一种求解非线性方程组的简化混沌算法

李海龙[1] 张美琦[2] 刘兰冬[3] 张 磊[3]
(1.中国矿业大学(北京)资源学院 2.北京航空航天大学数学与系统科学学院 3.中国矿业大学(北京)理学院)

[摘 要] 本文针对非线性方程组的求解问题提出一种将简化牛顿迭代法与混沌映射相结合的简化算法。利用混沌运动的遍历性选初值,使简化牛顿迭代法跳出局部最优,最终搜索出逼近非线性方程组的全部解。本文的简化算法可大大减少计算量,数值算例表明了该算法的正确性与有效性。

[关键词] 非线性方程组 混沌映射 牛顿迭代法 简化牛顿迭代法

1.引言

在工程计算中的许多问题常常归结为解非线性方程组

$$F(X)=0 \tag{1}$$

其中 $X=[x_1,x_2,\cdots,x_n]^T\in R^n$

非线性方程的求解问题中最经典的算法是文献[1]中的牛顿迭代法

$$X^{(k+1)}=X^{(k)}-[F'(X^{(k)})]^{-1}F(X^{(k)}) \tag{2}$$

其最大的特点是具有二阶敛速,收敛速度快,迭代函数明了,但用牛顿法求解非线性方程组时是局部最优的,即初始点的选取决定了算法的收敛性及收敛速度。且计算量大,并只能得到一个解。

混沌法能使牛顿迭代算法跳出局部最优。混沌并不是错综复杂,杂乱无章的一片混乱,而是具有精致的内在结构,是一种"奇异吸引子",能把系统的运动吸引并束缚在特定的范围内,系统运动没有什么特定的轨道而言,系统各点在混沌区内自由游荡,随机跳跃,远近不定。因而,无法预测其未来状态。初始条件极其微小的差异都会引起无法估量的巨大变化。由于混沌运动的遍历性,在一定的范围内系统按其"自身的规律"不重复地遍历所有状态。因此,利用混沌变量进行优化搜索无疑能跳出局部最优。文献[2]提出一种将解非线性方程组的牛顿迭代法与混沌映射相结合的算法能求解非线性方程组的所有解,但由于牛顿迭代法的计算量比较大,本文构造了一个混沌简化算法,既能克服牛顿迭代法的局部最优和对初值依赖性大的缺点,又能减少牛顿迭代法求解非线性方程组计算量,是求解非线性方程组的又一新方法。

2.混沌映射

混沌系统是指比较复杂的系统,它的外在表现和纯粹的随机运动很相似,即都不可预测。但和随机运动不同的是,混沌运动在动力学上是确定的,其行为表现为不确定性——不可重复、不可预测。混沌系统对无限小的初值变动和微扰也具敏感性,无论多小的扰动在长时间以后,也会使系统彻底偏离原来的演化方向,混沌在状态空间上具有遍历性。下面给出一种混沌定义:

定义(文献[6]) 设(X,f)是紧致系统,d是X的一个拓扑度量。设 $X_0\subset X$ 非空。如果存在不可数集合 $S\subset X_0$,满足:

① $\limsup d(f^n(x),f^n(y))\geq 1,\forall x,y\in S,x\neq y$;

② $\liminf d(f^n(x),f^n(y))=0,\forall x,y\in S$;

我们说f在 X_0 上是在Li-Yorke意义下是混沌的。

混沌在状态空间上具有遍历性。混沌的遍历理论认为:任一个没有稳定、周期轨道的S单峰映象f必有一个绝对连续不变测度的混沌轨道,因为它有一个测度,它必是混沌的。

$$x_{n+1}=\mu x_n(1-x_n) \tag{3}$$

式中 x_n 表示混沌变量x在n次迭代时的值。$\mu>3.569945672$ 时,系统(3)完全处于混沌状态[5]。

若取 $\mu=4$ 则有

$$x_{n+1}=4x_n(1-x_n) \tag{4}$$

文献[3]已经证明系统(4)在(0,1)范围内遍历。本文利用这种遍历性作为寻找非线性方程组初值的混沌公式,就可以找到非线性方程组的所有初始近似解。

3.简化混沌算法

3.1 简化牛顿迭代法

我们已经看到,牛顿迭代法(2)虽然收敛速度快,但每步要求一次Jacobi迭代矩阵 $F'(X^{(k)})$,要计算 n^2 个偏导 $\frac{\partial f_i(x^{(k)})}{\partial x_j}(i,j=1,2,\cdots,n)$,其计算量是很大的。如果将计算的 $F'(X^{(k)})$ 改为固定的 $F'(X^{(0)})$,则只需计算一次Jacobi矩阵 $F'(X^{(0)})$ 即可,那么在解线性方程 $F'(X^{(0)})\Delta X^{(k)}=-F(X^{(k)})$ 时,系数矩阵是固定的,解方程组时,减少了矩阵分解(LU分解)的次数,大大减少了计算量,由此得到的迭代公式为

$$X^{(k+1)}=X^{(k)}-[F'(X^{(0)})]^{-1}F'(X^{(k)}),k=0,1,2,\cdots \tag{5}$$

这种迭代方法称为简化牛顿迭代法。

3.2 混沌选初值——往返映射

本文的简化混沌算法先在精确解X的可能范围(设精确解的分量 x_i 在(a,b)内)内随意选一初值,然后由此初值通过式(4)迭代得到新的初值,如此反复迭代得到更多的初值。但我们知道当(0,1)时系统(4)才有遍历性,所以要把初值的分量 $x_i\in(a,b)$ 映射到(0,1)上的 y_i,通过(4)迭代得到 $y_i'\in(0,1)$,然后再通过逆变换把 y_i' 映射到(a,b)上的 x_i',x_i' 即为下一初值。其具体过程(往返映射)如下:

往返映射:

对X(i)的分量 $x_i\in(a,b)$ 可由式(4)以及下面的(6)、(7)与混沌变量 $y_i\in(0,1)$ 进行往返映射,

即:

$$y_i=(x_i-a)/(b-a) \tag{6}$$

把 y_i 代入(4)得新的 y_i,再进行式(6):

$$x_i=a+y_i(b-a) \tag{7}$$

得到的 x_i 即为下一初值的第i个分量。

3.3 基于混沌初值的简化牛顿迭代算法

由于混沌初值的牛顿迭代算法的计算量很大,我们用简化法可使计算量大大减少。其算法步骤如下:

(1)输入a,b与初始近似解X(0),变换初值的最高次数 $Max(X_0)$,解一个非线性方程组的所能用的最高迭代次数Max,误差限 $\varepsilon_1,\varepsilon_2$

(2)对 $i=0,\cdots Max(X_0)-1$,若 $\|F(X(i))\|_\infty<\varepsilon_1$,做 $X^{(0)}=X(i)$,$F'(X^{(0)})$,后做(2.1),否则至(2.2)

(2.1)对 $k=0,\cdots,Max-1$ 做(2.1.1)-(2.1.4)

(2.1.1)计算 $F(X^{(k)})$;

(2.1.2)解非线性方程组 $F'(X^{(0)})Y=-F(X^{(k)})$ 得到Y;

(2.1.3)若 $\|Y\|<\varepsilon_2$ 则输出 $X^{(k)}$,则至(2.2);

(2.1.4) $X^{(k+1)}\leftarrow X^{(k)}+Y$;

(2.2)X(i)的各分量 x_i 由公式(6)得到 y_i,进而得到Y(i);

(2.3)Y(i)的各分量 y_i 由公式(4)得到新的 y_i,进而得到新的Y(i);

(2.4)Y(i)的各分量 y_i 由公式(7)得到 x_i 进而得到下一个初值X(i+1),并返回(2)的循环。

4.实例分析

例.(2008年全国大学生建模竞赛题目)数码相机定位在交通监管(电子警察)等方面有广泛的应用。所谓数码相机定位是指用数码相机摄制物体的相片确定物体表面某些特征点的位置。最常用的定位方法是双目定位,即用两部相机来定位。对物体上一个特征点,用两部固定于不同位置的相机摄得物体的像,分别获得该点在两部相机像平面上的坐标。只要知道两部相机精确的相对位置,就可用几何的方法得到该特征点在固定一部相机的坐标系中的坐标,即确定了特征点的位置。于是对双目定位,精确地确定两部相机的相对位置就是关键,这一过程称为系统标定。

标定的一种做法是:在一块平板上画若干个点,同时用这两部相机照相,分别得到这些点在它们像平面上的像点,利用这两组像点的几何关系就可以得到这两部相机的相对位置。然而,无论在物平面或像平面上我们都无法直接得到没有几何尺寸的"点"。实际的做法是在物平面上画若干个圆(称为靶标),它们的圆心就是几何的点了。而它们的像一般会变形,所以必须从靶标上的这些圆的像中把圆心的像精确地找到,标定就可实现。

有人设计靶标如下，取 1 个边长为 100mm 的正方形，分别以四个顶点（对应为 A、C、D、E）为圆心，12mm 为半径作圆。以 AC 边上距离 A 点 30mm 处的 B 为圆心，12mm 为半径作圆，如图 1 所示。

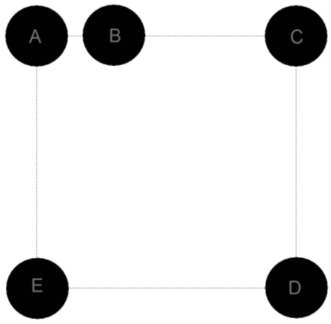

图 1 靶标示意图

用一位置固定的数码相机摄得其像，如图 2 所示。

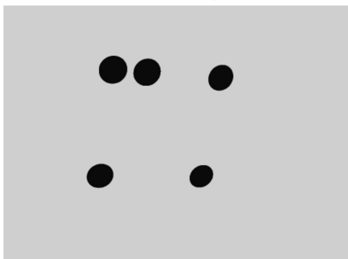

图 2 靶标的像

问：对由图 1、图 2 分别给出的靶标及其像，计算靶标上圆的圆心在像平面上的像坐标，该相机的像距（即焦点到像平面的距离）是 1577 个像素单位(1 毫米约为 3.78 个像素单位)，相机分辨率为 1024× 786。

经过分析我们得到如下的非线性方程组：

$$\begin{cases}(k_0x_0-k_1x_1)^2+(k_0y_0-k_1y_1)^2+(k_0z_0-k_1z_1)^2=m^2\\(k_0x_0-k_2x_2)^2+(k_0y_0-k_2y_2)^2+(k_0z_0-k_2z_2)^2=l^2\\(k_2x_2-k_1x_1)^2+(k_2y_2-k_1y_1)^2+(k_2z_2-k_1z_1\pi\theta)^2=n^2\\k_0(A\times x_0+B\times y_0+C\times z_0)+D=0\\k_1(A\times x_1+B\times y_1+C\times z_1)+D=0\\k_2(A\times x_2+B\times y_2+C\times z_2)+D=0\end{cases}\tag{8}$$

其中已知量为：
$$\begin{aligned}&(x_0,y_0,z_0)=(323,578.5,1157);\\&(x_1,y_1,z_1)=(640,554.5,1157);\\&(x_2,y_2,z_2)=(284.5,266,1157);\\&(D,m,l,n)=(1,378,378,534.5727)\end{aligned}$$

前三个等式是各点的位置约束条件，后三个等式是平面约束条件。其中除了 k_0,k_1,k_2,A,B,C 是未知的，要求其它量都是已知的点或数。k_0,k_1,k_2 分别是是点 1、2、3 在物理象限坐标与视觉象限坐标的向量长度比。A,B,C,D 分别是视觉象限平面的确定系数，其中 D 是已知，A,B,C 未知。因此我们现在要求解的就是 k_0,k_1,k_2,A,B,C。

其精确解为：

$X_1=(1.224295\ \ 1.203926\ \ 1.255233\ \ -0.000048\ \ -0.000059\ \ -0.000663)^T$

$X_2=(-1.224295\ \ -1.203926\ \ -1.255233\ \ 0.000048\ \ 0.000059\ \ 0.000663)^T$

$X_3=(1.181583\ \ 0.959748\ \ 1.348737\ \ -0.000637\ \ -0.000257\ \ -0.000425)^T$

$X_4=(-1.181583\ \ -0.959748\ \ -1.348737\ \ 0.000637\ \ 0.000257\ \ 0.000425)^T$

取初值 $\varepsilon_1=0.01$，$\varepsilon_2=0.000001$，Max=3000，$Max(X_0)=300$　$X(0)=(1.9,1.9,1.9,1.9,1.9,1.9)^T$ 分别用混沌初值的牛顿迭代算法和混沌初值的简化牛顿迭代算法解方程组，足可说明此两种方法用来解非线性方程组的可行性。其中，

$X(0)=(1.686400\ \ 1.998400\ \ 1.686400\ \ 1.686400\ \ 1.686400\ \ 1.686400)^T$

$X(1)=(-1.984562\ \ -1.115732\ \ -1.985019\ \ -1.985019\ \ -1.985019\ \ -1.985019)^T$

$X(2)=(1.432935\ \ 1.991079\ \ 1.902254\ \ 1.902254\ \ 1.902254\ \ 1.902254)^T$

$X(3)=(1.031487\ \ 1.795133\ \ 1.509485\ \ 1.509485\ \ 1.509485\ \ 1.509485)^T$

$X(4)=(0.759218\ \ 1.692956\ \ 1.508177\ \ 1.508177\ \ 1.508177\ \ 1.508177)^T$

$X(5)=(-0.988420\ \ -1.990097\ \ -1.575382\ \ -1.575382\ \ -1.575382\ \ -1.575382)^T$

$X(6)=(1.344326\ \ 1.998051\ \ 1.461576\ \ 1.461576\ \ 1.461576\ \ 1.461576)^T$

$X(7)=(-1.432546\ \ -1.521368\ \ -1.710003\ \ -1.710003\ \ -1.710003\ \ -1.710003)^T$

$X(8)=(-1.400642\ \ -1.367221\ \ -1.281528\ \ -1.281528\ \ -1.281528\ \ -1.281528)^T$

$X(9)=(1.998540\ \ 1.982916\ \ 1.872061\ \ 1.872061\ \ 1.872061\ \ 1.872061)^T$

$X(10)=(-1.994164\ \ -1.931954\ \ -1.504612\ \ -1.504612\ \ -1.504612\ \ -1.504612)^T$

$X(11)=(1.874445\ \ 0.955642\ \ 1.152574\ \ 1.152574\ \ 1.152574\ \ 1.152574)^T$

求解例题的两种方法的比较如下表（只需变换初值 12 次便可）：

方法	求出所有根的初值的最小的变换次数	初值	解此初值对应的线性方程组的迭代次数	分解 Jacobi 矩阵次数（LU 分解次数）	求出所有解的总时间(s)
本文的简化混沌算法	12	X(0)	125	1	1.142546s
		X(1)	99	1	
		X(2)	360	1	
		X(3)	409	1	
		X(4)	711	1	
		X(5)	484	1	
		X(6)	197	1	
		X(7)	203	1	
		X(8)	46	1	
		X(9)	96	1	
		X(10)	92	1	
		X(11)	140	1	
文献[2]混沌算法	12	X(0)	97	97	2.774755s
		X(1)	32	32	
		X(2)	61	61	
		X(3)	43	43	
		X(4)	36	36	
		X(5)	35	35	
		X(6)	34	34	
		X(7)	35	35	
		X(8)	26	26	
		X(9)	31	31	
		X(10)	31	31	
		X(11)	29	29	

5.结论

混沌初值的牛顿法求解非线性方程组避免了收敛性对初始点的依赖性，且可以找到全部解，而本文的方法又在混沌初值的牛顿法基础上减少了计算量与时间。数值例子很好地说明了混沌初值的牛顿简化法的有效性与时效性，它迭代次数少，计算速度快，是解非线性方程组的一种较有效的算法，特别在求解大型工程问题时更能显现简化算法的优越性。

参考文献

[1] 奥特加，莱因博尔特.多元非线性方程组迭代解法 [M] .北京：科学出版社，1983:613- 615.

[2] 刘健，袁建平. 一种求解非线性方程组的混沌算法 [J] .哈尔滨商业大学学报，2001,17(1):31- 34.

[3] 卢侃，孙建华等.M- t 动力学 [M] .上海:上海翻译出版公司，1990.

[4] 黄象鼎，曾钟钢，马亚南.非线性数值分析的理论与方法 [M] .武汉：武汉大学出版社，2004.

[5] Daniel Kaplan,Leon Glass.Understanding Nonlinear Dynamics [M] .北京:世界图书出版公司北京公司, 2004.

[6] 钱树华. 一种求解非线性方程组的混沌优化算法 [J] .邢台职业技术学院学报,2006,01- 0070- 03.