

非凸优化问题的两步正则化牛顿法

朱俊霖

长沙理工大学数学与统计学院, 湖南 长沙

收稿日期: 2023年7月21日; 录用日期: 2023年8月13日; 发布日期: 2023年8月21日

摘要

本文提出了非凸的无约束优化问题的一种在信赖域框架下的两步正则化牛顿算法, 其在适当条件下证明了该方法具有局部收敛性。在局部误差界的条件下, 该方法具有三阶收敛速度。此外我们还进行了数值实验, 数值结果显示, 与单步正则化牛顿法相比我们有更少的迭代次数更快的迭代速度, 说明两步正则化牛顿法比后者更有效。

关键词

非凸优化, 正则化牛顿法, 局部误差界, 信赖域

The Two-Step Regularized Newton Method for Non-Convex Optimization Problems

Junlin Zhu

School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha Hunan

Received: Jul. 21st, 2023; accepted: Aug. 13th, 2023; published: Aug. 21st, 2023

Abstract

In this paper, we propose a two-step regularized Newton algorithm for solving non-convex unconstrained optimization problems within the trust region framework. Under appropriate conditions, we prove that this method possesses local convergence. Under the condition of a local error bound, the method exhibits third-order convergence rate. Additionally, numerical experiments are conducted, and the results show that our two-step regularized Newton method outperforms the single-step regularized Newton method in terms of fewer iterations and faster convergence speed, indicating its higher efficiency.

Keywords

Non-Convex Optimization, Regularized Newton Method, Local Error Bound, Trust Region



1. 引言

本文主要考虑如下的优化问题

$$\min_{x \in R^n} f(x) \quad (1)$$

其中 $f: R^n \rightarrow R$ 是二次连续的可微函数。牛顿法是求解问题(1)的经典算法之一, 其具有如下的迭代格式:

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

其中 $\nabla f(x^k)$ 为一阶梯度, $\nabla^2 f(x^k)$ 为二阶海瑟矩阵。

牛顿法是求解问题(1)的经典方法, 他的突出特点是: 当初始点离最优解较近且目标函数在最优处的 Hessian 阵非奇异而且在最优解附近的 Hessian 矩阵满足 Lipschitz 条件时, 经典牛顿法具有快速的二次收敛速度, 收敛速度快是牛顿法的显著特点。但是当初始点的选择离最优点较远的时候, 牛顿法可能不收敛, 通常采取正则化来保证牛顿法的全局收敛性。此外, 当 Hessian 矩阵奇异的时候, 牛顿法无法使用。2001 年 Yamashita 和 Fukushima [1] 证明了当 $\|\nabla f(x)\|$ 在问题(1)的最优解 x^* 的领域内, 在提供局部误差界的假设下, 证明了求解非线性方程组的 Levenberg-marquardt 方法在比非奇异假设更弱的局部误差界的条件下仍然具有局部的二次收敛速度。因此, 研究在比 Hessian 阵非奇异假设更弱的局部误差界条件下牛顿算法有着研究价值, 本文将进一步研究求解非凸优化问题的两步正则化牛顿法。

对于奇异凸优化问题, 2004 年 Li [2] 等人通过近似计算如下线性方程组得到搜索方向 d_k :

$$(G_k + \mu_k I)d_k + g_k = r_k, \|r_k\| \leq \alpha \mu_k \|d_k\|.$$

其中 $\mu_k = C \|g(x_k)\|$, Li 等人证明了该算法在求解奇异凸优化问题时具有全局收敛性并且在局部误差界的条件下其仍具有二次收敛速度。

2009 年 Ueda 和 Yamashita, 在文献[3]中对 Levenberg-marquardt 正则化进行了推广, 当目标函数 f 是非凸函数的时候, 使用 Armijo 步长规则, 设 x_k 为第 k 个迭代点, 将梯度 $\nabla f(x^k)$ 和二阶海瑟矩阵 $\nabla^2 f(x^k)$, 分别用 g_k 和 H_k 表示, 令正则化参数 $\mu_k = c_1 \psi_k I + c_2 \|g_k\|^\delta I$, 其中, $\psi_k = \max(0, -\lambda_{\min}(H_k))$, λ 为矩阵 H_k 的特征值, $d_k = -(H_k + c_1 \psi_k I + c_2 \|g_k\|^\delta I)^{-1}$ 为下降方向, $0 < \delta < 1$, 并证明了该算法在适当的条件下有着全局收敛性, 并且具有超线性收敛速度, 并加以延申, 当 $f(x)$ 是强凸的且 $\delta = 1$ 时该方法二阶收敛。

Homeier 在文献[4] [5] 里根据牛顿法提出了一种改进的牛顿法——两阶段牛顿法, 算法如下:

$$\begin{cases} y_n = x_n - F'(x_n)^{-1} F(x_n), & n \geq 0, x_0 \in D; \\ x_{n+1} = x_n - F'(z_n)^{-1} F(x_n), & z_n = \frac{x_n + y_n}{2}, \end{cases}$$

用以解决具有 F -可导的非线性方程 $F(x) = 0$ 的近似解问题, 其在特定的情况下有着全局收敛性以及局部收敛性, 为我们提出两步牛顿法提供了理论支持。

Zhou 和 Chen 在文献[6]中提出了一个基于凸函数无约束优化问题的两步正则化牛顿法,

$$\begin{aligned} (G_k + \lambda_k I)d &= -g_k, \\ (G_k + \lambda_k I)\bar{d} &= -g(y_k), y_k = x_k + d_k, \end{aligned}$$

$$x_{k+1} = x_k + s_k,$$

该方法结合了正则化牛顿法与信赖域方法, 通过证明该算法求解凸的无约束优化问题在具有全局收敛和局部误差界的条件下具有三次收敛速度。

根据上述研究, 我们发现凸函数的收敛性讨论的较多, 而非凸函数的讨论较少, 而关于非凸函数的无约束优化问题的两步正则化牛顿算法国内外研究还是甚少, 所以本文将对此进行研究并去证明他的有效性, 具有全局收敛性且是具有三阶收敛速度的。

2. 非凸优化问题的两步正则化牛顿法

首先, f 是 $R^n \rightarrow R$ 的非凸函数且二阶连续可微, 将梯度 $\nabla f(x^k)$ 和二阶海瑟矩阵 $\nabla^2 f(x^k)$, 分别用 g_k 和 H_k 表示。用 $\|\cdot\|$ 表示2-范数。

该正则化牛顿法主要方案如下, 在每次迭代中, 求解如下方程:

$$(H_k + \lambda_k I)d = -g_k, \quad (2)$$

得到牛顿步长 d_k , 其中 $c_1 > 1$, λ_k 为合适的正则化参数, 有:

$$\begin{aligned} \lambda_k &= c_1 \psi_k + c_2 \|g_k\|, \\ \psi_k &= \max(0, -v_{\min}(H_k)). \end{aligned}$$

v 为矩阵 H_k 的特征值。

再求解:

$$(H_k + \lambda_k I)d = -g(y_k), \quad y_k = x_k + d_k, \quad (3)$$

得到的近似的牛顿步长 \bar{d}_k 。

设 d_k 和 \bar{d}_k 由(2)和(3)给出, 因为 $H_k + c_1 \psi_k I + \lambda_k I$ 正定, 因此 d_k 是 $f(x)$ 在 x_k 处的下降方向, 但 $d_k + \bar{d}_k$ 可能不是, 下面我们来讨论其是否为下降方向。

我们定义实际减少为:

$$Ared_k = f(x_k) - f(x_k + d_k + \bar{d}_k). \quad (4)$$

称为在第 k 次迭代时 $f(x)$ 的实际减少。

我们记牛顿步长 d_k 为下列问题的极小值,

$$\min_{d \in R^n} \varphi_{k,1}(d) = \frac{1}{2} d^T H_k d + g_k^T d + \frac{1}{2} \lambda_k \|d\|^2. \quad (5)$$

令:

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

由[7]知, d_k 也是如下最值问题的解:

$$\min_{d \in R^n} \frac{1}{2} d^T H_k d + g_k^T d, \quad \text{s.t. } \|d_k\| \leq \Delta_{k,1}. \quad (6)$$

根据[7], 得:

$$\varphi_{k,1}(0) - \varphi_{k,1}(d) \geq \frac{1}{2} \|g_k\| \min \left\{ \|d_k\|, \frac{\|g_k\|}{\|H_k\|} \right\}. \quad (7)$$

仿照 d_k , 根据相似的式子也定义了 \bar{d}_k ,

$$\min_{d \in R^n} \varphi_{k,2}(d) = \frac{1}{2} d^T H_k d + g(y_k)^T d + \frac{1}{2} \lambda_k \|d\|^2.$$

\bar{d}_k 为下列信赖域问题的解,

$$\min_{d \in R^n} \frac{1}{2} d^T H_k d + g(y_k)^T d, \quad \text{s.t. } \|d_k\| \leq \Delta_{k,2}.$$

其中:

$$\Delta_{k,2} = \|\bar{d}_k\| = \left\| -(H_k + \lambda_k I)^{-1} g(y_k) \right\|.$$

我们有,

$$\varphi_{k,2}(0) - \varphi_{k,2}(d) \geq \frac{1}{2} \|g(y_k)\| \min \left\{ \|\bar{d}_k\|, \frac{\|g(y_k)\|}{\|H(y_k)\|} \right\}. \quad (8)$$

我们定义预测减少为:

$$Pred_k = \varphi_{k,1}(0) - \varphi_{k,1}(d) + \varphi_{k,2}(0) - \varphi_{k,2}(d). \quad (9)$$

且:

$$Pred_k \geq \frac{1}{2} \|g_k\| \min \left\{ \|d_k\|, \frac{\|g_k\|}{\|H_k\|} \right\} + \frac{1}{2} \|g(y_k)\| \min \left\{ \|\bar{d}_k\|, \frac{\|g(y_k)\|}{\|H(y_k)\|} \right\} \quad (10)$$

根据定义 $Pred_k$ 总是非负的, 定义比率:

$$r_k = \frac{Ared_k}{Pred_k}. \quad (11)$$

下面详细给出本论文的求解非凸无约束优化问题(1)的两步正则化牛顿法。

Algorithm1

Step1: 给出初始点 $x_1 \in R^n$, $0 \leq \varepsilon \ll 1$ 。

Step2: 若 $\|g_k\| \leq \varepsilon$, 停止运算并输出 x_k 为近似极小点。

Step3: 通过如下方程式求解 d_k ,

$$(H_k + \lambda_k I)d = -g_k,$$

其中:

$$\lambda_k = c_1 \psi_k + c_k \|g_k\|, \quad \psi_k = \max(0, -v_{\min}(H_k))$$

令:

$$y_k = x_k + d_k.$$

通过下列方程式求解 \bar{d}_k

$$(H_k + \lambda_k I)d = -g(y_k).$$

令:

$$s_k = d_k + \bar{d}_k.$$

Step4: 计算 $r_k = \frac{Ared_k}{Pred_k}$, 令:

$$x_{k+1} = \begin{cases} x_k + s_k, & r_k \geq p_0 \\ x_k, & \text{其余情况} \end{cases}$$

Step5: 更新 c_{k+1}

$$c_{k+1} = \begin{cases} 4c_k, & r_k < p_1 \\ c_k, & r_k \in [p_1, p_2] \\ \max\left\{\frac{c_k}{4}, m\right\}, & r_k > p_2 \end{cases}$$

令 $k = k + 1$, 再转到步2。

3. 收敛性分析

在本节中，我们主要研究该正则化牛顿法的全局收敛性，但由于 f 不是凸函数，由就算该正则化牛顿法全局收敛并不意味着该方法找到了全局最优解。并且在一定条件下有着超线性收敛性。

我们提出了如下的假设。

假设1: $g(x)$ 和 $H(x)$ 都是lipschitz连续的，则有 $L > 0$ ，有如下式子成立。

$$\|g(x) - g(y)\| \leq L\|x - y\|. \quad (12)$$

$$\|H(x) - H(y)\| \leq L\|x - y\|. \quad (13)$$

由lipschitz连续，我们有：

$$\|g(y) - g(x) - H(x)(y - x)\| \leq L\|y - x\|^2. \quad (14)$$

假设2:

1) 存在该问题的局部最优解 x^* 。

2) $\|g(x)\|$ 在 x^* 的某个领域上提供了一个局部误差界，即存在两个正常数 c_1 和 b_1 ，使得：

$$\|g(x)\| \geq c_1 \text{dist}(x, X), \quad \forall x \in N(x^*, b_1) = \{x \mid \|x - x^*\| \leq b_1\} \quad (15)$$

3) 海瑟矩阵 $H(x)$ 是局部lipschitz连续的，即存在常数 $b_2 \in (0, 1)$ 和 $L_H > 0$ ，有：

$$\|H(y) - H(x)\| \leq L_H \|y - x\|, \quad \forall x \in N(x^*, b_1). \quad (16)$$

$$\|H(y)(x - y) - (g(x) - g(y))\| \leq \frac{1}{2} L_H \|y - x\|^2, \quad \forall x \in N(x^*, b_1). \quad (17)$$

由于 f 是两次连续可微的，因此存在正常数 L_1, L_2 ：

$$\|g(x)\| \leq L_1, \quad \forall x \in N(x^*, b_1). \quad (18)$$

$$\|g(x) - g(y)\| \leq L_2 \|x - y\|, \quad \forall x \in N(x^*, b_1). \quad (19)$$

在本文的后面部分，我们用 $\bar{x} \in X$ 满足，

$$\|\bar{x} - x\| = \text{dist}(x, X) = \inf_{y \in X} \|y - x\|. \quad (20)$$

定理1: 若假设1成立，如果 f 有界，根据AlgorithmI，我们有该算法终止于有限迭代或者满足 $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$ 。

证明：采用反证法，假设定理为假，则存在一个整数 \bar{k} ，对 $\forall k \geq \bar{k}$ 有：

$$\|g_k\| \geq \tau. \quad (21)$$

在不失一般性的条件下，我们能假设 $\bar{k} = 1$ ，我们令 $T = \{k \mid x_k \neq x_{k+1}\}$ ，我们有：

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

下面我们来考虑如下两种情况：

情况1：若 T 是无限的，我们能找到一个整数 k_1 ，有：

$$x_{k_1} = x_{k_1+1} = x_{k_1+2} = \dots$$

通过上述算法的步4，我们有：

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

由步5和式(21)我们有:

$$c_k \rightarrow \infty, \lambda_k \rightarrow \infty.$$

因为 $x_k = x_{k+1}$, 对任意 $k \geq k_1$ 都成立, 得到:

$$\|d_k\| = \left\| -(H_k + c_1 \psi_k I + \lambda_k I)^{-1} g_k \right\| \leq \lambda_k^{-1} \|g_k\| \rightarrow 0.$$

从步3, 我们得到,

$$\begin{aligned} \|\bar{d}_k\| &= \left\| -(H_k + \lambda_k I)^{-1} g(y_k) \right\| \\ &\leq \left\| (H_k + c_1 \psi_k I + c_k I)^{-1} (g(y_k) - g_k - H_k d_k) \right\| \\ &\quad + \left\| (H_k + c_1 \psi_k I + c_k I)^{-1} g_k \right\| + \left\| (H_k + c_1 \psi_k I + c_k I)^{-1} H_k d_k \right\| \\ &\leq L c_k^{-1} \|d_k\|^2 + 2 \|d_k\| \\ &\leq m \|d_k\|. \end{aligned}$$

得到

$$\begin{aligned} &|Ared_k - Pred_k| \\ &= \left| f(x_k) - f(x_k + d_k + \bar{d}_k) - (\varphi_{k,1}(d) - \varphi_{k,1}(d) + \varphi_{k,2}(d) - \varphi_{k,2}(d)) \right| \\ &\leq \left| f(y_k + \bar{d}_k) - f(y_k) - \frac{1}{2} \bar{d}^T H_k d - g(y_k)^T \bar{d} \right| + \left| f(y_k) - f(x_k) - d^T H_k d + g_k^T d \right| \\ &= o(\|d_k\|^2) + o(\|\bar{d}_k\|^2) = o(\|d_k\|^2). \end{aligned}$$

从(10)和(21)得到:

$$Pred_k \geq \frac{1}{2} \tau \min \left\{ \|d_k\|, \frac{\tau}{L} \right\} \geq \frac{1}{2} \tau \|d_k\|.$$

那么对于足够大的 k , 下式都成立,

$$|r_k - 1| = \frac{|Ared_k - Pred_k|}{Pred_k} = \frac{o(\|d_k\|^2)}{\|d_k\|} \rightarrow 0.$$

这意味着 $r_k \rightarrow 1$, 因此, 根据算法中的步5, 那么有一个正数 b_1 有: $c_k \leq b_1$ 。

情况2: 若 T 是有界的, 那么有,

$$\begin{aligned} \infty &> f(x_1) - \liminf_{k \rightarrow \infty} f(x_k) \\ &\geq \sum_{i=1}^{\infty} f(x_i) - f(x_{i+1}) = \sum_{k \in T} p_0 Pred_k \\ &\geq \sum_{k \in T} p_0 \left(\frac{1}{2} \|g_k\| \min \left\{ \|d_k\|, \frac{\|g_k\|}{L} \right\} + \frac{1}{2} \|g(y_k)\| \min \left\{ \|\bar{d}_k\|, \frac{\|g(y_k)\|}{\|H(y_k)\|} \right\} \right) \\ &\geq \sum_{k \in T} p_0 \frac{\tau}{2} \min \left\{ \|d_k\|, \frac{\tau}{L} \right\}. \end{aligned}$$

得:

$$\lim_{k \rightarrow \infty, k \in T} d_k = 0. \quad (22)$$

由Step5可知,

$$\lambda_k \rightarrow \infty.$$

与情况1类似, 我们有

$$\|\bar{d}_k\| \leq b_2 \|d_k\|, \quad \forall k \in T. \quad (23)$$

根据(23)我们有,

$$\begin{aligned} \|s_k\| &= \|d_k + \bar{d}_k\| \leq (1 + b_2) \|d_k\|, \quad \forall k \in T. \\ \sum_{k \in T} \|s_k\| &< \infty. \end{aligned}$$

这意味着,

$$x_k \rightarrow x^*, \quad d_k \rightarrow 0, \quad \bar{d}_k \rightarrow 0.$$

因为 $(H_k + \mu_k \|g_k\|^\sigma I)d = -g_k$, 由(12) (14) (23), 有:

$$\begin{aligned} c_k \|d_k\| &= \|g_k + H_k d_k + c_1 \psi_k d_k\| \\ &\geq \|g_k\| - \|H_k\| \|d_k\| - \|c_1 \psi_k\| \|d_k\| \\ &\geq \tau - L \|d_k\|. \end{aligned}$$

有:

$$c_k \geq \frac{\tau}{\|d_k\|} - L \rightarrow \infty.$$

由(11)可知, $r_k \rightarrow 1$ 。同理有那么有一个正数 b_2 有: $c_k \leq b_1$ 。

根据这两个情况所得 c_k 均我们的假设相矛盾。定理得证。

根据定理1, 我们可以得到AlgorithmI是全局收敛的, 且若 f 是凸函数, 定理1可以保证AlgorithmI产生的解集 $\{x_k\}$ 收敛于全局最优解。但 f 是非凸函数, 全局收敛并不能保证找到全局最优解, 为了证明该算法的超线性收敛性, 我们给出如下的引理。

引理1: 若假设2成立, 那么我们有

$$\begin{aligned} \|d_k\| &= o(\|\bar{x}_k - x_k\|). \\ \|\bar{d}_k\| &= o(\|\bar{x}_k - x_k\|). \\ \|s_k\| &= o(\|\bar{x}_k - x_k\|). \end{aligned}$$

证明: 因为 $x_k \rightarrow x^* \in X$, 我们有:

$$\begin{aligned} \|d_k\| &= \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} g_k \right\| \\ &= \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} (g_k - \nabla f(\bar{x}_k) - H_k(x_k - \bar{x}_k) + H_k(x_k - \bar{x}_k)) \right\| \\ &\leq \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} (g_k - \nabla f(\bar{x}_k) - H_k(x_k - \bar{x}_k)) \right\| \\ &\quad + \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} H_k(x_k - \bar{x}_k) \right\| \end{aligned}$$

$$\begin{aligned}
 &\leq \left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} \right\| \left\| g_k - \nabla f(\bar{x}_k) - H_k (x_k - \bar{x}_k) \right\| \\
 &\quad + \left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} \right\| \left\| H_k (x_k - \bar{x}_k) \right\| \\
 &\leq \left(\frac{L_H}{2} \|x_k - \bar{x}_k\|^2 + H_k \|x_k - \bar{x}_k\| \right) \left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} \right\| \\
 &= \frac{L_H}{2} \text{dist}(x_k, X)^2 \left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\|^\delta I \right)^{-1} \right\| \\
 &\quad + \text{dist}(x_k, X) \left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\|^\delta I \right)^{-1} H_k \right\|.
 \end{aligned}$$

其中

$$\begin{aligned}
 \left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} &= v_{\max} \left(\left(H_k + c_1 \psi_k I + c_k \|g_k\|^\delta I \right)^{-1} \right) \\
 &= \frac{1}{v_{\min} \left(H_k + c_1 \psi_k I + c_k \|g_k\|^\delta I \right)} \\
 &\leq \frac{1}{c_k \|g_k\|} \leq \frac{1}{c_k L \text{dist}(x_k, X^*)}.
 \end{aligned}$$

下面我们考虑 $\left\| \left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} H_k \right\|$ ，我们称 $v_k^{(l)}$ 为 H_k 中第 l 大的特征值。那么 $\left(H_k + c_1 \psi_k I + c_k \|g_k\| I \right)^{-1} H_k$ 的特征值为

$$\frac{v_k^{(l)}}{v_k^{(l)} + c_1 \psi_k + c_k \|g_k\|}, \quad 1 \leq l \leq n.$$

下面我们分两种情况讨论：1) $v_k^{(l)} \geq 0$ ，2) $v_k^{(l)} < 0$ 。

1) $\frac{|v_k^{(l)}|}{|v_k^{(l)} + c_1 \psi_k + c_k \|g_k\|} \leq 1.$

2) 因为：

$$\begin{aligned}
 -\psi_k = v_{\min}(H_k) &\leq v_k^{(l)} < 0, \\
 v_k^{(l)} - v_{\min}(H_k) &\geq 0, \\
 |v_k^{(l)}| &\leq |v_{\min}(H_k)|.
 \end{aligned}$$

因此，我们有

$$\begin{aligned}
 \frac{|v_k^{(l)}|}{|v_k^{(l)} + c_1 \psi_k + c_k \|g_k\|} &= \frac{|v_k^{(l)}|}{\left| \left(v_k^{(l)} - v_{\min}(H_k) \right) - (c_1 - 1)v_{\min}(H_k) + c_k \|g_k\| \right|} \\
 &\leq \frac{|v_k^{(l)}|}{v_k^{(l)} - v_{\min}(H_k) + (c_1 - 1)|v_{\min}(H_k)| + c_k \|g_k\|} \\
 &\leq \frac{1}{c_1 - 1}.
 \end{aligned}$$

那么我们有

$$\frac{|v_k^{(l)}|}{|v_k^{(l)} + c_1 \psi_k + c_k \|g_k\|} \leq \max\left(1, \frac{1}{c_1 - 1}\right).$$

因此, 我们有

$$\begin{aligned} \|d_k\| &\leq \frac{L_H}{2c_k L} \text{dist}(x_k, X^*) + \max\left(1, \frac{1}{c_1 - 1}\right) \text{dist}(x_k, X^*) \\ &\leq L_1 \text{dist}(x_k, X^*). \end{aligned}$$

同理

$$\begin{aligned} \|\bar{d}_k\| &= \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} g(y_k) \right\| \\ &= \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} (g(y_k) - \nabla f(x_k) - H_k(y_k - x_k) + H_k(y_k - x_k)) \right\| \\ &\leq \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} (g(y_k) - \nabla f(x_k) - H_k(y_k - x_k)) \right\| \\ &\quad + \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} H_k(y_k - x_k) \right\| \\ &\leq \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} \right\| \left\| g(y_k) - \nabla f(x_k) - H_k(y_k - x_k) \right\| \\ &\quad + \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} \right\| \left\| H_k(y_k - x_k) \right\| \\ &\leq \left(\frac{L_H}{2} \|y_k - x_k\|^2 + H_k \|y_k - x_k\| \right) \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} \right\| \\ &= \frac{L_H}{2} \|d_k\|^2 \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} \right\| + \|d_k\| \left\| (H_k + c_1 \psi_k I + c_k \|g_k\| I)^{-1} H_k \right\| \\ &\leq m \text{dist}(x_k, X) + \max\left(1, \frac{1}{c_1 - 1}\right) \text{dist}(x_k, X) \\ &\leq \max\left(1, \frac{1}{c_1 - 1}\right) \text{dist}(x_k, X) \\ \|s_k\| &= \|d_k + \bar{d}_k\| = o(\|\bar{x}_k - x_k\|). \end{aligned}$$

引理成立。

引理2: 若假设2成立, 那么我们总有一个常数 c_2 , 有 $c_k \leq c_2$ 。

证明: 从

$$\begin{aligned} \varphi_{k,1}(0) - \varphi_{k,1}(d_k) &\geq \frac{1}{2} \|g_k\| \min\left\{ \|d_k\|, \frac{\|g_k\|}{\|G_k\|} \right\} \\ &\geq \frac{1}{2} c_1 \|x_k - \bar{x}_k\| \min\left\{ \|d_k\|, \frac{c_1}{L} \|x_k - \bar{x}_k\| \right\} \\ &\geq c_2 \|x_k - \bar{x}_k\|. \end{aligned}$$

对于某个 c_2 来说成立

从(11)我们得到

$$|r_k - 1| = \left| \frac{Ared_k - Pred_k}{Pred_k} \right| = \frac{o(\|d_k\|^2)}{\|x_k - \bar{x}_k\| \min\{\|x_k - \bar{x}_k\|, \|d_k\|\}} \rightarrow 0.$$

因此有 $r_k \rightarrow 1$ ，我们由算法中的步骤5的更新规则推断出，存在一个常数 c_2 ，有 $c_k \leq c_2$ 。引理得证。

引理3: 若假设2成立，我们有

$$\psi_k \leq L_H \text{dist}(x_k, X)$$

证明:

情况1): 当 $H_k \geq 0$ 时，我们有 $\psi_k = 0$ 。

情况2): 当 $H_k < 0$ 时，我们假设 $v_{\min}(H_k) < 0$ ，我们称 $\overline{v_k^{(l)}}$ 为 $\nabla^2 f(\bar{x}_k)$ 中第 l 大的特征值。

因为 $\bar{x}_k \in X$ ，我们有 $\overline{v_k^{(l)}} \geq 0$ ，根据假设，非凸函数 $f(x)$ 满足二阶连续的条件，那么 $\nabla^2 f(\bar{x}_k)$ 是对称矩阵，则存在正交矩阵 Q_k ，有：

$$Q_k^T \nabla^2 f(\bar{x}_k) Q_k = \text{diag}(\overline{v_k^{(l)}}).$$

其中 $\text{diag}(\overline{v_k^{(l)}})$ 表示的为对角线元素为 $\overline{v_k^{(l)}}$ 的对角矩阵，

$$\begin{aligned} v_{\min}(H_k) I - Q_k^T H_k Q_k &= v_{\min}(H_k) I - Q_k^T (\nabla^2 f(\bar{x}_k) + (H_k - \nabla^2 f(\bar{x}_k))) Q_k \\ &= v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) - Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k. \end{aligned}$$

因为 $Q_k^T H_k Q_k$ 总会有特征值 $v_{\min}(H_k)$ ，因此我们有 $v_{\min}(H_k) I - Q_k^T H_k Q_k$ 是奇异的，因此 $v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) - Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k$ 也是奇异的。因为 $v_{\min}(H_k) < 0$ 且 $\overline{v_k^{(l)}} \geq 0$ ， $v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}})$ 是非奇异的。我们令：

$$M = \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} * v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) - Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k.$$

根据引理，我们有：

$$\begin{aligned} 1 &\leq \|I - M\| \\ &= \left\| I - \left(I - \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} \right) Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k \right\| \\ &= \left\| \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k \right\| \\ &\leq \left\| \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} \right\| \left\| Q_k^T (H_k - \nabla^2 f(\bar{x}_k)) Q_k \right\| \\ &= \left\| \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} \right\| \left\| H_k - \nabla^2 f(\bar{x}_k) \right\|. \end{aligned}$$

我们分别考虑 $\left\| \left(v_{\min}(H_k) I - \text{diag}(\overline{v_k^{(l)}}) \right)^{-1} \right\|$ 和 $\left\| H_k - \nabla^2 f(\bar{x}_k) \right\|$ 。

因为 $v_{\min}(H_k) < 0$ ，且 $\overline{v_k^{(l)}} \geq 0$ 。

$$\begin{aligned} \left(v_{\min}(H_k)I - \text{diag}\left(\overline{v_k^{(l)}}\right) \right)^{-1} &= \max_{1 \leq l \leq n} \left| \left(v_{\min}(H_k)I - \text{diag}\left(\overline{v_k^{(l)}}\right) \right)^{-1} \right| \\ &= \frac{1}{\min_{1 \leq l \leq n} \left| v_{\min}(H_k)I - \text{diag}\left(\overline{v_k^{(l)}}\right) \right|} \\ &\leq \frac{1}{|v_{\min}(H_k)|} = \frac{1}{\psi_k}. \end{aligned}$$

根据假设2, 我们有

$$\|H_k - \nabla^2 f(\bar{x}_k)\| \leq L_H \|x_k - \bar{x}_k\| = L_H \text{dist}(x_k, X).$$

因此, 我们有:

$$\psi_k \leq L_H \text{dist}(x_k, X).$$

引理得证。

根据上述引理, 我们有:

$$\begin{aligned} \lambda_k = c_1 \psi_k + c_k \|g_k\| &\geq c_k \|g_k\| \geq m \text{dist}(x_k, X) = mc \|x_k - \bar{x}_k\|, \\ \lambda_k = c_1 \psi_k + c_k \|g_k\| &\leq L_H \text{dist}(x_k, X) + c_k \|g_k\| \leq c \|x_k - \bar{x}_k\| \end{aligned}$$

即: λ_k, c_k 等价于 $\|x_k - \bar{x}_k\|$ 。

$$\lambda_k = \|x_k - \bar{x}_k\|, c_k = \|x_k - \bar{x}_k\|. \quad (24)$$

引理4: 若假设2成立, 我们有

$$\text{dist}(x_{k+1}, X) = o(\text{dist}(x_k, X)^2).$$

证明: 由(18)可知

$$\begin{aligned} k_1 \text{dist}(x_{k+1}, X) &\leq \|g_{k+1}\| = \|g(y_k + \bar{d}_k)\| \\ &\leq \|g(y_k + \bar{d}_k) - g(y_k) - G(y_k)\bar{d}_k\| + \|g(y_k) + G(y_k)\bar{d}_k\| \\ &\leq L\|\bar{d}_k\|^2 + \|g(y_k) + G(x_k)\bar{d}_k\| + \|(G(y_k) - G(x_k))\bar{d}_k\| \\ &\leq L\|\bar{d}_k\|^2 + (\psi_k + \lambda_k)\|\bar{d}_k\| + L\|d_k\|\|\bar{d}_k\| \\ &= o(\text{dist}(x_k, X)^2). \end{aligned}$$

定理2: 若假设成立, 我们假设 $\{x_k\}$ 是由AlgorithmI生成的序列, 那么我们称 $\text{dist}(x_k, X)$ 是二阶收敛于0, 且 $\{x_k\}$ 是局部收敛于最优解 x^* 的。

证明: 由引理4, 我们总能找到一个序列 $\{x_k\}$ 使得 $\text{dist}(x_k, X)$ 2阶趋近于0的。

而根据[1], 我们总能找到序列 $\{x_k\}$, 在其中找出两个元素 $x_p, x_q \in \{x_k\}$, 使得 $|x_p - x_q| \leq \varepsilon$ 。

那么我们可知, $\{x_k\}$ 是一个柯西序列那么其肯定是一个收敛数列, 定理得证。

引理5 [7]: 如果 $\{x_k\}$ 是超线性收敛于 x^* 的, 那么有:

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x^*\|} = 1.$$

由引理4我们有

$$\|\bar{x}_k - x_k\| \leq \|\bar{x}_{k+1} - x_{k+1} + x_{k+1} - x_k\| \leq 2\|s_k\|$$

由引理5可得, $\|s_k\|$ 等价于 $\|x_k - x^*\|$ 。因此有 $\|x_k - x^*\|$ 等价于 $\|\bar{x}_k - x_k\|$ 。

我们知道 $H(x^*) \neq 0$, 又 $H(x^*)$ 是对称的, 而且, $H(x^*) + c_1 \psi_* I$, 是半正定的, 且 $\psi_* \geq 0$, 此时 $G(x^*) = H(x^*) + c_1 \psi_* I \neq 0$, 因此存在正交矩阵 U 使得:

$$G(x^*) = (U_1, U_2) \begin{pmatrix} \Sigma_1^* & 0 \\ 0 & 0 \end{pmatrix} (U_1, U_2)^T,$$

其中 Σ_1^* 是一个正对角矩阵, 类似的, 我们假设 $H(x)$ 的奇异值分解为:

$$G(x) = (U_1, U_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} (U_1, U_2)^T = U_1 \Sigma_1 U_1^T + U_2 \Sigma_2 U_2^T.$$

其中 $\text{Rank}(\Sigma_1) = \text{Rank}(\Sigma_1^*)$, Σ_2 随着 $x \rightarrow x^*$ 收敛于 0 。在后文中, 我们将 $G(x_k)$ 简写成 G_k , 则根据矩阵扰动理论[8] [9], 我们有:

$$\|\Sigma_1 - \Sigma_2\| + \|\Sigma_2\| \leq \|G_k - G(x^*)\| \leq L \|x_k - x^*\|.$$

即: $\|\Sigma_1 - \Sigma_1^*\| \leq L \|\bar{x}_k - x_k\|$, $\|\Sigma_2\| \leq \|\bar{x}_k - x_k\|$,

引理6: 若假设成立, 我们有:

$$\|g(y_k)\| = o(\|x_k - \bar{x}_k\|^2), \|U_2 U_2^T g(y_k)\| = o(\|x_k - \bar{x}_k\|^3).$$

证明: 根据式(24), 我们有 $\|g_k + G_k d_k\| = \lambda_k \|d_k\| = o(\|x_k - \bar{x}_k\|^2)$ 。

相似的, 我们有: $\|g_k + G_k \bar{d}_k\| = \lambda_k \|\bar{d}_k\| = o(\|x_k - \bar{x}_k\|^2)$ 。

因此, 我们有,

$$\|g(y_k)\| = \|g(x_k + d_k) - g_k - G_k d_k\| + \|g_k + G_k d_k\| = o(\|x_k - \bar{x}_k\|^2).$$

且

$$\|\bar{y}_k - y_k\| \leq L \|g(y_k)\| = o(\|x_k - \bar{x}_k\|^2).$$

令 $M_k = U_1 \Sigma_1 U_1^T$, $t_k = -M_k^+ g(y_k)$ 。那么我们易得 t_k 是如下最小二乘算法的解:

$$\min \|g(y_k) + M_k t_k\|.$$

因此:

$$\begin{aligned} \|U_2 U_2^T g(y_k)\| &= \|g(y_k) + M_k t_k\| \leq \|g(y_k) + M_k (\bar{y}_k - y_k)\| \\ &\leq \|g(y_k) - G(y_k)(\bar{y}_k - y_k)\| + \|(G(y_k) - G_k)(\bar{y}_k - y_k)\| + \|(G_k - M_k)(\bar{y}_k - y_k)\| \\ &\leq L \|\bar{y}_k - y_k\|^2 + L \|d_k\| \|\bar{y}_k - y_k\| + \|\Sigma_2\| \|\bar{y}_k - y_k\| \\ &= o(\|x_k - \bar{x}_k\|^3). \end{aligned}$$

引理7: 若假设成立, 我们有: $\|\bar{d}_k\| = o(\|x_k - \bar{x}_k\|^2)$ 。

证明: $\bar{d}_k = -(H_k + \lambda_k I)^{-1} g(y_k) = -(G_k + c_k I)^{-1} g(y_k)$
 $= -U_1 (\Sigma_1 + c_k I)^{-1} U_1^T g(y_k) - U_2 (\Sigma_2 + c_k I)^{-1} U_2^T g(y_k).$

因为 $x_k \rightarrow x^*$ ，令 $\Sigma_1 \rightarrow \Sigma_1^*$ ，那么我们有 Σ_1^{-1} 是一致有界的，即存在一个常数 c_4 ，有： $\|\Sigma_1^{-1}\| \leq c_4$ 。因此我们有：

$$\|\bar{d}_k\| \leq \|\Sigma_1^{-1}\| \|U_1 U_1^T g(y_k)\| + c_k^{-1} \|U_2 U_2^T g(y_k)\| \leq c_4 \|g(y_k)\| + c_k^{-1} \|U_2 U_2^T g(y_k)\| = o(\|x_k - \bar{x}_k\|^2).$$

定理3: 若假设成立，我们假设 $\{x_k\}$ 是由AlgorithmI生成的序列，那么我们称 $dist(x_k, X)$ 是三阶收敛于0，且 $\{x_k\}$ 是局部收敛于最优解 x^* 的。

证明:

$$\begin{aligned} \|\bar{x}_{k+1} - x_{k+1}\| &\leq \|g(x_{k+1})\| = \|g(y_k + \bar{d}_k)\| \\ &\leq \|g(y_k + \bar{d}_k) - g(y_k) - G(y_k)\bar{d}_k\| + \|g(y_k) + G(y_k)\bar{d}_k\| \\ &\leq L\|\bar{d}_k\|^2 + \|(G(y_k) - G_k)\bar{d}_k\| + \|g(y_k) + G_k\bar{d}_k\| \\ &\leq L\|\bar{d}_k\|^2 + L\|d_k\|\|\bar{d}_k\| + \lambda_k\|\bar{d}_k\| \\ &= o(\|x_k - \bar{x}_k\|^2). \end{aligned}$$

定理得证。

4. 数值实验

下面求解两个不等式约束优化问题的数值算例，验证AlgorithmI的有效性，并与[3]进行比较，算法在MATLAB R2022a编程实现，数值实验在windows系统中进行。实验中，我们取终止条件为 $\varepsilon_{stop} \leq 10^{-7}$ ，取初始的正则化参数为， $c_1 = 2$ ， $c_k = 2$ ，比率的取值为 $p_0 = 0.1$ ， $p_1 = 0.2$ ， $p_2 = 0.3$ 。

考虑如下无约束优化问题且统一算例的初始值相同，实验结果见下表，每组表都给出了选取的初值 x_1 ，并在所取初值相等的情况下比较，所需迭代次数 k 、迭代点 x^* 和此时函数的值 $f(x^*)$ 。

例1：考虑如下无约束优化问题

$$\min f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_1^2 x_2^2$$

初始值	算法	迭代次数	x^*	函数值
[7,9]	[3]	7	[0,8.6]	1.5e-14
	AlgorithmI	5	[0,6.1]	1.38e-16
[2,3]	[3]	7	[0,2.7998]	3.9e-15
	AlgorithmI	5	[0,2.806]	6.64e-20

例2：考虑如下无约束优化问题

$$\min f(x) = \sin x_1 + \cos x_2 + x_1 x_2^2$$

初始值	算法	迭代次数	x^*	函数值
[3,3]	[3]	16	[4.712, 6.37e-17]	1.17e-17
	AlgorithmI	6	[9.51, 3.82e-17]	2.11e-17
[7,7]	[3]	15	[4.712, 1.6e-17]	3.12e-21
	AlgorithmI	9	[4.712, 3.82e-15]	2.39e-21

例3: 考虑如下无约束优化问题

$$\min f(x) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 + x_1x_3$$

初始值	算法	迭代次数	x^*	函数值
[1,1,1]	[3]	8	[8.6e-12, 8.6e-12, 8.6e-12]	4.47e-20
	AlgorithmI	4	[3.76e-18, 3.76e-14, 3.76e-14]	8.5e-34
[6,6,6]	[3]	13	[8.95e-11, 8.95e-11, 8.95e-11]	4.8e-18
	AlgorithmI	9	[4.3e-11, 4.3e-11, 4.3e-11]	1.1e-18

从上述例子我们可以看出, 上述例子对于非凸函数的无约束优化问题的最优解是有效的,而在取不同的初始值时, 迭代效果依旧强于单步正则化牛顿法, 于是我们说该算法是可行的。

5. 结论

本文根据非凸无约束优化问题的单步正则化牛顿法进行改进提出了非凸无约束优化问题的两步正则化牛顿法, 并对方法的收敛性进行研究。证明了在一定的条件下该方法是全局收敛的并且在局部误差界下具有三次收敛速度。理论分析证明, 该算法有着良好的性质并证明了对应算法的全局收敛性。根据所提出的方法进行了部分数值实验, 并将运算结果与单步的正则化牛顿法进行了对比, 我们的方法有着更好的性质, 进一步验证了本文算法的有效性。

参考文献

- [1] Yamashita, N. and Fukushima, M. (2001) On the Rate of Convergence of the Levenberg-Marquardt Method. https://doi.org/10.1007/978-3-7091-6217-0_18
- [2] Li, D.H., Fukushima, M., Qi, L., et al. (2004) Regularized Newton Methods for Convex Minimization Problems with Singular Solutions. *Computational Optimization and Applications*, **28**, 131-147. <https://doi.org/10.1023/B:COAP.0000026881.96694.32>
- [3] Ueda, K. and Yamashita, N. (2010) Convergence Properties of the Regularized Newton Method for the Unconstrained Nonconvex Optimization. *Applied Mathematics & Optimization*, **62**, 27-46. <https://doi.org/10.1007/s00245-009-9094-9>
- [4] Homeier, H.H.H. (2003) A Modified Newton Method for Rootfinding with Cubic Convergence. *Journal of Computational and Applied Mathematics*, **157**, 227-230. [https://doi.org/10.1016/S0377-0427\(03\)00391-1](https://doi.org/10.1016/S0377-0427(03)00391-1)
- [5] Homeier, H.H.H. (2004) A Modified Newton Method with Cubic Convergence: The Multivariate Case. *Journal of Computational and Applied Mathematics*, **169**, 161-169. <https://doi.org/10.1016/j.cam.2003.12.041>
- [6] Zhou, W. and Chen, X. (2013) On the Convergence of a Modified Regularized Newton Method for Convex Optimization with Singular Solutions. *Journal of Computational and Applied Mathematics*, **239**, 179-188. <https://doi.org/10.1016/j.cam.2012.09.030>
- [7] Sun, W. and Yuan, Y.X. (2006) Optimization Theory and Methods: Nonlinear Programming. Springer Science & Business Media, Berlin.
- [8] 孙继广. 矩阵扰动分析[M]. 北京: 科学出版社, 2001.
- [9] Bogavski, V.N. and Povzner, A. (1991) Matrix Perturbation Theory. In: *Algebraic Methods in Nonlinear Perturbation Theory*. Springer, New York. <https://doi.org/10.1007/978-1-4612-4438-7>