Advances in Applied Mathematics 应用数学进展, 2016, 5(4), 813-817
Published Online November 2016 in Hans. http://www.hanspub.org/journal/aam
http://dx.doi.org/10.12677/aam.2016.54093

A LS Algorithm for Nonlinear Equations

Linghua Huang

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

Received: Nov. 12th, 2016; accepted: Nov. 26th, 2016; published: Nov. 30th, 2016

Abstract

This paper presents a LS conjugate gradient algorithm for nonlinear equations and the given algorithm has the following features: 1) the search direction satisfies the sufficient descent property; 2) the direction also has the trust region property; 3) the proposed algorithm possesses the global convergence; 4) numerical results show that the new algorithm is effective.

Keywords

Nonlinear Equations, Conjugate Gradient, Convergence

一个求解非线性方程组问题的LS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：2016年11月12日；录用日期：2016年11月26日；发布日期：2016年11月30日

摘 要

本文给出一个求解非线性方程组问题的LS算法，该方法具有如下特点：1) 搜索方向自动满足充分下降性；2) 方向具有信赖域的特征；3) 算法拥有全局收敛性；4) 数值结果表明新方法是有效的。

文章引用：黄玲花. 一个求解非线性方程组问题的 LS 算法[J]. 应用数学进展, 2016, 5(4): 813-817.
http://dx.doi.org/10.12677/aam.2016.54093

关键词

方程组，共轭梯度，收敛性

1. 引言

考虑如下问题：

$$F(x)=0, x\in\Re^n, \tag{1.1}$$

其中 $F:\Re^n\to\Re^n$ 是一个连续函数，上述问题是一个非线性方程组，一个求解该问题的一个方法时将它化为如下的最优化问题模型：

$$\min P(x), x\in\Re^n, \tag{1.2}$$

其中定义 $P(x)=\frac{1}{2}\|F(x)\|^2$，$\|\bullet\|$ 是欧式范数，不难看出(1.1)和(1.2)的解是等价的。从中可见(1.1)就转化为了一个最小值的优化问题(1.2)，本文将针对(1.2)提出其求解方法。下面的数值迭代方法经常被使用，公式为：

$$x_{k+1}=x_k+\alpha_k d_k, \tag{1.3}$$

其中 d_k 代表搜索方向，α_k 是步长，x_k 是当前迭代点或第 k 次迭代点，根据 d_k 和 α_k 就产生下一个迭代点 x_{k+1}，依次下去便会产生一个点列 $\{x_k\}$，分析此点列收敛到某一聚点 x^* 或得到 $P(x_k)\to 0$。

共轭梯度算法是一类非常有效的数值优化方法，它的方向定义如下：

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

根据 $\beta_k\in R$ 的选取不同，便称为相应的共轭梯度法，目前已有很多公式[1] [2] [3] [4]等。目前已有很多学者将共轭梯度法应用于非线性方程组问题(见文献[5] [6]等)，我们给出一个修正的 LS 共轭梯度算法，该方法具有显著的优点。下一部分，我们给出修正的 LS 方法和具体算法步骤，第三部分证明算法的全局收敛性，最后给出数值检验结果。

2. 公式和算法

下面给出修正的 LS 公式，表达式为：

$$d_{k+1}=\begin{cases}-F_{k+1}, & \text{if} \quad k=0\\ -F_{k+1}+\dfrac{F_{k+1}^T y_k d_k-F_{k+1}^T d_k y_k}{\max\left(2\mu_1\|d_k\|\|y_k\|,\mu_2\left|d_k^T F_k\right|\right)} & \text{if} \quad k\geq 1\end{cases} \tag{2.1}$$

其中 $F_{k+1}=F(x_{k+1}), F_k=F(x_k), y_k=F_{k+1}-F_k$，$\mu_1>0,\mu_2>0$ 是常数。此公式设计的部分思想是基于文章[6]。在此公式的基础上，我们给出一个修正的 LS 共轭梯度算法，步骤如下：

算法 1 (修正的 LS 算法)

步 0：给定初始点 $x_0\in\Re^n$，常数 $\mu>0,\mu_1>0,\mu_2>0,\varepsilon\in(0,1)$。令 $k:=1$。

步 1：若 $\|F_k\|<\varepsilon$，则停止；否则进行步 2。

步 2：通过式(2.1)计算搜索方向 d_k。

步 3：选出满足下面的线搜索条件的步长 α_k：

$$-F\left(x_k+\alpha_k d_k\right)^{\mathrm{T}} d_k \ge \mu\alpha_k \left\|F\left(x_k+\alpha_k d_k\right)\right\| \left\|d_k\right\|^2, \tag{2.2}$$

步 4：令迭代公式为 $p_k = x_k + \alpha_k d_k$ 。

步 5：若 $\|F_{k+1}\| \le \varepsilon$ ，停止并令 $x_{k+1} = p_k$ ；否则令

$$x_{k+1} = x_k - \frac{F\left(p_k\right)^{\mathrm{T}}\left(x_k - p_k\right)}{\left\|F\left(p_k\right)\right\|^2} F\left(p_k\right) \tag{2.3}$$

步：令 $k := k+1$ ，转步 1。

注：线搜索(2.2)是由文献[5]给出，(2.3)的思想来源于文献[7]。

3. 下降性、信赖域性质和收敛性分析

首先分析算法 1 的充分下降性和信赖域的性质，通过下面的引理实现。

引理 3.1. 若搜索方向 d_k 由式(2.1)产生，有

$$F_{k+1}^T d_{k+1} = -\|F_{k+1}\|^2 \tag{3.1}$$

和

$$c_1\|F_k\| \le \|d_k\| \le c_2\|F_k\| \tag{3.2}$$

成立，其中 $c_2 \ge c_1 > 0$ 是常数。

证明：根据(2.1)式，当 $k=1$ 时，(3.1)和(3.2)显然成立。当 $k>1$ 时，由(2.1)得到

$$\begin{aligned} F_{k+1}^T d_{k+1} &= F_{k+1}^T\left(-F_{k+1} + \frac{F_{k+1}^T y_k d_k - F_{k+1}^T d_k y_k}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)}\right) \\ &= -\|F_{k+1}\|^2 + \frac{F_{k+1}^T y_k F_{k+1}^T d_k - F_{k+1}^T d_k F_{k+1}^T y_k}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)} \\ &= -\|F_{k+1}\|^2, \end{aligned}$$

(3.1)式成立。下证(3.2)式。根据(3.1)式，可推出

$$-\|F_{k+1}\|^2 = F_{k+1}^T d_{k+1} \ge -\|F_{k+1}\|\|d_{k+1}\|,$$

从上式可得(3.1)式的左边，其中取 $c_1 = 1$ 。对于(3.2)的右边，通过(2.1)式获得下式

$$\begin{aligned} \|d_{k+1}\| &= \left\|-F_{k+1} + \frac{F_{k+1}^T y_k d_k - F_{k+1}^T d_k y_k}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)}\right\| \\ &\le \|-F_{k+1}\| + \left\|\frac{F_{k+1}^T y_k d_k - F_{k+1}^T d_k y_k}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)}\right\| \\ &\le \|F_{k+1}\| + \frac{\|F_{k+1}\|\|y_k\|\|d_k\| + \|F_{k+1}\|\|y_k\|\|d_k\|}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)} \\ &= \|F_{k+1}\| + \frac{2\|F_{k+1}\|\|y_k\|\|d_k\|}{\max\left(2\mu_1\|d_k\|\|y_k\|, \mu_2\left|d_k^T F_k\right|\right)} \\ &\le \left(1 + \frac{1}{\mu_1}\right)\|F_{k+1}\|, \end{aligned}$$

其中最后一个不等式根据 $\frac{1}{\max\left(2\mu_1\|d_k\|\|y_k\|,\mu_2\left|d_k^T F_k\right|\right)}\le\frac{1}{2\mu_1\|d_k\|\|y_k\|}$ 得到，取 $c_2=\left(1+\frac{1}{\mu_1}\right)$ 便得到(3.2)的右边。综上，(3.1)和(3.2)均成立。证毕！

上述引理中的(3.1)是充分下降性，(3.2)是信赖域性质，这两个性质对算法的全局收敛性起着非常重要的作用。从中也可看出，所建议的搜索方向不需要任何附加条件，自动满足这两个性质。为获得算法 1 的全局收敛性，我们需要下面的假设条件。

假设 1：(A) 所研究的问题(1.1)有解；

(B) $F(x)$ 在 $\Re^n$ 上满足下面的 Lipschitz 性质，即

$$\|F(x)-F(y)\|\le L\|x-y\|,\quad \forall x,y\in\Re^n,$$

其中 $L>0$ 是个常数。

在假设 1 和引理 3.1 结论的基础上，我们可以建立算法的全局收敛性。因为全局收敛性的证明与文献[6]的证明过程基本相同，此处不再详述，只给出定理。

定理 3.1. 假如序列 $\{x_k,d_k,\alpha_k,F_k\}$ 由算法 1 产生，同时假设 1 成立，则有

$$\lim_{k\to+\infty}\inf\|F_k\|=0$$

成立。

此定理被称为算法 1 的全局收敛性定理。

4. 数值结果

为了验证算法的有效性，我们将对下面的问题进行检验(表 1)，这些问题同文献[5]一样。

利用 Matlab R2009a 编写程序，在 Windows 7 系统，6.00GB 内存，64 位操作系统，CPU 为 Intel (R) Xeon (R), E5507 2.27 GHz 上运行。参数选取是：$\mu=0.001,\mu_1=0.01,\mu_2=0.5,\varepsilon=1e-4$。表二中参数含义：NI 表示迭代次数，NF 表示函数值次数，Dim 表示问题的维数，Time 表示运行的 CPU 时间，Fval 表示程序终止时的范数值。

Table 1. Equations problem
表 1. 方程组问题

No.	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	Extended Freudentein and Roth function
8	Discrete boundary value problem
9	Troesch problem

No.	Dim	NI/NF	Time	Fval
1	1000 2000	32/120 17/68	0.202801 0.312002	7.996961e−005 9.097385e−005
2	1000 2000	35/70 33/66	0.156001 0.405603	8.879691e−005 9.802225e−005
3	1000 2000	4/5 4/5	0.031200 0.031200	3.596900e−005 3.102323e−005
4	1000 2000	131/263 134/269	0.436803 1.372809	9.821319e−005 9.637759e−005
5	1000 2000	111/318 113/318	0.561604 1.591210	9.079656e−005 7.925594e−005
6	1000 2000	51/102 52/104	0.202801 0.577204	8.338954e−005 9.544534e−005
7	1000 2000	1/2 1/2	0.000000 0.031200	0.000000e+000 0.000000e+000
8	1000 2000	27/55 26/52	0.124801 0.312002	9.504233e−005 9.885319e−005
9	1000 2000	0/1 0/1	0.000000 0.000000	0.000000 0.000000

从上面表格中的数值结果可以看出，算法 1 对上述问题的求解非常有效，最终的范数值很接近零，表明对非线性方程组的精确解逼近程度很高，个别问题可以求到精确解(如问题 7 和问题 9)；随着问题维数的增加，CPU 时间和迭代次数以及函数值次数改变不是很大，说明算法 1 的有效性；个别问题的 CUP 时间是零，说明其求解几乎没有占用系统的 CPU 时间，就得出了结果。

参考文献 (References)

[1] Fletcher, R. and Reeves, C.M. (1964) Function Minimization by Conjugate Gradients. *Computer Journal*, **7**, 149-154. https://doi.org/10.1093/comjnl/7.2.149

[2] Polak, E. and Ribière, G. (1968) Note sur la convergence de méthodes de directions conjuguées. *Rev. franaise Informat.recherche Opérationnelle*, **16**, 35-43.

[3] Wei, Z.X., Yao, S.W. and Liu, L.Y. (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

[4] Yuan, G.L. (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

[5] 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

[6] 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

[7] 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.