{p-1} + \sqrt{n} c_2 m_1 \|\nabla f(x_k)\|^{\frac{p+1}{2}}.\end{aligned} \tag{27}$$

Now, recall that $\|d_k\| \le m_2 \mathrm{dist}(x_k, X^*)$, $\|\nabla f(x_k)\| \le u_2 \mathrm{dist}(x_k, X^*)$, and $\mathrm{dist}(x_k, X^*) \le r_2$, where the second and third inequalities follow from (20) and (21), respectively. Thus, we can easily verify that

$$\begin{aligned}&\frac{c_1 m_1}{2} \|d_k\|^2 + c_1^{\frac{p+1}{2}} m_1 \|\nabla f(x_k)\|^{\frac{3-p}{2}} \|d_k\|^{p-1} + \sqrt{n} c_2 m_1 \|\nabla f(x_k)\|^{\frac{p+1}{2}} \\ &\le \left(\frac{c_1 m_1 m_2^2 r_2^{\frac{3-p}{2}}}{2} + c_1^{\frac{p+1}{2}} u_2^{\frac{3-p}{2}} m_2^{p-1} + \sqrt{n} c_2 m_1 u_2^{\frac{p+1}{2}} \right) \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}.\end{aligned} \tag{28}$$

Therefore, combining (27) and (28) guarantees the existence of $m_4 > 0$ satisfying $\mathrm{dist}(x_{k+1}, X^*) \le m_4 \mathrm{dist}(x_k, X^*)^{\frac{p+1}{2}}$. □

Finally, we establish local and superlinear convergence of Algorithm 1. Although we can prove the theorem using the above lemmas in a manner similar to [2, Theorem 3.2], the proof is given in Appendix B for completeness of the paper.

Theorem 2. *Suppose that* (A1), (A2), and (A7) *hold. If an initial point x_0 is chosen sufficiently close to x^*, then any sequence $\{x_k\}$ generated by Algorithm 1 converges to some global optimum $\bar{x} \in X^*$ superlinearly. Moreover, if $p = 3$ is satisfied, then $\{x_k\}$ converges to $\bar{x} \in X^*$ quadratically.*

6. Numerical experiments

This section provides numerical experiments to confirm the performance of Algorithm 1. In particular, we compare results among the following four regularization terms in problem (11): the quadratic regularization $\frac{\mu_k}{2}\|d\|^2$, elastic-net regularization $\frac{\mu_k}{2}\|d\|^2 + \rho_k \|d\|_1$, cubic regularization $\frac{\mu_k}{3}\|d\|^3$, and cubic–L^1 regularization $\frac{\mu_k}{3}\|d\|^3 + \rho_k \|d\|_1$. The termination criterion ε of Algorithm 1 was set as $\varepsilon := 10^{-6}$. The parameters μ_k and ρ_k were updated as follows:

$$\mu_k := c_k^{\frac{p-1}{2}} \|\nabla f(x_k)\|^{\frac{3-p}{2}}, \quad \rho_k := \min\left\{ \frac{q}{\sqrt{n}} \|\nabla f(x_k)\|, c \|\nabla f(x_k)\|^{\frac{p+1}{2}} \right\},$$

where $c := 0.5$, $q := 0.01$, $c_0 := 100$, and

$$c_k := \max\left\{ m_k, \frac{c_{k-1}}{2} \right\}, \quad m_k := \frac{\|\nabla f(x_k) - \nabla f(x_{k-1}) - \nabla^2 f(x_{k-1})(x_k - x_{k-1})\|}{\|x_k - x_{k-1}\|^2}$$

Table 1
Performance of Algorithm 1 on problem (29).

	quadratic	elastic net	cubic	cubic–L^1
Max of ♯ite	24	24	22	23
Min of ♯ite	23	23	21	21
Averaged ♯ite	23.2	23.6	21.8	21.8
Max of $\|\nabla f(x^*)\|$	9.46e-07	2.88e-07	9.27e-07	8.38e-07

Table 2
Performance of Algorithm 1 on problem (30).

	quadratic	elastic net	cubic	cubic–L^1
Max of ♯ite	19	19	20	21
Min of ♯ite	18	18	16	18
Averaged ♯ite	18.4	18.4	18.6	19.6
Max of $\|\nabla f(x^*)\|$	6.77e-07	8.93e-07	9.81e-07	7.94e-07

for $k \in \mathbb{N}$. Note that the updating rule of c_k is inspired by [5, Section 3]. The initial point $x_0 \in \mathbb{R}^n$ was set as $x_0 := 0$. All the programs were implemented with MATLAB R2023a and ran on a machine with an Intel Core–i9–9900K 3.60GHz and 128GB RAM. Moreover, we utilized `fminunc`, which is a MATLAB optimizer for unconstrained optimization problems, to solve problem (11) at each iteration except for the quadratic and elastic net regularization. Regarding the quadratic regularization, we solved $(\nabla^2 f(x_k) + \mu_k I)d = -\nabla f(x_k)$ instead of problem (11). For the elastic net regularization, we utilized `quadprog`, which is a MATLAB optimizer for quadratic programming problems, to solve the following quadratic programming problem obtained by reformulating the corresponding subproblem:

$$\begin{aligned} \underset{(d,t)\in\mathbb{R}^n\times\mathbb{R}^n}{\text{minimize}} \quad & \langle \nabla f(x_k), d\rangle + \frac{1}{2}\langle \nabla^2 f(x_k)d, d\rangle + \frac{\mu_k}{2}\|d\|^2 + \rho_k \sum_{j=1}^{n} t_j, \\ \text{subject to} \quad & -t_j \le d_j \le t_j \quad \forall j \in \{1,2,\dots,n\}. \end{aligned}$$

The following two types of test problems were solved. The first one is the following convex optimization problem [2]:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2}\sum_{i=1}^{n-1}(x_i - x_{i+1})^2 + \frac{1}{12}\sum_{i=1}^{n-1}\alpha_i (x_i - x_{i+1})^4, \tag{29}$$

where $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ are positive constants. In the experiments, we set $n = 200$ and $\alpha_i = 1$ for each $i \in \mathbb{N}$. The second problem is the log-sum-exp problem [5]:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad \kappa \log\left[\sum_{i=1}^{m} \exp\left(\frac{a_i^\top x - b_i}{\kappa}\right)\right], \tag{30}$$

where $a_1, a_2, \dots, a_m \in \mathbb{R}^n$ are constant vectors and $\kappa, b_1, b_2, \dots, b_m \in \mathbb{R}$ are constants. For the test problem, we set $n = 200$, $m = 500$, and $\kappa = 0.5$. Moreover, a_i $(i = 1,2,\dots,m)$ and b_i $(i = 1,2,\dots,m)$ were generated randomly from $[-1,1]^n$ and $[-1,1]$, respectively. For each test problem, five instances were randomly made and solved by Algorithm 1.

Tables 1 and 2 provide the averaged, maximum, and minimum number of iterations and the maximum value of $\|\nabla f(x^*)\|$ obtained by solving the five instances for problems (29) and (30), respectively. Note also that x^* described in the tables indicates the final iteration point calculated by Algorithm 1. We can confirm that all the instances were solved because there are no maximum values of $\|\nabla f(x^*)\|$ being greater than 10^{-6}. No significant differences were observed among the four regularization terms concerning the number of iterations as seen in Tables 1 and 2. Meanwhile, the computational time of the elastic net, cubic, and cubic–L^1 regularization was higher than the quadratic regularization because we utilized generic optimizers to solve their subproblems except for the quadratic regularization. Since there are no significant differences in the number of iterations, solving subproblems and evaluating the objective function affect the computational time. Thus, it is important to develop efficient methods for solving the subproblems and to compute search directions that can reduce the cost of evaluating the objective function f, e.g., search directions with sparsity.

7. Concluding remarks

In this paper, we have proposed Algorithm 1, which is an RNM with generalized regularization terms. The proposed method is based on the RNM proposed by Mishchenko [5], but it is a generalization of the existing one regarding regularization. Therefore, not only the quadratic and cubic RNMs but also novel RNMs with other regularization, such as the elastic net, are included in Algorithm 1. We have proven global $\mathcal{O}(k^{-2})$ and local superlinear convergence of Algorithm 1. Moreover, we have examined the performance of Algorithm 1 by using several regularization terms.

One of future research is to propose an accelerated GRNM that globally converges in the order of $\mathcal{O}(k^{-3})$. For the existing methods, such as the RNM [5] and cubic RNM [6], their accelerated methods have been proposed in [1,7]. In these studies, accelerated schemes have been proposed and they have been incorporated in the proposed accelerated methods. Thus, we think that such an accelerated scheme can be applied to Algorithm 1 and its accelerated method can also be proposed. As another challenge, it would be worthwhile to develop a specialized optimizer to solve subproblems of a specific regularization other than the quadratic one. Specifically, we can consider regularization that utilizes the sparsity of search directions by including the L^1 regularization term in subproblem (11) because Ariizumi et al. [19] proposed a Levenberg-Marquardt method equipped with such a regularization and succeeded in quickly finding solutions for nonlinear equations.

Acknowledgements

The authors received support from the Japan Society for the Promotion of Science KAKENHI for 21K17709 and 21K11769. This work was partially supported by the joint project of Kyoto University and Toyota Motor Corporation, titled “Advanced Mathematical Science for Mobility Society”. We would like to thank the referees for their valuable comments which gave us a lot of constructive suggestions.

Appendix A

In this appendix, we show that the parameters described in Examples 1 and 2 satisfy assumptions (A4)–(A6) stated in Theorem 1. We now discuss Example 1. It can be verified that

$$p \in (1,3], \quad q = 0, \quad 3 - (1+q)^{\frac{3-p}{p-1}} = 2, \quad 1 + 2q + (1+q)^{\frac{2}{p-1}} = 2, \quad \left(1 + 2q + (1+q)^{\frac{2}{p-1}}\right)\theta = \frac{3}{4}.$$

Thus, we can easily verify that assumptions (A4) and (A5) hold. Moreover, assumption (A6) is obtained from

$$\theta^k = \left(\frac{3}{8}\right)^k = \mathcal{O}(k^{-2}) \ \ (k \to \infty), \quad \left(\left(1 + 2q + (1+q)^{\frac{2}{p-1}}\right)\theta\right)^{\frac{k}{2}} = \left(\frac{3}{4}\right)^{\frac{k}{2}} = \mathcal{O}(k^{-2}) \ \ (k \to \infty).$$

Next, we consider Example 2. In this case, the parameter q depends on $p \in (1,3]$, and hence we denote $q = q(p)$. Moreover, we use the following notation:

$$s(p) := 3 - (1+q(p))^{\frac{3-p}{p-1}}, \quad t(p) := 1 + 2q(p) + (1+q(p))^{\frac{2}{p-1}}.$$

For any $x \in (1,3)$, we define

$$\begin{aligned} q_1(x) &:= \frac{1}{10}(2^{\frac{x-1}{3-x}} - 1) > 0, \qquad & t_1(x) &:= 1 + 2q_1(x) + (1+q_1(x))^{\frac{2}{x-1}}, \\ q_2(x) &:= \frac{1}{20}2^{\frac{3-x}{x-1}} > 0, & t_2(x) &:= 1 + 2q_2(x) + (1+q_2(x))^{\frac{2}{x-1}}. \end{aligned}$$

Recall that $q(p)$ and $t(p)$ can be represented as follows:

$$q(p) = \begin{cases} q_1(p) & \text{if } p \in (1,2], \\ q_2(p) & \text{if } p \in (2,3], \end{cases} \qquad t(p) = \begin{cases} t_1(p) & \text{if } p \in (1,2], \\ t_2(p) & \text{if } p \in (2,3]. \end{cases} \tag{A.1}$$

By the definitions of q_1 and q_2, we obtain

$$\begin{aligned} \frac{d}{dp}q_1(p) &= \frac{1}{10(3-p)^2}2^{\frac{2}{3-p}}\log 2 > 0 \quad \forall p \in (1,2], \\ \frac{d}{dp}q_2(p) &= -\frac{1}{20(p-1)^2}2^{\frac{2}{p-1}}\log 2 < 0 \quad \forall p \in (2,3]. \end{aligned} \tag{A.2}$$

Thus, the first equality of (A.1) implies that q is monotonically increasing for $p \in (1,2]$ and is monotonically decreasing for $p \in (2,3]$. Thus, using $\theta = \frac{1}{5}$ yields

$$0 < q(p) \le q(2) = \frac{1}{10} < \theta < 1 \quad \forall p \in (1,3]. \tag{A.3}$$

Since $q(p) = \min\{q_1(p), q_2(p)\} \le q_1(p) = \frac{1}{10}(2^{\frac{p-1}{3-p}} - 1) < 2^{\frac{p-1}{3-p}} - 1$ for $p \in (1,3]$, we have

$$s(p) = 3 - (1+q(p))^{\frac{3-p}{p-1}} > 3 - 2 = 1 > 0 \quad \forall p \in (1,3]. \tag{A.4}$$

Noting $\frac{2}{p-1} \ge 1$ and $q(p) > 0$ derives

$$t(p) = 1 + 2q(p) + (1+q(p))^{\frac{2}{p-1}} \ge 2 + 3q(p) \ge 1 \quad \forall p \in (1,3]. \tag{A.5}$$

Utilizing (A.2) implies

$$\frac{d}{dp}t_1(p) = \left(2 + \frac{2}{p-1}(1+q_1(p))^{\frac{3-p}{p-1}}\right)\frac{d}{dp}q_1(p) > 0 \quad \forall p \in (1,2],$$
$$\frac{d}{dp}t_2(p) = \left(2 + \frac{2}{p-1}(1+q_2(p))^{\frac{3-p}{p-1}}\right)\frac{d}{dp}q_2(p) < 0 \quad \forall p \in (2,3].$$

Hence, the second equality of (A.1) derives $t(p) \le t(2) = \frac{241}{100}$ for $p \in (1,3]$. It then follows from (A.5) and $\theta = \frac{1}{5}$ that

$$0 < \frac{1}{5} \le \theta t(p) \le \frac{241}{500} < 1 \quad \forall p \in (1,3]. \tag{A.6}$$

Therefore, the assumptions are ensured by (A.3), (A.4), (A.5), and (A.6).

Appendix B

This appendix provides the proof of Theorem 2.

Proof. We define r_3 and r_4 as follows:

$$r_3 := \min\left\{r_2, \left(\frac{m_3}{3m_2m_4}\right)^{\frac{2}{p-1}}\right\}, \quad r_4 := \frac{1}{2+m_2}\min\left\{r_3, \left(\frac{m_3}{3m_2m_4}\right)^{\frac{2}{p-1}}\right\},$$

where r_2, m_2, m_3, and m_4 are positive constants described in Lemma 3. Assume that the initial point x_0 is selected from $B(x^*, r_4)$, thus it satisfies $\|x_0 - x^*\| \le r_4$.

The proof is divided into two parts: The former part will prepare two inequalities regarding $\|d_k\|$, and the latter part will prove fast convergence of $\{x_k\}$ using those inequalities. We first show

$$\|d_k\| \le \frac{1}{3}\|d_{k-1}\|, \quad \|d_k\| \le \frac{r_4m_2}{3^k} \quad \forall k \in \mathbb{N}. \tag{B.1}$$

Let $k \in \mathbb{N}$ be arbitrary. Using (a) and (b) of Lemma 3 yields

$$\|d_k\| \le m_2m_4\mathrm{dist}(x_{k-1}, X^*)^{\frac{p+1}{2}} \le \frac{m_2m_4}{m_3}r_3^{\frac{p-1}{2}}\|d_{k-1}\| \le \frac{1}{3}\|d_{k-1}\|, \tag{B.2}$$

where note that the last inequality follows from the definition of r_3. Now, let us show that

$$x_\ell \in B(x^*, r_3)\ \forall \ell \in \{0,1,\dots,k\} \implies \|d_\ell\| \le \frac{r_4m_2}{3^\ell}\ \forall \ell \in \{0,1,\dots,k\}. \tag{B.3}$$

From (B.2) and (a) of Lemma 3, we have

$$\|d_0\| \le m_2\mathrm{dist}(x_0, X^*) \le m_2\|x_0 - x^*\| \le r_4m_2,$$
$$\|d_\ell\| \le \frac{1}{3}\|d_{\ell-1}\| \le \cdots \le \frac{1}{3^\ell}\|d_0\| \le \frac{r_4m_2}{3^\ell} \quad \forall \ell \in \{1,2,\dots,k\},$$

namely, (B.3) can be verified.

From now on, we prove by mathematical induction that $x_k \in B(x^*, r_3)$ for all $k \in \mathbb{N}$. Let us consider the case where $k = 1$. Item (a) of Lemma 3 implies $\|x_1 - x^*\| \le \|x_0 - x^*\| + \|d_0\| \le r_4 + m_2\mathrm{dist}(x_0, X^*) \le r_4(1+m_2) \le r_3$. Next, let $k \in \mathbb{N}$ be arbitrary, and we assume that $x_j \in B(x^*, r_3)$ for $j \in \{0,1,\dots,k\}$. By (B.3) and item (a) of Lemma 3, we obtain

$$\|x_{k+1} - x^*\| \le \|x_k - x^*\| + \|d_k\| \le \cdots \le \|x_0 - x^*\| + \sum_{\ell=0}^{k}\|d_\ell\| \le r_4 + \frac{r_4m_2}{2}\left(1 - \frac{1}{3^{k+1}}\right) \le \frac{2+m_2}{2}r_4 \le r_3.$$

Thus, we verify that $x_k \in B(x^*, r_3)$ for $k \in \mathbb{N}$. It then follows from (B.3) that $\|d_k\| \le \frac{r_4m_2}{3^k}$ for $k \in \mathbb{N}$, namely, the desired inequalities of (B.1) are proven.

The second part shows the local fast convergence of $\{x_k\}$. We arbitrarily take $i \in \mathbb{N}$ and $j \in \mathbb{N}$ with $i \gg j$. Using (B.1) yields

$$\|x_i - x_j\| \le \|x_{i-1} - x_j\| + \|d_{i-1}\| \le \|x_{i-2} - x_j\| + \sum_{\ell=i-2}^{i-1}\|d_\ell\| \le \cdots \le \sum_{\ell=j}^{i-1}\|d_\ell\| \le r_4m_2\sum_{\ell=j}^{i-1}\frac{1}{3^\ell} = r_4m_2\left(\frac{1}{3^{j-1}} - \frac{1}{3^{i-1}}\right) \le \frac{r_4m_2}{3^{j-1}}.$$

This fact implies that $\{x_k\}$ is a Cuachy sequence, that is, there exists $\bar{x} \in \mathbb{R}^n$ such that $x_k \to \bar{x}$ as $k \to \infty$. Meanwhile, it follows from (a) of Lemma 3 and (B.1) that $\{\mathrm{dist}(x_k, X^*)\}$ converges to zero. We note that $\|\hat{x}_k\| \le \|\hat{x}_k - x_k\| + \|x_k\| = \mathrm{dist}(x_k, X^*) + \|x_k\|$, namely, $\{\hat{x}_k\}$ is bounded. Hence, there exist $\tilde{x} \in X^*$ and $\mathcal{K} \subset \mathbb{N}$ such that $\hat{x}_k \to \tilde{x}$ as $\mathcal{K} \ni k \to \infty$. These facts imply that $\|\bar{x} - \tilde{x}\| \le \|x_k - \bar{x}\| + \mathrm{dist}(x_k, X^*) + \|\hat{x}_k - \tilde{x}\| \to 0$ as $\mathcal{K} \ni k \to \infty$, that is, $\{x_k\}$ converges to some global optimum $\bar{x} = \tilde{x} \in X^*$.

Hereinafter, we show that $\{x_k\}$ converges to $\bar{x}$ superlinearly. Combining items (a) and (b) of Lemma 3 derives

$$\|d_{j+1}\| \le m_2 \operatorname{dist}(x_{j+1}, X^*) \le m_2 m_4 \operatorname{dist}(x_j, X^*)^{\frac{p+1}{2}} \le m_2 m_4 m_3^{-\frac{p+1}{2}} \|d_j\|^{\frac{p+1}{2}}. \tag{B.4}$$

Since the first inequality of (B.1) holds, it can be verified that

$$\|d_\ell\| \le \frac{1}{3^{\ell-j-1}} \|d_{j+1}\| \quad \forall \ell \in \{j+1, j+2, \ldots, i-1\},$$
$$\|d_\ell\| \le \frac{1}{3^{\ell-j}} \|d_j\| \quad \forall \ell \in \{j+1, j+2, \ldots, i-1\}.$$

By these inequalities, we obtain

$$\|x_{j+1} - x_i\| = \left\| \sum_{\ell=j+1}^{i-1} d_\ell \right\| \le \sum_{\ell=j+1}^{i-1} \|d_\ell\| \le \sum_{\ell=j+1}^{i-1} \frac{1}{3^{\ell-j-1}} \|d_{j+1}\| = \frac{3}{2}\left(1 - \frac{1}{3^{i-j-1}}\right)\|d_{j+1}\| \le 2\|d_{j+1}\|,$$

and

$$\|x_j - x_i\| = \left\| \sum_{\ell=j}^{i-1} d_\ell \right\| \ge \|d_j\| - \sum_{\ell=j+1}^{i-1} \|d_\ell\| \ge \left(1 - \sum_{\ell=j+1}^{i-1} \frac{1}{3^{\ell-j}}\right)\|d_j\| = \frac{1}{2}\left(1 + \frac{1}{3^{i-j-1}}\right)\|d_j\| \ge \frac{1}{2}\|d_j\|.$$

Hence taking the limit $i \to \infty$ implies

$$\|x_{j+1} - \bar{x}\| \le 2\|d_{j+1}\|, \quad \|d_j\| \le 2\|x_j - \bar{x}\|. \tag{B.5}$$

Exploiting (B.4) and (B.5) derives

$$\|x_{j+1} - \bar{x}\| \le 2^{\frac{p+3}{2}} m_2 m_4 m_3^{-\frac{p+1}{2}} \|x_j - \bar{x}\|^{\frac{p+1}{2}}.$$

Therefore, from $\frac{p+1}{2} \in (1,2]$, the sequence $\{x_k\}$ converges to $\bar{x}$ superlinearly. Moreover, if $p = 3$ holds, then $\frac{p+1}{2} = 2$, that is, the rate of convergence is quadratic. □

Data availability

No data was used for the research described in the article.

References

[1] N. Doikov, K. Mishchenko, Y. Nesterov, Super-universal regularized Newton method, SIAM J. Optim. 34 (2024) 27–56.
[2] D. Li, M. Fukushima, L. Qi, N. Yamashita, Regularized Newton methods for convex minimization problems with singular solutions, Comput. Optim. Appl. 28 (2004) 131–147.
[3] Y.-J. Li, D.-H. Li, Truncated regularized Newton method for convex minimizations, Comput. Optim. Appl. 43 (2009) 119–131.
[4] R.A. Polyak, Regularized Newton method for unconstrained convex optimization, Math. Program. 120 (2009) 125–145.
[5] K. Mishchenko, Regularized Newton method with global $\mathcal{O}(1/k^2)$ convergence, SIAM J. Optim. 33 (2023) 1440–1462.
[6] Y. Nesterov, B.T. Polyak, Cubic regularization of Newton method and its global performance, Math. Program. 108 (2006) 177–205.
[7] Y. Nesterov, Accelerating the cubic regularization of Newton's method on convex problems, Math. Program. 112 (2008) 159–181.
[8] M.-C. Yue, Z. Zhou, A.M.-C. So, On the quadratic convergence of the cubic regularization method under a local error bound condition, SIAM J. Optim. 29 (2019) 90–932.
[9] D. Goldfarb, Y. Ren, A. Bahamou, Practical quasi-Newton methods for training deep neural networks, in: Advances in Neural Information Processing Systems, 2020, pp. 2386–2396.
[10] L. Grippo, F. Lampariello, S. Luclidi, A nonmonotone line search technique for Newton's method, SIAM J. Numer. Anal. 23 (1986) 707–716.
[11] A. Rodomanov, Y. Nesterov, Greedy quasi-Newton methods with explicit superlinear convergence, SIAM J. Optim. 31 (2021) 785–811.
[12] R. Crane, F. Roosta, DINO: distributed Newton-type optimization method, in: Proceedings of the 37th International Conference on Machine Learning, 2020, pp. 2174–2184.
[13] N. Doikov, P. Richtárik, Randomized block cubic Newton method, in: Proceedings of the 35th International Conference on Machine Learning, PMLR, 2018, pp. 1290–1298.
[14] R.M. Gower, D. Kovalev, F. Lieder, P. Richtárik, RSN: randomized subspace Newton, in: Advances in Neural Information Processing Systems, 2019, pp. 616–625.
[15] F. Hanzely, N. Doikov, P. Richtárik, Y. Nesterov, Stochastic subspace cubic Newton method, in: Proceedings of the 37th International Conference on Machine Learning, 2020, pp. 4027–4038.
[16] U. Marteau-Ferey, F. Bach, A. Rudi, Globally convergent Newton methods for ill-conditioned generalized self-concordant losses, in: Advances in Neural Information Processing Systems, 2019, pp. 7636–7646.
[17] A. Rodomanov, D. Kropotov, A superlinearly-convergent proximal Newton-type method for the optimization of finite sums, in: Proceedings of the International Conference on Machine Learning, PRML, 2016, pp. 2597–2605.
[18] S. Soori, K. Mishchenko, A. Mokhtari, M.M. Dehnavi, M. Gürbüzbalaban, DAve-QN: a distributed averaged quasi-Newton method with local superlinear convergence rate, in: Proceedings of the International Conference on Artificial Intelligence and Statistics, 2020, pp. 1965–1976.
[19] S. Ariizumi, Y. Yamakawa, N. Yamashita, Convergence properties of Levenberg-Marquardt methods with generalized regularization terms, Appl. Math. Comput. 463 (2024) 128365.
[20] D.P. Bertsekas, Convex Optimization Algorithms, Athena Scientific, Massachusetts, 2015.