非光滑优化问题分析

刘金波[1,2]，杭丹[2]，刘万利[1]

（1. 中国矿业大学理学院，江苏 徐州 221008;

2. 徐州空军学院基础部，江苏 徐州 221000）

摘要：牛顿法是求解非线性方程组的经典高阶算法，它主要是解决大型线性方程组（牛顿方程组）的精确解问题。为了解决光滑的无约束优化方法的终止判别条件具有缺陷性的问题，本文首先对非精确牛顿法的线性收敛性条件进行了论述。同时，由于非精确牛顿法的线性收敛条件较为严格，所以，在非精确牛顿法的基础上，本文引出了非精确修正牛顿法, 并对其线性收敛性的条件进行了严格的证明。非精确修正牛顿法是对非精确牛顿法的有益的发展和补充, 对解决某些实际问题很有用处。

关键词：运筹学；非光滑优化问题；收敛性

中图分类号：0221.2

Analysis of Non-smooth Optimization

Liu Jinbo[1,2], Hang Dan[2], Liu Wanli[1]

(1. College of Science,China University of Mining and Technology, JiangSu XuZhou 221008;

2. Department of Basic Course Xuzhou Airforce College, JiangSu XuZhou 221000)

Abstract: Newton method for solving nonlinear equations is a classic high-end algorithm, which is mainly to solve large linear equations (Newton's equations) the exact solutions of the problem. In order to solve the smooth termination of unconstrained optimization methods Criterion has a defect in nature, this paper inexact Newton method for linear convergence conditions are discussed. At the same time, due to inexact Newton method for linear convergence conditions are strict, so in inexact Newton method, the paper raises the inexact Newton method modified, and the linear convergence of the conditions of a strict proof. Inexact Modified Newton method is inexact Newton method for the development of useful and complementary, to be useful to solve some practical problems.

Keywords: Operations research; non-smooth Optimization; convergence

0 引言

牛顿法是求解非线性方程组的经典高阶算法。当 x_k 远离 x^* 时，实际上不必花费庞大的工作量以求解大型线性方程组（牛顿方程组）$F'(x_k)s_k = -F(x_k)$ 的精确解。类似的，$F'(x_k)$ 也可以被某些近似值所替代。本论文将从几个角度来分析并说明这个问题。

1 无约束优化问题

我们考虑无约束优化问题：$\min\limits_{x\in X} f(x)$ （1.1）

其中 $f(x)$ 是定义在 Banach 空间的不可微函数。假定函数 $f(x)$ 符合 Lipschitz 条件，给出以下定理：

定理 1.1:如果 $f(x)$ 在 x^* 处达到局部极大或局部极小，且 $f(x)$ 在 x^* 附近是 Lipschitz 的，则必有 $0 \in \partial f(x^*)$. （1.2）

定理证明：见[1].

定理 1.2:设 $f(x)$ 在 X^* 附近是凸的和 Lipschitz 的,且 $0 \in \partial f(x^*)$,则 x^* 是 $f(x)$ 的局部极小点.

作者简介：刘金波，(1979-)，男，江苏徐州人，助教，硕士在读，研究方向为概率论与数理统计. E-mail: jcbljb2008@sina.com

定理证明：见[2].

由定理我们知 x^* 是问题（1.1）之解则必有（1.2）,因此我们称满足（1.2）之点 x 为问题（1.1）的稳定点.

非光滑优化问题(1.1)的求解,如果利用解可微问题的方法(假定在每个迭代点上 $f(x)$ 都可微)有个很大的难点，那就是算法终止的条件不易给出.我们知道,当 x 充分靠近一连续可微函数 $f(x)$ 的极小点时 $\nabla f(x)$ 一定非常小,所以光滑的无约束优化方法的终止判别条件常常是

$$\nabla f(x) \le \varepsilon \qquad (1.3)$$

但对于非光滑函数 $f(x)$,我们并没有类似的结果.

例如,考虑 $X = R^1$, $f(x) = |x|$,则对任何不是解的 x, $f(x)$ 都可微,有

$$|\partial f(x)| = |\nabla f(x)| = 1 \qquad (1.4)$$

为了解决这个问题，我们需要进一步给出非精确牛顿法的线性收敛性.

2 非精确牛顿法的线性收敛性

当 x_0 充分接近 x^* 和 η_k 满足 $0 \le \eta_k < 1$ 时,非精确牛顿法在 $\|\cdot\|_*$ 意义下是线性收敛的.此收敛性有如下定理:

定理 2.1: 假设 $\eta_k \le \eta_{\max} < \eta < 1$,那么存在 $\varepsilon > 0$,使得若 $\|x_0 - x^*\| \le \varepsilon$,则非精确牛顿法是线性收敛的,亦即

$$\left\|x_{k+1} - x^*\right\|_* \le \eta \left\|x_k - x^*\right\| \qquad (2.1)$$

其中, $\|y\|_* = \left\|F'(x^*)y\right\|$.

定理证明：见[3].

这里指出，因 $F'(x^*)$ 是非奇异的，故

$$\frac{\|y\|}{\left\|F'(x^*)^{-1}\right\|} \le \|y\|_* \le \left\|F'(x^*)\right\|\|y\| \qquad (2.2)$$

利用定理 1，注意到（2.1）式和（2.2）式，就可以得到 $Q-$因子[4]如下：

$$Q_1(x_k, x^*) = \limsup_{k\to\infty} \frac{\left\|x_{k+1} - x*\right\|}{\left\|x_k - x^*\right\|} \le \eta cond\left(F'(x^*)\right) \qquad (2.3)$$

其中 $cond\left(F'(x^*)\right) = \left\|F'(x^*)\right\|\left\|F'(x^*)^{-1}\right\|$.

因此,在 x_0 充分接近 x^* 的情况下,当 $\eta cond\left(F'(x^*)\right) < 1$ 时,得到 $Q_1 < 1$,从而非精确牛顿法至少是线性收敛的.但当 $Q_1 > 1$ 时,非精确牛顿法是次线性收敛的.于是,就有一般的非精确牛顿法的线性收敛性的结论如下:

推论 2.2: 假设 $\eta_k \le \eta_{\max} < \eta < 1$ 和 $\rho = \eta cond\left(F'(x^*)\right) < 1$,

那么存在 $\varepsilon > 0$,使得若 $\|x_0 - x^*\| \le \varepsilon$,则非精确牛顿法是线性收敛的,亦即[5]

$$\left\|x_{k+1}-x^*\right\| \le \rho\eta\left\|x_k - x^*\right\| \tag{2.4}$$

3 非精确修正牛顿法的线性收敛性

当 x_0 充分接近 x^* 和 η_k 满足 $0 \le \eta_k < 1$ 时,非精确修正牛顿法在 $\|\cdot\|_*$ 意义下是线性收敛的. 因为 F 在 x_* 的邻域内是连续可微的,故当 x_k 在 x_* 的某邻域内时,就有

$$F(x_k) - F(x^*) = \int_0^1 F'\left(x^* + t(x_k - x^*)\right)dt\,(x_k - x^*) \tag{3.1}$$

$$F(x_k) - F(x^*) - F'\left(x^*(x_k - x^*)\right) = \int_0^1 F'\left[x^* + t(x_k - x^*) - F'(x^*)\right]dt\,(x_k - x^*) \tag{3.2}$$

$$\frac{1}{2}\left\|F'(x^*)\right\|\left\|x_k - x^*\right\| \le \left\|F(x_k)\right\| \le 2\left\|F'(x^*)\right\|\left\|x_k - x^*\right\| \tag{3.3}$$

定理 3.1 纽曼引理:设 $E \in R^{n\times n}, I \in R^{n\times n}$ 是单位矩阵. $\|\cdot\|$ 是满足 $\|I\| = 1$ 的相容矩阵范数, 如果 $\|E\| < 1$,则 $(I - E)$ 非奇异[6],且,

$$(I-E)^{-1} = \sum_{i=0}^{\infty} E^i \qquad \left\|(I-E)^{-1}\right\| \le \sum_{i=0}^{\infty}\left\|E^i\right\| = \frac{1}{1-\|A\|}. \tag{3.4}$$

定理 3.2:假设 $\eta_k \le \eta_{\max} < \eta < 1$,那么存在 $\varepsilon > 0$,使得若 $\left\|x_0 - x^*\right\| \le \varepsilon$,则非精确修正牛顿法是线性收敛的,亦即

$$\left\|x_{k+1}-x^*\right\|_3 \le \eta\left\|x_k - x^*\right\|_3 \tag{3.5}$$

其中, $\|y\|_* = \left\|F'(x^*)y\right\|$.

证明:因为 $F'(x^*)$ 是非奇异的,并且 $\eta_{\max} < \eta$,所以存在 $0 < \gamma < \left\|F'(x^*)^{-1}\right\|^{-1}$ 足够小,使得

$$\frac{\eta_{\max} + \left\|F'(x^*)^{-1}\right\|(\eta_{\max}+2)\gamma}{1 - \left\|F'(x^*)^{-1}\right\|\gamma} \le \eta$$

又因为 F 在 x^* 的邻域内是连续可微的,所以存在 $\delta > 0$ 使得,当 $\left\|y - x^*\right\| \le cond\left(F'(x^*)\right)\delta$ 时,

$$\left\|F'(y) - F'(x^*)\right\| \le \gamma \tag{3.6}$$

又注意到,当 $\left\|x_0 - x^*\right\| \le cond\left(F'(x^*)\right)\delta$ 时,

$$\left\|F'(x_*) - F'(x^0)F'(x^*)^{-1}\right\| \le \left\|F'(x^*)^{-1}\right\|\gamma < 1$$

由(3.4)式就有
$$\left\|\left[I - \left(F'(x_*) - F'(x^0)\right)F'(x^*)^{-1}\right]^{-1}\right\| \le \frac{1}{1 - \left\|F'(x^*)^{-1}\right\|\gamma} \tag{3.7}$$

设 $\left\|x_0 - x^*\right\| \le \delta$. 下面由数学归纳法证明(3.5)式.

由归纳法假设,有
$$\left\|x_0 - x^*\right\| \le \left\|F'(x^*)^{-1}\right\|\left\|x_k - x^*\right\| \le cond\left(F'(x^*)\right)\delta \tag{3.8}$$

故当 $y = x_k$ 时,（3.6）式成立.由（3.2）式和（3.6）式,有

$$\begin{aligned}\|F(x_k)\| &\le \|F'(x^*)\|\|x_k - x^*\| + \int_0^1 [F'(x^* + t(x_k - x^*)) - F'(x^*)]dt(x_k - x^*) \\ &\le \|x_k - x^*\|_* + \gamma\|x_k - x^*\|\end{aligned} \tag{3.9}$$

又由（1.4）式和（3.1）式,有

$$\begin{aligned}x_{k+1} - x^* &= F'(x_0)^{-1}[F'(x_0)(x_k - x^*) - (F(x_k) - F(x_*)) + \gamma_k] \\ &= [F'(x_*) - (F'(x_*) - F'(x_0))]^{-1} \cdot \{[F'(x_0) - F'(x_*) - \\ &\quad (F'(x_*) + t(x_k - x^*)) - F'(x_*)]dt \cdot (x_k - x^*) + \gamma_k\}\end{aligned} \tag{3.10}$$

将(3.10）式左乘 $F'(x_*)$,并且取模,有

$$\begin{aligned}\|x_{k+1} - x^*\|_* \le \left\|\left[I - (F'(x^*) - F'(x_0))F'(x_*)^{-1}\right]^{-1}\right\| \cdot \{\textstyle\int_0^1 [F'(x_0) - F'(x_*) + \\ \|(F'(x_*) + t(x_k - x^*)) - F'(x_*)]dt\| \cdot (x_k - x^*) + \|\gamma_k\|\}\end{aligned} \tag{3.11}$$

由（3.6）,（3.10）,（1.4）,（3.9）式和（2.2）式,得到

$$\begin{aligned}\|x_{k+1} - x^*\|_* &\le \frac{1}{1 - \|F'(x^*)^{-1}\|\gamma \cdot (2\gamma\|x_k - x^*\| + \eta_k\|F(x_k)\|)} \\ &\le \frac{\eta_{\max} + \|F'(x^*)^{-1}\|(\eta_{\max} + 2)\gamma}{1 - \|F'(x^*)^{-1}\|\gamma} \cdot (x_k - x^*)_*\end{aligned} \tag{3.12}$$

最后,由 γ 的任意性得到（3.5）式.

推论 3.3: 假设 $\eta_k \le \eta_{\max} < \eta < 1$ 和 $\rho = \eta cond(F'(x^*)) < 1$,那么存在 $\varepsilon > 0$,使得若 $\|x_0 - x^*\| \le \varepsilon$,则非精确修正牛顿法是线性收敛的,亦即

$$\|x_{k+1} - x^*\| \le \rho\eta\|x_k - x^*\| \tag{3.13}$$

证明:利用定理 2.1,注意到（2.2）式（3.5）式,以及 $\eta_k \le \eta_{\max} < \eta < cond(F'(x^*))^{-1}$,就得到 Q-因子如下:

$$Q_1(x_k, x^*) = \limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \le cond(F'(x^*)) < 1 \tag{3.14}$$

由 $Q_1 < 1$ 知,非精确修正牛顿法是线性收敛的.

4 结论

本文给出了非精确修正牛顿法收敛的判别条件,它是对非精确牛顿法的有益的发展和补充,对解决某些实际问题很有用处.

[参考文献] (References)

[1] Ortega J M,Rheinboldt W G. Iterative Solution of Nonlinear Equations in Several Variables,New York:Academic Press[M].1970, 21(1): 156～187.
[2] 袁亚湘,孙文瑜.最优化理论与方法 [M]. 北京：科学出版社，2003.

[3] 杨凤红，唐云，何淼.大型稀疏非线性方程组的不精确牛顿法[J].延边大学学报，2003
[4] D.R.Han and B.S. He, A new accuracy criterion for approximate proximal point algorithms,Journal of Mathematical Analysis and Applications [M]263(2001):343～354.
[5] K.G.Murty,Linear Complementarity,Linear and Nonlinear Programming,Helderman,Berlin,1988.
[6] L.Mathiesen,Computational experience in solving equilibarum models by a sequence of linear complementarity problems,Operations Research33(1985):144～162.