Open Journal of Optimization, 2016, 5, 44-52
Published Online March 2016 in SciRes. http://www.scirp.org/journal/ojop
http://dx.doi.org/10.4236/ojop.2016.51006

A Regularized Newton Method with Correction for Unconstrained Convex Optimization

Liming Li, Mei Qin*, Heng Wang

College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: *usstqinmaynew@sina.com

Received 17 January 2016; accepted 12 March 2016; published 15 March 2016

Abstract

In this paper, we present a regularized Newton method (M-RNM) with correction for minimizing a convex function whose Hessian matrices may be singular. At every iteration, not only a RNM step is computed but also two correction steps are computed. We show that if the objective function is LC^2, then the method posses globally convergent. Numerical results show that the new algorithm performs very well.

Keywords

Regularied Newton Method, Correction Technique, Trust Region Technique, Unconstrained Convex Optimization

1. Introduction

We consider the unconstrained optimization problem [1]-[3]

$$\min_{x\in R^n} f(x), \tag{1.1}$$

where $f: R^n \to R$ is twice continuously differentiable, whose gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ are denoted by $g(x)$ and $H(x)$ respectively. Throughout this paper, we assume that the solution set of (1.1) is nonempty and denoted by *X*, and in all cases $\|\cdot\|$ refers to the 2-norm.

It is well known that $f(x)$ is convex if and only if $H(x)$ is symmetric positive semidefinite for all $x \in R^n$. Moreover, if $f(x)$ is convex, then $x \in X$ if and only if *x* is a solution of the system of nonlinear

*Corresponding author.

How to cite this paper: Li, L.M., Qin, M. and Wang, H. (2016) A Regularized Newton Method with Correction for Unconstrained Convex Optimization. *Open Journal of Optimization*, **5**, 44-52. http://dx.doi.org/10.4236/ojop.2016.51006

equations

$$g(x)=0. \tag{1.2}$$

Hence, we could get the minimizer of $f(x)$ by solving (1.2) [4]-[8]. The Newton method is one of efficient solution methods. At every iteration, it computes the trial step

$$d_k^N=-H_k^{-1}g_k, \tag{1.3}$$

where $g_k=g(x_k)$ and $H_k=H(x_k)$. As we know, if H_k is Lipschitz continuous and nonsingular at the solution, then the Newton method has quadratic convergence. However, this method has an obvious disadvantage when the H_k is singular or near singular. In this case, we may compute the Moore-Penrose step [7] $d_k^{MP}=-H_k^{+}g_k$. But the computation of the singular value decomposition to obtain H_k^{+} is sometimes prohibitive. Hence, computing a direction that is close to d_k^{MP} may be a good idea.

To overcome the difficulty caused by the possible singularity of H_k, [9] proposed a regularized Newton method, where the trial step is the solution of the linear equations

$$(H_k+\lambda_k I)d=-g_k, \tag{1.4}$$

where I is the identity matrix. μ_k is a positive parameter which is updated from iteration to iteration.

Now we need to consider another question, "how to choose the regularized parameter μ_k?" Yamashita and Fukushima [10] chose $\lambda_k=\|g_k\|^2$ and showed that the regularized Newton method has quadratic convergence under the local error bound condition which is weaker than nonsingularity. Fan and Yuan [11] took $\lambda_k=\|g_k\|^{\delta}$ with $\delta\in[1,2]$ and showed that the Levenberg-Marqulardt method preserves the quadratic convergence. Numerial results ([12] [13]) show that the choice of $\lambda_k=\|F_k\|$ performs more stable and preferable.

Inspired by the regularized Newton method [13] with correction for nonlinear equations, we propose a modified regularized Newton method in this paper. At every iteration, the modified regularized Newton method first solves the linear equations

$$(H_k+\lambda_k I)d=-g_k \quad \text{with } \lambda_k=\mu_k\|g_k\| \tag{1.5}$$

to obtain the Newton step d_k, where μ_k is updated from iteration to iteration, and solves the linear equations

$$(H_k+\lambda_k I)d=-g_k+\lambda_k d_k \tag{1.6}$$

to obtain the approximate Newton step s_k.

It is easy to see

$$s_k=d_k+\widetilde{d_k}, \quad \widetilde{d_k}=\lambda_k(H_k+\lambda_k I)^{-1}d_k. \tag{1.7}$$

Then it solves the linear equations

$$(H_k+\lambda_k I)s=-g(y_k) \quad \text{with } y_k=x_k+s_k \tag{1.8}$$

to obtain the approximate Newton step $\widetilde{s_k}$.

The aim of this paper is to study the convergence properties of the above modified regularized Newton method and do a numerical experiment to test its efficiency.

The paper is organized as follows. In Section 2, we present a new regularized Newton algorithm with correction by trust region technique, and then prove the global convergence of the new algorithm under some suitable conditions. In Section 3, we test the regularized Newton algorithm with correction and compared it with a regularized Newton method. Finally, we conclude the paper in Section 4.

2. The Algorithm and Its Global Convergence

In this section, we first present the new modified regularized Newton algorithm by using trust region technique, then prove the global convergence. First, we give the modified regularized Newton algorithm.

Let s_k and $\widetilde{s_k}$ be given by (1.6) and (1.8), respectively. Since the matrix $H_k+\lambda_k I$ is symmetric and positive definite, s_k is a descent direction of $f(x)$ at x_k, however $s_k+\widetilde{s_k}$ may not be. Hence we prefer to use a trust region technique.

Define the actual reduction of $f(x)$ at the kth iteration as

$$Ared_k = f(x_k) - f\left(x_k + s_k + \widetilde{s_k}\right). \tag{2.1}$$

Note that the regularization step d_k is the minimizer of the convex minimization problem

$$\min_{d \in R^n} \frac{1}{2} d^{\mathrm{T}} H_k d + g_k^{\mathrm{T}} d + \frac{1}{2} \lambda_k \|d\|^2 .$$

If we let

$$\Delta_{k,1} = \|d_k\| = \left\| -\left(H_k + \lambda_k I\right)^{-1} g_k \right\|,$$

then it can be proved [4] that d_k is also a solution of the trust region problem

$$\min_{d \in R^n} \varphi(d) = \frac{1}{2} d^{\mathrm{T}} H_k d + g_k^{\mathrm{T}} d,$$
$$s.t. \quad \|d\| \le \Delta_{k,1}.$$

By the famous result given by Powell in [14], we know that

$$\varphi(0) - \varphi(d_k) \ge \frac{1}{2} \|g_k\| \min\left\{ \|d_k\|, \frac{\|g_k\|}{\|H_k\|} \right\}. \tag{2.2}$$

By some simple calculations, we deduce from (1.7) that

$$\begin{aligned}
\varphi(d_k) - \varphi(s_k) &= g_k^{\mathrm{T}} d_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k - g_k^{\mathrm{T}} s_k - \frac{1}{2} s_k^{\mathrm{T}} H_k s_k \\
&= -g_k^{\mathrm{T}} \widetilde{d_k} - \frac{1}{2} \widetilde{d_k}^{\mathrm{T}} H_k \widetilde{d_k} - \widetilde{d_k}^{\mathrm{T}} H_k d_k \\
&= \lambda_k \widetilde{d_k}^{\mathrm{T}} d_k - \frac{1}{2} \widetilde{d_k}^{\mathrm{T}} H_k \widetilde{d_k} \\
&= \frac{1}{2} \widetilde{d_k}^{\mathrm{T}} H_k \widetilde{d_k} + \lambda_k \widetilde{d_k}^{\mathrm{T}} \widetilde{d_k} \\
&\ge 0,
\end{aligned}$$

so, we have

$$\varphi(0) - \varphi(s_k) \ge \varphi(0) - \varphi(d_k). \tag{2.3}$$

Similar to d_k, $\widetilde{s_k}$ is not only the minimizer of the problem

$$\min_{s \in R^n} g(y_k)^{\mathrm{T}} s + \frac{1}{2} s^{\mathrm{T}} \left(H_k + \lambda_k I\right) s$$

but also a solution to the trust region problem

$$\min_{s \in R^n} \phi(s) = \frac{1}{2} s^{\mathrm{T}} H_k s + g(y_k)^{\mathrm{T}} s,$$
$$s.t. \quad \|s\| \le \Delta_{k,2},$$

where $\Delta_{k,2} = \left\| -\left(H_k + \lambda_k I\right)^{-1} g(y_k) \right\| = \left\| \widetilde{s_k} \right\|$. Therefore we also have

$$\phi(0) - \phi\left(\widetilde{s_k}\right) \ge \frac{1}{2} \|g(y_k)\| \min\left\{ \left\|\widetilde{s_k}\right\|, \frac{\|g(y_k)\|}{\|H_k\|} \right\}. \tag{2.4}$$

Based on the inequalities (2.2), (2.3) and (2.4), it is reasonable for us to define the new predicted reduction as

$$Pred_k = \varphi(0) - \varphi(s_k) + \phi(0) - \phi\left(\widetilde{s_k}\right), \tag{2.5}$$

which satisfies

$$Pred_k \geq \frac{1}{2}\|g_k\|\min\left\{\|d_k\|, \frac{\|g_k\|}{\|H_k\|}\right\} + \frac{1}{2}\|g(y_k)\|\min\left\{\|\widetilde{s_k}\|, \frac{\|g(y_k)\|}{\|H_k\|}\right\}. \tag{2.6}$$

The ratio of the actual reduction to the predicted reduction

$$r_k = \frac{Ared_k}{Pred_k}, \tag{2.7}$$

plays a key role in deciding whether to accept the trial step and how to adjust the regularized parameter.

The regularized Newton algorithm with correction for unconstrained convex optimization problems is stated as follows.

Algorithm 2.1

Step 1. Given $x_0 \in R^n$, $\varepsilon \geq 0$, $\mu_0 > m > 0$, $0 < p_0 \leq p_1 \leq p_2 < 1$. Set $k := 0$.

Step 2. If $\|g_k\| \leq \varepsilon$, then stop.

Step 3. Compute $\lambda_k = \mu_k \|g_k\|$.

Solve

$$(H_k + \lambda_k I)d = -g_k \tag{2.8}$$

to obtain d_k.

Solve

$$(H_k + \lambda_k I)d = -g_k + \lambda_k d_k \tag{2.9}$$

to obtain s_k and set

$$y_k = x_k + s_k.$$

Solve

$$(H_k + \lambda_k I)s = -g(y_k) \tag{2.10}$$

to obtain $\widetilde{s_k}$ and set

$$t_k = s_k + \widetilde{s_k}$$

Step 4. Compute $r_k = \frac{Ared_k}{Pred_k}$. Set

$$x_{k+1} = \begin{cases} x_k + t_k, & \text{if } r_k \geq p_0, \\ x_{k,} & \text{otherwise.} \end{cases} \tag{2.11}$$

Step 5. Choose μ_{k+1} as

$$\mu_{k+1} = \begin{cases} 4\mu_k, & \text{if } r_k < p_1, \\ \mu_k, & \text{if } r_k \in [p_1, p_2], \\ \max\{\mu_k/4, m\}, & \text{if } r_k + p_2. \end{cases} \tag{2.12}$$

Set $k := k+1$ and go step 2.

Before discussing the global convergence of the algorithm above, we make the following assumption.

Assumption 2.1. $g(x)$ and $H(x)$ are both Lipschitz continuous, that is, there exists a constant $L_1 > 0$, $L_2 > 0$ such that

$$\|g(y) - g(x)\| \leq L_1 \|y - x\|, \forall x, y \in R^n \tag{2.13}$$

and

$$\|H(y) - H(x)\| \leq L_2 \|y - x\|, \forall x, y \in R^n. \tag{2.14}$$

It follows from (2.14) that

$$\|g(y)-g(x)-H(x)(y-x)\| \le L_2\|y-x\|^2, \forall x, y \in R^n. \tag{2.15}$$

The following lemma given below shows the relationship between the positive semidefinite matrix and symmetric positive semidefinite matrix.

Lemma 2.1. A real-valued matrix A is positive semidefinite if and only if $\left(A+A^{\mathrm{T}}\right)/2$ is positive semidefinite.

Proof. See [4]. ◇

Next, we give the bounds of a positive definite matrix and its inverse.

Lemma 2.2. Suppose A is positive semidefinite. Then,

$$\|A+\varphi I\| \ge \varphi$$

and

$$\left\|(A+\varphi I)^{-1}\right\| \le \varphi^{-1}$$

hold for any $\varphi > 0$.

Proof. See [13]. ◇

Theorem 2.1. Under the conditions of Assumption 2.1, if f is bounded below, then Algorithm 2.1 terminates in finite iterations or satisfies

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{2.16}$$

Proof. We prove by contradiction. If the theorem is not true, then there exists a positive τ and an integer $\tilde{k}$ such that

$$\|g_k\| \ge \tau, \forall k \ge \tilde{k}. \tag{2.17}$$

Without loss of generality, we can suppose $\tilde{k}=1$. Set $T=\{k \mid x_k \ne x_{k+1}\}$. Then

$$\{1,2,\cdots\} = T \cup \{k \mid x_k = x_{k+1}\}.$$

Now we will analysis in two cases whether T is finite or not.

Case (1): T is finite. Then there exists an integer k_1 such that

$$x_{k_1} = x_{k_1+1} = x_{k_1+2} = \cdots.$$

By (2.11), we have

$$r_k < p_0, \forall k \ge k_1.$$

Therefore by (2.12) and (2.17), we deduce

$$\mu_k \to \infty, \lambda_k \to \infty. \tag{2.18}$$

Since $x_{k+1} = x_k$, $\forall k \ge k_1$, we get from (2.8) and (2.18) that

$$\|d_k\| = \left\|-(H_k+\lambda_k I)^{-1} g_k\right\| \le \lambda_k^{-1}\|g_k\| \to 0. \tag{2.19}$$

Duo to (1.7), we get

$$\|s_k\| = \left\|d_k+\widetilde{d_k}\right\| \le 2\|d_k\|, \quad \|s_k\| \to 0.$$

From (2.10), we obtain

$$\begin{aligned}
\left\|\widetilde{s_k}\right\| &= \left\|-(H_k+\lambda_k I)^{-1} g(y_k)\right\| \\
&\le \left\|(H_k+\lambda_k I)^{-1}\left(g(y_k)-g_k-H_k s_k\right)\right\| \\
&\quad + \left\|(H_k+\lambda_k I)^{-1} g_k\right\| + \left\|(H_k+\lambda_k I)^{-1} H_k s_k\right\| \\
&\le L_2\lambda_k^{-1}\|s_k\|^2 + \|d_k\| + \|s_k\| \\
&\le \gamma_1\|d_k\|,
\end{aligned} \tag{2.20}$$

where γ_1 is a positive constant.

It follows from (2.1) and (2.5) that

$$\begin{aligned}\left|Ared_k - Pred_k\right| &= \left|f(x_k) - f\left(x_k + s_k + \widetilde{s_k}\right) - \left(\varphi(0) - \varphi(s_k) + \phi(0) - \phi\left(\widetilde{s_k}\right)\right)\right| \\ &\le \left|f\left(y_k + \widetilde{s_k}\right) - f(y_k) - \frac{1}{2}\widetilde{s_k}^{\mathrm{T}} H_k \widetilde{s_k} - g(y_k)^{\mathrm{T}}\widetilde{s_k}\right| \\ &\quad + \left|f(y_k) - f(x_k) - \frac{1}{2}s_k^{\mathrm{T}} H_k s_k - g_k^{\mathrm{T}} s_k\right| \\ &\le o\left(\|s_k\|^2\right) + o\left(\left\|\widetilde{s_k}\right\|^2\right).\end{aligned} \tag{2.21}$$

Moreover, from (2.6), (2.17), (2.13) and (2.19), we have

$$Pred_k \ge \frac{1}{2}\tau \min\left\{\|d_k\|, \frac{\tau}{L_1}\right\} \ge \frac{1}{2}\tau\|d_k\|, \tag{2.22}$$

for sufficiently large *k*.

Duo to (2.21) and (2.22), we get

$$\begin{aligned}|r_k - 1| &= \left|\frac{Ared_k - Pred_k}{Pred_k}\right| \\ &\le \left|\frac{f(x_k) - f\left(x_k + s_k + \widetilde{s_k}\right) - \left(\varphi(0) - \varphi(s_k) + \phi(0) - \phi\left(\widetilde{s_k}\right)\right)}{\frac{1}{2}\tau \min\left\{\|d_k\|, \frac{\tau}{L_1}\right\}}\right| \\ &\le \frac{o\left(\|s_k\|^2\right) + o\left(\left\|\widetilde{s_k}\right\|^2\right)}{\|d_k\|} \to 0,\end{aligned} \tag{2.23}$$

which implies that $r_k \to 1$. Hence, there exists positive constant γ_2 such that $\mu_k \le \gamma_2$, holds for all large *k*, which contradicts to (2.18).

Case (2): *T* is infinite. Then we have from (2.6) and (2.17) that

$$\begin{aligned}\infty &> f(x_1) - \liminf_{k\to\infty} f(x_k) \ge \sum_{i=1}^{\infty}\left(f(x_i) - f(x_{i+1})\right) \\ &= \sum_{k\in T}\left(f(x_k) - f(x_{k+1})\right) \ge \sum_{k\in T} p_0 Pred_k \\ &\ge \sum_{k\in T} p_0\left(\frac{1}{2}\|g_k\|\min\left\{\|d_k\|, \frac{\|g_k\|}{\|H_k\|}\right\} + \frac{1}{2}\|g(y_k)\|\min\left\{\left\|\widetilde{s_k}\right\|, \frac{\|g(y_k)\|}{\|H_k\|}\right\}\right) \\ &\ge \sum_{k\in T} p_0\frac{\tau}{2}\min\left\{\|d_k\|, \frac{\tau}{L_1}\right\},\end{aligned} \tag{2.24}$$

which implies that

$$\lim_{k\to\infty, k\in T} d_k = 0. \tag{2.25}$$

The above equality together with the updating rule of (2.12) means

$$\lambda_k \to \infty. \tag{2.26}$$

Similar to (2.20), it follows from (2.25) and (2.26) that

$$\left\|\widetilde{s_k}\right\| \le \gamma_3\|d_k\|, \|s_k\| \le 2\|d_k\|, \ \forall k \in T$$

for some positive constant γ_3. Then we have

$$\|t_k\| \le \|s_k\| + \|\widetilde{s_k}\| \le (\gamma_3 + 2)\|d_k\|, \ \forall k \in T.$$

This equality together with (2.24) yields

$$\sum_{k\in T} \|t_k\| < \infty,$$

which implies that

$$x_k \to x^*. \tag{2.27}$$

It follows from (2.8), (2.27), (2.26) and (2.20) that

$$s_k \to 0, \ \widetilde{s_k} \to 0. \tag{2.28}$$

Since $(H_k + \mu_k \|g_k\| I) d_k = -g_k$ from (2.8), we have from (2.13), (2.17) and (2.28) that

$$1 \le \frac{\|H_k\|}{\|g_k\|}\|d_k\| + \mu_k \|d_k\| \le \frac{L_1}{\tau}\|d_k\| + \mu_k \|d_k\|,$$

which means

$$\mu_k \to \infty. \tag{2.29}$$

By the same analysis as (2.23) we know that

$$r_k \to 1. \tag{2.30}$$

Hence, there exists a positive constant $\gamma_4 > m$ such that $\mu_k \le \gamma_4$ holds for all sufficiently large *k*, which gives a contradiction to (2.29). The proof is completed. ◇

3. Numerical Experiments

In this section, we test the performance of Algorithm 2.1 on the unconstrained nonlinear optimization problem, and compared it with a regularized Newton method without correction. The function to be minimized is

$$f(x) = \frac{1}{2}\sum_{i=1}^{n-1}(x_i - x_{i+1})^2 + \frac{1}{12}\sum_{i=1}^{n-1}\alpha_i (x_i - x_{i+1})^4, \tag{3.1}$$

where $\alpha_i \ge 0 \ (i = 1,\cdots,n-1)$ are constants. It is clear that function $f(x)$ is convex and the minimizer set of $f(x)$ is

$$S = \{x \in R^n \mid x_1 = x_2 = \cdots = x_n\}$$

The Hessian $\nabla^2 f(x)$ is given by

$$\nabla^2 f(x) = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix} + \begin{pmatrix} a_1 & -a_1 & & & \\ -a_1 & a_1 + a_2 & -a_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -a_{n-2} & a_{n-2} + a_{n-1} & -a_{n-1} \\ & & & -a_{n-1} & a_{n-1} \end{pmatrix},$$

where $a_i = a_i(x) = a(x_i - x_{i+1})^2$, $(i = 1,2,\cdots,n-1)$. Matrix $\nabla^2 f(x)$ is positive semidefinite for all *x*, but singular as the sum of every column is zero. Since the Hessian H_k is always singular, the Newton method cannot be used to solve nonlinear Equations (1.2). But by using the regularization technique, both regularized Newton method and Algorithm 2.1 work quite well.

The aims of the experiments are as follows: to check whether Algorithm 2.1 converges quadratically as stated in Section 3 and also to see how well the technique of correction works. We set $p_0 = 0.001$, $p_1 = 0.25$, $p_2 = 0.75$, $p_1 = 0.25$, $p_3 = 4$, $\mu_0 = 10^{-2}$ and $m = \varepsilon = 10^{-5}$ for Algorithm 2.1.

Table 1 reports the norms of g_k at every iteration when $n = 10$, $\alpha_i = 1(i = 1,2,\cdots,n-1)$,

$\beta_i = 1(i = 1, 2, \cdots, n-1)$ and $x_0 = (1, 2, \cdots, 10)^T$. Algorithm 2.1 only take four iterations to obtain the minimizer of $f(x)$; $\|g_k\|$ decreases very quickly. The results show the sequence $\{\|g_k\|\}$ quadratic convergence. The iteration is as follows

$$x_1 = (3.2142, 3.4357, 3.8874, 4.4813, 5.1523, 5.8477, 6.5187, 7.1126, 7.5643, 7.7858)^T$$

$$x_2 = (5.2666, 5.2742, 5.3159, 5.3859, 5.4690, 5.5391, 5.6161, 5.6841, 5.7258, 5.7334)^T$$

$$x_3 = (5.4991, 5.4991, 5.4992, 5.4995, 5.4998, 5.5002, 5.5005, 5.5008, 5.5009, 5.5009)^T$$

$$x_4 = (5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000)^T$$

$$x_5 = (5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000, 5.5000)^T.$$

We may observe that the whole sequence $\{x_k\}$ converges to $x^* = (5, 5, \cdots, 5, 5)^T$.

We also ran the regularized Newton algorithm (RNA) without correction, that is, we do not solve the linear equations (2.9)-(2.10) and just set the solution of (2.8) to be the trial step. Then, we tested the regularized Newton algorithm without correction and Algorithm 2.1 for various of *n*, α_i and different choices of the starting point. The results are listed in **Table 2**. α_i: the selected value of α_i; Dim: the dimension *n* of the problem; x_0: the *i*th element x_0; niter: the number of iterations required; $\|\nabla f\|$: the final value of $\|\nabla f(x_k)\|$; x^*: the final value of x_k. We use $\|\nabla f(x_k)\| \le 10^{-5}$ as the stopping criterion.

Table 1. Results of Algorithm 2.1 to test quadratic convergence.

k	0	1	2	3	4	5
$\|g_k\|$	1.8856	0.4890	0.0315	1.0368e−05	5.6523e−15	0

Table 2. Results of RNA and Algorithm 2.1.

α_i	*Dim*	x_0	niter	$\|\nabla f\|$	x^*
0	10	*i*	3/1	1.63e−07	5.5
		1/*i*	2/1	1.85e−06	0.2929
	50	*i*	13/4	4.28e−06	25.5
		1/*i*	19/6	4.28e−06	0.09
	100	*i*	16/3	1.62e−06	50.5
		1/*i*	5/2	1.15e−08	0.519
	500	*i*	49/6	6.31e−07	250.5
		1/*i*	18/8	2.08e−08	0.0136
1	10	*i*	5/1	2.66e−06	5.5
		1/*i*	8/2	5.6e−12	0.2929
	50	*i*	7/3	3.84e−10	25.5
		1/*i*	25/14	4.86e−09	0.09
	100	*i*	8/2	4.95e−06	50.5
		1/*i*	11/5	2.30e−06	0.519
	500	*i*	38/19	3.45e−07	250.5
		1/*i*	11/5	5.48e−06	0.0136
i	10	*i*	9/4	9.56e−08	5.5
		1/*i*	5/1	9.00e−06	0.2929
	50	*i*	39/16	2.77e−07	25.5
		1/*i*	19/10	3.10e−08	0.09
	100	*i*	59/35	4.85e−07	50.5
		1/*i*	19/10	2.81e−06	0.519
	500	*i*	45/23	6.20e−07	250.5
		1/*i*	19/10	1.56e−06	0.0136

Moreover, we can see for the same α_i, n and x_0, the number of iterations of Algorithm 2.1 is always less than that of RNA. And the correction term does help to improve RNA when the initial point is far away from the minimizer. These facts indicate that the introduction of correction is really useful and could accelerate the convergence of the regularized Newton method.

4. Concluding Remarks

In this paper, we propose a regularized Newton method with correction for unconstrained convex optimization. At every iteration, not only a RNM step is computed but also two correction steps are computed which make use of the previous available Jacobian instead of computing the new Jacobian. Numerical experiments suggest that the introduction of correction is really useful.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (11426155) and the Hujiang Foundation of China (B14005).

References

[1] Yang, W.W., Yang, Y.T., Zhang, C.H. and Cao, M.Y. (2013) A Newton-Like Trust Region Method for Large-Scale Unconstrained Nonconvex Minimization. *Abstract and Applied Analysis*, **2013**.

[2] Polyak, R.A. (2009) Regularized Newton Method for Unconstrained Convex Optimization. *Mathematical Programming Series B*, **120**, 125-145. http://dx.doi.org/10.1007/s10107-007-0143-3

[3] Shen, C.G., Chen, X.D. and Liang, Y.M. (2012) A Regularized Newton Method for Degenerate Unconstrained Optimization Problems. *Optimization Letters*, **6**, 1913-1933. http://dx.doi.org/10.1007/s11590-011-0386-z

[4] Sun, W. and Yuan, Y. (2006) Optimization Theory and Methods. Springer Science and Business Media, LLC, New York.

[5] Zhou, W. and Li, D. (2008) A Globally Convergent BFGS Method for Nonlinear Monotone Equations without Any merit Functions. *Mathematics of Computation*, **77**, 2231-2240. http://dx.doi.org/10.1090/S0025-5718-08-02121-2

[6] Argyros, I.K. and Hilout, S. (2012) On the Convergence of Damped Newton Method. *Applied Mathematics and Computation*, **219**, 2808-2824. http://dx.doi.org/10.1016/j.amc.2012.09.011

[7] Nashed, M.Z. and Chen, X. (1993) Convergence of Newton-Like Method for Singular Operator Equations Using Outer Inverses. *Numerische Mathematik*, **66**, 235-257. http://dx.doi.org/10.1007/BF01385696

[8] Kelley, C.T. (1999) Iterative Methods for Optimization. In: *Frontiers in Applied Mathematics*, Vol. 18, SIAM, Philadelphia, 2. http://dx.doi.org/10.1137/1.9781611970920

[9] Sun, D. (1999) A Regularization Newton Method for Solving Nonlinear Complementarity Problems. *Applied Mathematics and Optimization*, **40**, 315-339. http://dx.doi.org/10.1007/s002459900128

[10] Li, D.H., Fukushima, M., Qi, L. and Yamashita, N. (2004) Regularized Newton Methods for Convex Minimization Problems with Singular Solutions. *Computational Optimization and Applications*, **28**, 131-147. http://dx.doi.org/10.1023/B:COAP.0000026881.96694.32

[11] Fan, J.Y. and Yuan, Y.X. (2005) On the Quadratic Convergence of the Levenberg-Marquardt Method without Nonsingularity Assumption. *Computing*, **74**, 23-39. http://dx.doi.org/10.1007/s00607-004-0083-1

[12] Fan, J.Y. and Pan, J.Y. (2009) A Note on the Levenberge-Marquardt Parameter. *Applied Mathematics and Computation*, **207**, 351-165. http://dx.doi.org/10.1016/j.amc.2008.10.056

[13] Fan, J.Y. and Yuan, Y.X. (2014) A Regularized Newton Method for Monotone Nonlinear Equations and Its Application. *Optimization Methods and Software*, **29**, 102-119. http://dx.doi.org/10.1080/10556788.2012.746344

[14] Powell, M.J.D. (1975) Convergence Properties of a Class of Minimization Algorithms. In: Mangasarian, O.L., Meyer, R.R. and Robinson, S.M., Eds., *Nonlinear Programming*, Vol. 2, Academic Press, New York, 1-27. http://dx.doi.org/10.1016/b978-0-12-468650-2.50005-5