

A Three-Term Conjugate Gradient Algorithm for Nonlinear Equations Problems

Linghua Huang

School of Information and Statistics, Guangxi University of Finance and Economics, Nanning Guangxi
Email: linghuahuang@163.com

Received: Feb. 6th, 2017; accepted: Feb. 21st, 2017; published: Feb. 24th, 2017

Abstract

This paper designs a three-term conjugate gradient algorithm for nonlinear equations problems and the proposed algorithm has four advantages: (1) the sufficient descent property is satisfied for the search direction; (2) the trust region feature holds for the direction too; (3) the global convergence of the proposed algorithm is possessed; (4) the new algorithm can successfully solve nonlinear equations problems with 1000 dimensions.

Keywords

Nonlinear Equations Problems, Three-Term Conjugate Gradient, Global Convergence

非线性方程组问题的一个三项共轭梯度算法

黄玲花

广西财经学院信息与统计学院, 广西 南宁
Email: linghuahuang@163.com

收稿日期: 2017年2月6日; 录用日期: 2017年2月21日; 发布日期: 2017年2月24日

摘要

本文设计一个三项共轭梯度算法, 用来求解非线性方程组问题, 建议的算法优点是: (1) 搜索方向的充分下降性; (2) 搜索方向的信赖域特点; (3) 算法能保证具有全局收敛性; (4) 算法能成功求解一千维以上的非线性方程组问题。

关键词

方程组问题, 三项共轭梯度, 全局收敛性

Copyright © 2017 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

本文研究下面的非线性方程组问题:

$$q(x) = 0, \quad x \in \mathfrak{R}^n, \quad (1.1)$$

$q: \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ 是连续可微函数, 不难看出问题(1.1)的求解等价于求解如下形式的最小值优化问题模型:

$$\min E(x), \quad x \in \mathfrak{R}^n, \quad (1.2)$$

其中 $E(x) = \frac{1}{2} \|q(x)\|^2$, $\|\cdot\|$ 表示欧式范数。则非线性方程组问题(1.1)就转化为一个最小值的优化问题(1.2), 我们将对(1.2)提出其求解的新算法。求解一个最优化问题, 常采用数值迭代方法, 一般的迭代公式定义为

$$x_{k+1} = x_k + \alpha_k d_k, \quad (1.3)$$

d_k 是搜索方向, α_k 表示步长, x_k 代表第 k 次迭代点(或称为当前点), 有了 d_k 和 α_k , 产生第 $k+1$ 个迭代点 x_{k+1} (或称为下一个迭代点); 依次类推, 根据(1.3)产生一个迭代序列 $\{x_k\}$, 设计算法主要是分析迭代序列 $\{x_k\}$ 是否收敛到某一聚点 x^* 或得到 $E(x_k) \rightarrow 0$?

本文将针对(1.2)设计一个新的三项共轭梯度法, 众所周知共轭梯度算法是一类非常有效的求解目标函数最小值的优化方法[1] [2] [3] [4] [5], 特别是对大规模优化问题, 其搜索方向定义形式是:

$$d_{k+1} = \begin{cases} -q(x_{k+1}), & k=0 \\ -q(x_{k+1}) + \beta_k d_k & k \geq 1 \end{cases}$$

$\beta_k \in \mathbb{R}$ 表示共轭梯度参数, 它是共轭梯度法设计的关键。鉴于此方法设计简单、程序容易实现, 很多学者便成功将此类方法应用于(1.1)的求解(见文献[6] [7] [8]等), 受此启发, 我们将做进一步的研究, 设计一个三项共轭梯度算法, 新方法具有不错的性质和特点。第二部分, 给出设计好的三项共轭梯度方法和及其具体的算法步骤, 第三部分分析算法的下降性、信赖域特点和全局收敛性, 第四部分进行非线性方程组问题的检验。

2. 公式和算法

建议的三项共轭梯度公式如下:

$$d_{k+1} = \begin{cases} -q_{k+1}, & k=0 \\ -q_{k+1} + \frac{q_{k+1}^T y_k d_k - q_{k+1}^T d_k y_k}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|}, & k \geq 1 \end{cases} \quad (2.1)$$

$q_{k+1} = q(x_{k+1})$, $q_k = q(x_k)$, $y_k = q_{k+1} - q_k$, $\delta_1 > 0$, $\delta_2 > 0$, $\delta_3 > 0$ 是常数。上述公式的思想是受文章[7]

[8]启发设计的。下面给出该三项共轭梯度算法详细步骤:

算法 1: (三项共轭梯度算法)

步骤 0: 初始条件的赋值, 初始迭代点 $x_0 \in \mathfrak{R}^n$, 常数 $\delta > 0$, $\delta_1 > 0$, $\delta_2 > 0$, $\delta_3 > 0$, $\alpha \in (0, 1)$, 置 $k := 1$ 。

步骤 1: 条件 $\|q_k\| < \alpha$ 满足, 停止算法; 否则转步骤 2。

步骤 2: 解(2.1)获得搜索方向 d_k 。

步骤 3: 通过下面线搜索技术

$$-q(x_k + \alpha_k d_k)^T d_k \geq \delta \alpha_k \|q(x_k + \alpha_k d_k)\| \|d_k\|^2, \quad (2.2)$$

获得步长 α_k 。

步骤 4: 置 $t_k = x_k + \alpha_k d_k$ 。

步骤 5: 如果条件 $\|q_{k+1}\| \leq \alpha$ 满足, 算法停止并取 $x_{k+1} = t_k$; 否则置

$$x_{k+1} = x_k - \frac{q(t_k)^T (x_k - t_k)}{\|q(t_k)\|^2} q(t_k) \quad (2.3)$$

步骤 6: 置 $k := k + 1$, 转到步骤 1。

注: 线搜索技术(2.2)是文献[7]最先给出, (2.3)是投影技术, 由文献[9]给出。

3. 信赖域特点、充分下降性以及全局收敛性

我们首先通过下面的引理证明算法 1 的信赖域的特点和充分下降性。

引理 3.1: (2.1)产生的搜索方向 d_k 满足

$$q_{k+1}^T d_{k+1} = -\|q_{k+1}\|^2 \quad (3.1)$$

和

$$\beta_1 \|q_k\| \leq \|d_k\| \leq \beta_2 \|q_k\| \quad (3.2)$$

$\beta_2 \geq \beta_1 > 0$ 是常数。

证明: 根据(2.1)的定义形式, $k = 1$ 时, 不难获得(3.1)和(3.2)。下面讨论 $k > 1$ 时的情况, 同样利用(2.1)有

$$\begin{aligned} q_{k+1}^T d_{k+1} &= q_{k+1}^T \left(-q_{k+1} + \frac{q_{k+1}^T y_k d_k - q_{k+1}^T d_k y_k}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \right) = -\|q_{k+1}\|^2 + \frac{q_{k+1}^T y_k q_{k+1}^T d_k - q_{k+1}^T d_k q_{k+1}^T y_k}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \\ &= -\|q_{k+1}\|^2, \end{aligned}$$

从而获得(3.1)。根据(3.1)的结果, 易得

$$-\|q_{k+1}\|^2 = q_{k+1}^T d_{k+1} \geq -\|q_{k+1}\| \|d_{k+1}\|,$$

取 $\beta_1 \in (0, 1)$, 则(3.2)式的左边成立。关于(3.2)的右边不等式, 需要再次利用(2.1)来获得, 具体如下

$$\begin{aligned} \|d_{k+1}\| &= \left\| -q_{k+1} + \frac{q_{k+1}^T y_k d_k - q_{k+1}^T d_k y_k}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \right\| \leq \| -q_{k+1} \| + \left\| \frac{q_{k+1}^T y_k d_k - q_{k+1}^T d_k y_k}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \right\| \\ &\leq \|q_{k+1}\| + \frac{\|q_{k+1}\| \|y_k\| \|d_k\| + \|q_{k+1}\| \|y_k\| \|d_k\|}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \\ &= \|q_{k+1}\| + \frac{2\|q_{k+1}\| \|y_k\| \|d_k\|}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \leq \left(1 + \frac{2}{\delta_1} \right) \|F_{k+1}\|, \end{aligned}$$

最后一个不等式利用了 $\frac{1}{\delta_1 \|d_k\| \|y_k\| + \delta_2 \|q_k\|^2 + \delta_3 |d_k^T q_k|} \leq \frac{1}{\delta_1 \|d_k\| \|y_k\|}$ 获得, 任取 $\beta_2 \in \left(1 + \frac{2}{\delta_1}, +\infty\right)$, (3.2)

的右边不等式便能获得。综上所述, 关系式(3.1)和(3.2)都是成立的。证毕!

关系式(3.1)被称为充分下降性, 而关系式(3.2)被称为信赖域的特征。从证明过程可以看出, 我们建议的三项共轭梯度的搜索方向在没有其他假设条件的情况下, 自身就能拥有这两个优点, 关系式(3.1)和(3.2)对算法的收敛性起着至关重要的作用。为进一步获得三项共轭梯度算法的全局收敛性, 需要一些假设条件。

假设 A: (i) 假设问题(1.1)的解存在;

(ii) $q(x)$ 满足 Lipschitz 条件, 也就是

$$\|q(x) - q(y)\| \leq L \|x - y\|, \quad \forall x, y \in \mathfrak{R}^n,$$

$L > 0$ 是 Lipschitz 常数。

根据假设 A 和获得的引理 3.1, 便可以证明算法 1 的全局收敛性结论。此结论的证明与参考文献[7]的证明过程相似, 我们就不再详述。

定理 3.1: 若迭代序列 $\{x_k, d_k, \alpha_k, q_k\}$ 是由算法 1 产生的且假设 A 的条件满足, 则关系式

$$\liminf_{k \rightarrow +\infty} \|q_k\| = 0$$

成立。

此定理被称为算法 1 的全局收敛性定理。

4. 数值实验

为验证建议的三项共轭梯度算法是有效的, 本节将对下面的问题(见表 1)利用计算机编写程序计算。

我们将利用 Matlab R2009a 编写程序代码, 程序是在 Windows 7 系统、CPU 为 Intel (R) Xeon (R), E5507 2.27 GHz、64 位操作系统、6.00GB 内存上运行。算法中的参数选取分别是: $\delta = 0.002$, $\delta_1 = 0.02$, $\delta_2 = \delta_3 = 0.6$, $\alpha = 10^{-4}$ 。表 2 中第一行参数含义分别是: NI 代表迭代次数, Nq 代表函数值次数, Dim 代表所检验问题的维数, Cpu-Time 代表程序在计算机上运行的 CPU 时间, Val 代表程序终止时的范数值。

Table 1. Problems name

表 1. 问题名称

Nr.	名称
1	Exponential function 2
2	Trigonometric function
3	Logarithmic function
4	Broyden Tridiagonal function
5	Trigexp function
6	Strictly convex function 1
7	Variable dimensioned function
8	Extended Freudentein and Roth function
9	Discrete boundary value problem
10	Troesch problem

Table 2. Numerical results
表 2. 数值结果

Nr.	Dim	NI/Nq	Cpu-Time	Val
1	800	57/206	0.218401	9.958541e-005
	1500	24/93	0.249602	8.792103e-005
2	800	35/70	0.124801	9.849371e-005
	1500	34/68	0.249602	8.990512e-005
3	800	4/5	0.000000	3.878205e-005
	1500	4/5	0.046800	3.246054e-005
4	800	76/152	0.202801	8.833125e-005
	1500	78/156	0.468003	8.953038e-005
5	800	229/594	0.748805	9.804039e-005
	1500	227/590	1.778411	9.532705e-005
6	800	50/100	0.156001	9.109563e-005
	1500	52/104	0.343202	8.173231e-005
7	800	1/2	0.000000	0.000000e+000
	1500	1/2	0.031200	0.000000e+000
8	800	616/1693	1.934412	4.555189e-005
	1500	714/1957	5.444435	1.939103e-005
9	800	28/57	0.093601	8.779266e-005
	1500	26/53	0.202801	9.421179e-005
10	800	0/1	0.000000	0.000000e+000
	1500	0/1	0.000000	0.000000e+000

从表格中测验结果能够看出：(1) 给定的三项共轭梯度算法能成功求解这 10 个非线性方程组问题，范数值接近于零说明利用新算法几乎得到了其精确解，其中第十个问题就获得了精确解；(2) 随着检验问题规模的增大，NI 和 Nq 的变化不大，有的个别问题反而减少，但 Cpu-Time 均会适当增加，这也证明了算法的合理性；(3) 有些问题的 Cpu-Time = 0，这证明算法在运行时，几乎没有占用系统的 CPU 时间。

参考文献 (References)

- [1] Andrei, N. (2008) Another Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization. *Numerical Algorithms*, **47**, 143-156. <https://doi.org/10.1007/s11075-007-9152-9>
- [2] Fletcher, R., Reeves, C.M., et al. (1964) Function Minimization by Conjugate Gradients. *Computer Journal*, **7**, 149-154. <https://doi.org/10.1093/comjnl/7.2.149>
- [3] Polak, E. and Ribière, G. (1968) Note sur la convergence de méthodes de directions conjuguées. *Rev. Francaise Informat Recherche Opertionelle*, **16**, 35-43.
- [4] Wei, Z., Yao, S. and Liu, L. (2006) The Convergence Properties of Some New Conjugate Gradient Methods. *Applied Mathematics & Computation*, **183**, 1341-1350. <https://doi.org/10.1016/j.amc.2006.05.150>
- [5] Yuan, G. (2009) Modified Nonlinear Conjugate Gradient Methods with Sufficient Descent Property for Large-Scale Optimization Problems. *Optimization Letters*, **3**, 11-21. <https://doi.org/10.1007/s11590-008-0086-5>
- [6] Li, Q. and Li, D.H. (2011) A Class of Derivative-Free Methods for Large-Scale Nonlinear Monotone Equations. *IMA Journal of Numerical Analysis*, **31**, 1625-1635. <https://doi.org/10.1093/imanum/drq015>

-
- [7] Li, D. and Wang, L. (2011) A Modified Fletcher-Reeves-Type Derivative-Free Method for Symmetric Nonlinear Equations. *Numerical Algebra, Control and Optimization*, **1**, 71-82. <https://doi.org/10.3934/naco.2011.1.71>
- [8] Yuan, G. and Zhang, M. (2015) A Three-Terms Polak-Ribière-Polyak Conjugate Gradient Algorithm for Large-Scale Nonlinear Equations. *Journal of Computational & Applied Mathematics*, **286**, 186-195. <https://doi.org/10.1016/j.cam.2015.03.014>
- [9] Solodov, M.V. and Svaiter, B.F. (1998) A Globally Convergent Inexact Newton Method for Systems of Monotone Equations. *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Springer, US, 1411-1414.

期刊投稿者将享受如下服务:

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: orf@hanspub.org