第 40 卷第 6 期 湖北工程学院学报 VOL.40 NO.6
2020 年 11 月 JOURNAL OF HUBEI ENGINEERING UNIVERSITY NOV.2020

梯度下降神经网络方法求解雅可比矩阵奇异的非线性方程组

雍龙泉[1,2], 贾 伟[1]
(1.陕西理工大学 数学与计算机科学学院, 陕西 汉中 723001;2.陕西省工业自动化重点实验室, 陕西 汉中 723001)

摘 要:给出了一种雅可比矩阵奇异的非线性方程组的求解方法,通过把非线性方程组转化为一个无约束优化问题,采用梯度下降神经网络求解。对于唯一根的非线性方程组,该方法能够收敛到其唯一根;对于具有多个根的非线性方程组,该方法能够找到尽可能多的根。并且,该方法无需计算非线性方程组的雅克比矩阵,适用范围广。

关键词:非线性方程组; 无约束优化; 梯度下降神经网络; 雅克比矩阵

中图分类号:TP183;O221;O224;O241 **文献标志码**:A **文章编号**:2095-4824(2020)06-0069-05

1 考虑非线性方程组

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0, \\ f_2(x_1, x_2, \cdots, x_n) = 0, \\ \vdots \\ f_{n-1}(x_1, x_2, \cdots, x_n) = 0, \\ f_n(x_1, x_2, \cdots, x_n) = 0. \end{cases} \tag{1}$$

记向量 $x = (x_1, x_2, \cdots, x_n)$, 向量函数 $F(x) = (f_1(x), f_2(x), \cdots, f_n(x))^T$, 则方程组(1)等价于 $F(x) = 0$。求解非线性方程组的常见方法是牛顿法以及各类改进的牛顿法[1-11]。近年来,相继出现了五阶牛顿法[12-16]、七阶牛顿法[17-19]、八阶牛顿法[20-21]、九阶牛顿法等[22-23]。当然,收敛阶数越高,(从 x_n 到 x_{n+1} 每迭代一次的)计算量也就越大。因此,在构造高阶牛顿迭代法求解非线性方程时,既需要考虑收敛阶,更需要考虑计算效率[24]。牛顿型算法依赖于初始点的选取和函数的性态,事实上初始点的选取本身也是一个比较困难的问题;牛顿型算法在实际计算过程中,若其雅克比矩阵存在奇异点,即 $det(F'(x)) = 0$ 有解,则不能直接应用牛顿型算法。

本文假设问题(1)的解存在,建立了求解非线性方程组的梯度下降神经网络,通过求解唯一解、多个解的非线性方程组,结果表明该方法不依赖初始点。鉴于该方法无需考虑 $F(x)$ 的雅克比矩阵,因此对雅克比矩阵存在奇异点的非线性方程组也适用,且该方法通过改变初始点,能找到多个解。

2 梯度下降神经网络

为了避免计算 $F(x)$ 的雅克比矩阵,定义函数

$$E(x) = \frac{1}{2}\|F(x)\|^2, \tag{2}$$

称 $E(x)$ 为能量函数。于是,求解非线性方程组就转化为求解连续可微的优化问题 $\min E(x)$。

梯度神经网络是基于求解问题(2)的最速下降模型的连续化形式,近年来已广泛应用于求解

收稿日期:2020-07-30
基金项目:国家自然科学基金(11401357);陕西省教育厅重点科学研究计划项目(20JS021);陕省教育厅专项科研计划项目(17JK0146);陕西理工大学科研项目(SLGYQZX2002);陕西理工大学教学改革研究项目(SLGYJG2015)
作者简介:雍龙泉(1980-),男,陕西汉中人,陕西理工大学数学与计算机科学学院教授,博士。
贾 伟(1977-),男,四川绵阳人,陕西理工大学数学与计算机科学学院讲师,硕士。

非线性互补、二阶锥规划等问题[25-29]。梯度下降神经网络模型如下:

$$\frac{dx}{dt}=-\tau\nabla E(x)=-\tau[F'(x)]^{T}F(x),\tau>0 \quad (3)$$

参数 τ 表示梯度下降算法的步长,$\nabla E(x)$ 表示能量函数 $E(x)$ 的梯度。该神经网络的收敛性证明详见文献[30-31]。下面来计算一些常见的非线性方程组,以验证算法的有效性。

3 数值实验

给出 4 个非线性方程组问题(前 3 个具有唯一解,第 4 个具有 2 个解),通过将其转化为神经网络(3),采用四阶 Runge-Kutta 法求解微分方程组,程序采用 Matlab R2009a(内置函数 ode45)编写。设置 $E(x)\leqslant 1\times 10^{-10}$ 为终止条件。

问题 3 的初始点取为 $x^{(0)}=1\times \mathrm{rand}(n,1)$,这里 rand 表示 0~1 之间的随机数;其余问题的初始点取为 $x^{(0)}=1\times \mathrm{rand}(n,1)-1\times \mathrm{rand}(n,1)$,即初始点在任一象(卦)限随机选取,这对于具有多个解的非线性方程组,(多次运行)就有可能找到尽可能多的解。

算例 1 考虑非线性方程组

$$\begin{cases}x_1+e^{x_2}-\cos x_2=0\\3x_1-\sin x_1-x_2=0\end{cases}$$

其解为 $x^*=(0,0)^T$。表 1 给出了参数 τ 取不同值的计算结果;图 1 和图 2 分别给出了 $\tau=10$ 时近似解随时间的变化(轨线)及能量函数随时间的变化曲线。

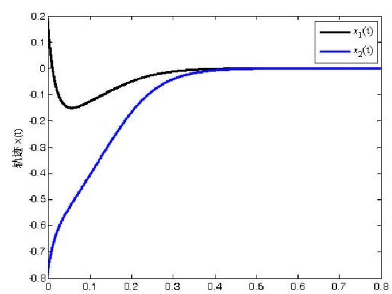


图 1 近似解随时间的变化曲线

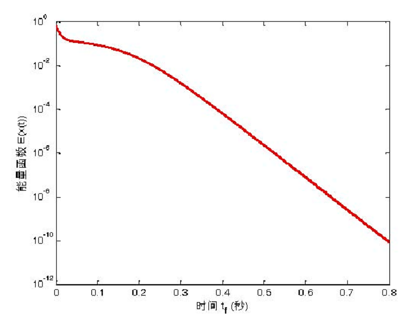


图 2 能量函数随时间的变化曲线

算例 2 非线性方程组

$$\begin{cases}10x_1+\sin(x_1+x_2)-1=0\\8x_2-\cos^2(x_3-x_2)-1=0\\12x_3+\sin x_3-1=0\end{cases}$$

其解为 $x^*=(0.068978,0.246442,0.076929)^T$。表 1 给出了参数 τ 取不同值的计算结果;图 3 和图 4 分别给出了 $\tau=10$ 时近似解随时间的变化(轨线)及能量函数随时间的变化曲线。

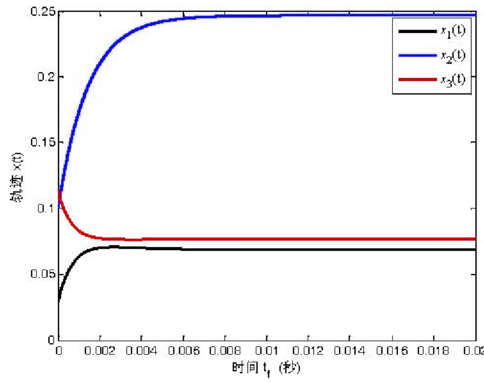


图 3 近似解随时间的变化曲线

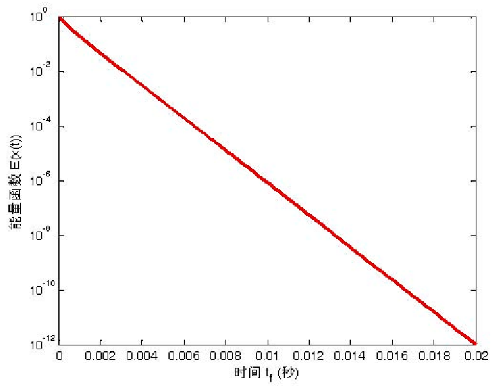


图 4 能量函数随时间的变化曲线

算例 3 考虑非线性方程组

$$\begin{cases}x_1+10x_2=0\\2(x_3-x_1)=0\\(x_2-2x_3)^2=0\\3(x_1-x_4)^2=0\end{cases}$$

其解为 $x^*=(0,0,0,0)^T$。该方程虽形式简单(第 1 个方程仅反映了 x_1 与 x_2 的关系,第 2 个方程仅反映了 x_1 与 x_3 的关系,第 3 个方程仅反映了 x_2 与 x_3 的关系,第 4 个方程仅反映了 x_1 与 x_4 的关系,变量之间的相互关联性较弱),但是求出精确解却不易。计算可得雅克比矩阵为

$$\begin{bmatrix}1&10&0&0\\-2&0&2&0\\0&2x_2-4x_3&-4x_2+8x_3&0\\6x_1-6x_4&0&0&-6x_1+6x_4\end{bmatrix}$$

雅克比矩阵的奇异点分布在 $x_1 - x_4 = 0$ 或 $x_2 - 2x_3 = 0$ 上。这里雅克比矩阵存在奇异点，故牛顿法不再适用。采用梯度下降神经网络求解，表 1 给出了参数 $\tau = 10$ 的计算结果；图 5 和图 6 分别给出了 $\tau = 10$ 时近似解随时间的变化(轨线)及能量函数随时间的变化曲线。

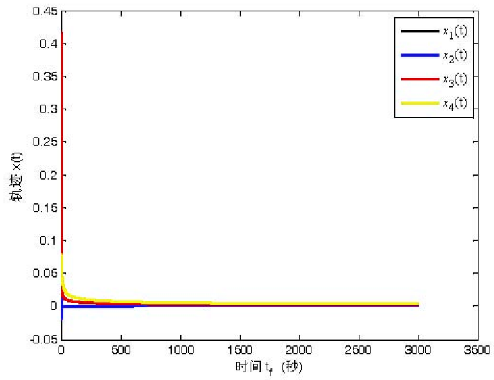


图 5　近似解随时间的变化曲线

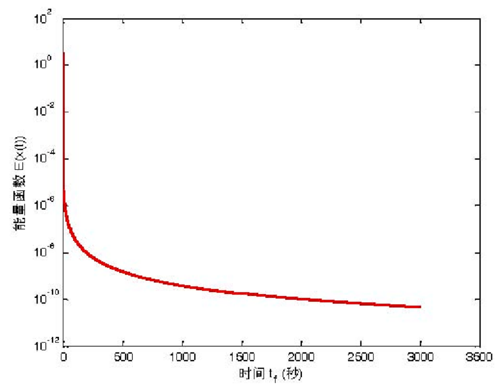


图 6　能量函数随时间的变化曲线

表 1　参数 τ 取不同值的计算结果

	参数 τ 值	t_f	$\|\|x(t_f) - x^*\|\|$	$E(x(t_f))$
算例 1	$\tau = 1$	6	6.112e-006	3.17e-011
	$\tau = 10$	0.8	9.939e-006	8.382e-011
算例 2	$\tau = 1$	0.2	6.961e-007	1.651e-011
	$\tau = 10$	0.02	1.718e-007	1.006e-012
算例 3	$\tau = 10$	3001	0.003253	4.405e-011

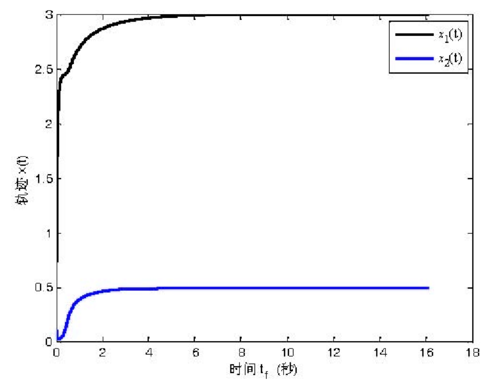
算例 4　考虑非线性方程组

$$\begin{cases} x_1(1 - x_2^2) - 2.25 = 0, \\ x_1(1 - x_2^3) - 2.625 = 0. \end{cases}$$

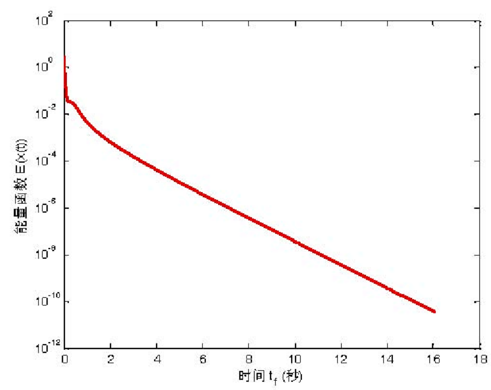
其解为 $x^* = (3, 0.5)^T$ 及 $x^* = (81/32, -1/3)^T$，计算可得雅克比矩阵为

$$\begin{bmatrix} 1 - x_2^2 & -2x_1x_2 \\ 1 - x_2^3 & -3x_1x_2^2 \end{bmatrix}$$

雅克比矩阵的奇异点分布在四条直线 $x_1 = 0$, $x_2 = 1$, $x_2 = 0$, $x_2 = -2$ 上。这里雅克比矩阵存在奇异点，故牛顿法不再适用。采用梯度下降神经网络求解，图 7 和图 8 分别给出了 $\tau = 10$ 时，初始点在不同象限时的近似解随时间的变化曲线及能量函数随时间的变化曲线。



(a) 近似解随时间的变化曲线

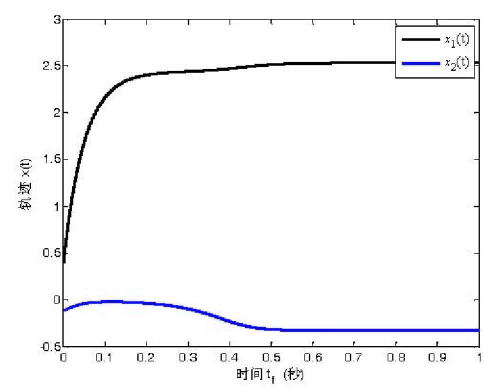


(b) 能量函数随时间的变化曲线

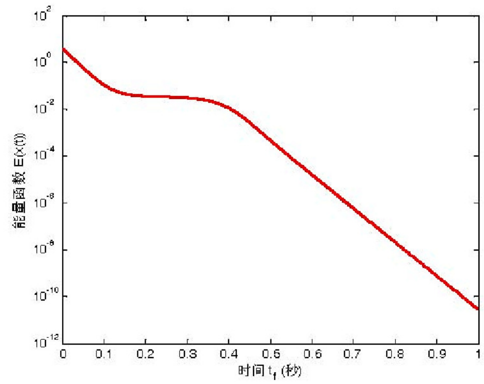
图 7　初始点在第 1 象限时，收敛到第 1 象限的解

4　结束语

本文计算过程表明：针对唯一解的非线性方程组，该方法能够收敛到其唯一解；针对具有多个解的非线性方程组，该方法能够找到尽可能多的解；且该算法无需考虑其雅克比矩阵，因此对雅克比矩阵是否存在奇异点都适用。由于线性两点边



(a) 近似解随时间的变化曲线



(b) 能量函数随时间的变化曲线

图 8　初始点在第 4 象限时,收敛到第 4 象限的解

(说明:初始点在那个象限,则收敛到该象限的解)

值问题离散化后得到线性方程组,非线性两点边值问题离散化后所得到的非线性方程组[32-34],下一步可以利用上面方法求解非线性两点边值问题。

[参　考　文　献]

[1] 李庆扬. 解非线性方程组的离散型延拓法[J]. 数值计算与计算机应用,1984,5(2):114-124.

[2] 李庆扬,莫孜中,祁力群. 非线性方程组的数值解法[M]. 北京:科学出版社,1987.

[3] 范金燕,袁亚湘. 非线性方程组数值方法[M]. 北京:科学出版社,2018.

[4] CHUN C. Iterative methods improving Newton's method by the decomposition method[J]. Computers & Mathematics with Applications, 2005, 50(10):1559-1568.

[5] 马元婧. 非线性方程组的一种修正牛顿法及其连续型[D]. 哈尔滨: 哈尔滨工业大学, 2009.

[6] 郭俊, 吴开腾, 张莉, 等. 一种新的求非线性方程组的数值延拓法[J]. 计算数学,2017,39(1):33-41.

[7] 伍渝江,陈亮. 非线性方程组的非交替 Newton-PHSS 迭代法[J]. 应用数学与计算数学学报,2017,31(2):153-162.

[8] 萨和雅. 非线性方程组的锥模型方法研究[D]. 呼和浩特: 内蒙古大学,2017.

[9] 罗亚中,袁端才,唐国金. 求解非线性方程组的混合遗传算法[J]. 计算力学学报, 2005, 22(1):109-114.

[10] LI Q , LI D H. A class of derivative-free methods for large-scale nonlinear monotone equations[J]. IMA Journal of Numerical Analysis, 2011, 31(4):1625-1635.

[11] 张娜. 求解非线性方程的智能优化算法[D]. 长春:吉林大学,2013.

[12] RAFIULLAH M. A fifth-order iterative method for solving nonlinear equations[J]. Numerical Analysis & Applications, 2011, 4(3):239-243.

[13] LIU T, CAI H. A Family of Fifth-order Iterative Methods for Solving Nonlinear Equations[J]. Communications in Mathematical Research, 2013, 29(3):255-260.

[14] SHARMA J R, GUPTA P. An efficient fifth order method for solving systems of nonlinear equations[J]. Computers & Mathematics with Applications, 2014, 67(3):591-601.

[15] LIANG J, LI X, WU Z, et al. Fifth-Order Iterative Method for Solving Multiple Roots of the Highest Multiplicity of Nonlinear Equation[J]. Algorithms, 2015, 8(3):656-668.

[16] 裕静静,江平,刘植. 两类五阶解非线性方程组的迭代算法[J]. 计算数学,2017,39(2):151-166.

[17] KOU J, LI Y, WANG X. Some variants of Ostrowski's method with seventh-order convergence[J]. Journal of Computational & Applied Mathematics, 2007, 209(2):153-159.

[18] AHMED S O. An improvement of Jarratt method with seventh-order convergence[J]. International Journal of Scientific Research and Innovative Technology, 2015, 2(5):97-102.

[19] NARANG M, BHATIA S, KANWAR V. New efficient derivative free family of seventh-order methods for solving systems of nonlinear equations[J]. Numerical Algorithms, 2017, 76(1):1-25.

[20] 王晓锋, 张铁. 一类求解非线性方程最优的 8 阶收敛迭代法[J]. 吉林大学学报(理学版), 2013, 51(4): 568-572.

[21] BI W, REN H, WU Q. Three-step iterative methods with eighth-order convergence for solving nonlinear equations[M]. Journal of Computational and Applied Mathematics, 2009, 225:105-112.

[22] HU Z, LIU G, TIAN L. An iterative method with ninth-order convergence for solving nonlinear equations[J]. International Journal of Contemporary Mathematical Sciences, 2011, 6(1-4):17-23.

[23] 张旭,檀结庆,艾列富. 一种求解非线性方程组的 3p 阶迭代方法[J]. 计算数学,2017,39(1):14-22.

[24] 陈香萍. 谱 HS 投影算法求解非线性单调方程组[J]. 运筹学学报,2018, 22(3):15-27.

[25] CHEN J S, KO C H, PAN S H. A neural network based on generalized Fischer-Burmeister function for nonlinear complementarity problems[J]. Information Sciences, 2010, 180: 697-711.

[26] KO C H, CHEN J S, YANG C Y. Recurrent neural networks for solving second-order cone programs[J]. Neurocomputing, 2011, 74(17):3646 - 3653.

[27] MIAO X, CHEN J S, KO C H. A neural network based on the generalized FB function for nonlinear convex programs with second-order cone constraints[J]. Neurocomputing, 2016, 203:62 - 72.

[28] SUN J H, WU X R, SAHEYA B, et al. Neural network for solving SOCQP and SOCCVI based on two discrete-type classes of SOC complementarity functions[J]. Mathematical Problems in Engineering, 2019, 2019(PT.4):1 - 18.

[29] ALCANTARA J H, CHEN J S. Neural networks based on three classes of NCP-functions for solving nonlinear complementarity problems [J]. Neurocomputing, 2019, 359(SEP.24):102 - 113.

[30] 赵华敏, 陈开周. 用神经网络解非线性方程组[J]. 西安电子科技大学学报(自然科学版), 2001, 28(1):35 - 38.

[31] 赵华敏, 陈开周. 解非线性方程组的神经网络方法[J]. 电子学报, 2002, 30(4):601 - 604.

[32] 雍龙泉. 一种改进的和声搜索算法求解线性两点边值问题[J]. 数学的实践与认识, 2019, 49(10):226 - 233.

[33] 雍龙泉. 一类高阶牛顿迭代法及其在非线性两点边值问题中的应用[J]. 重庆理工大学学报(自然科学), 2019, 33(7):242 - 247.

[34] 雍龙泉. 一种改进的和声搜索算法求解非线性方程组[J]. 重庆理工大学学报(自然科学版), 2020, 34(10): 231 - 237.

Solving Nonlinear Equations with Singular Jacobi Matrix Based on Gradient Descend Neural Network

Yong Longquan[1,2], Jia Wei[1]

(1. *School of Mathematics and Computer Science, Shaanxi University of Technology, Hanzhong, Shaanxi* 723001, *China*; 2. *Shaanxi Key Laboratory of Industrial Automation, Hanzhong, Shaanxi* 723001, *China*)

Abstract: A method is presented for nonlinear equations with singular Jacobi matrix. After transforming nonlinear equations into an unconstrained optimization problem, the nonlinear equations can be solved by gradient descend neural network model. For nonlinear equations with unique root, this method can converge to its unique root; and for nonlinear equations with several roots, this method can find the roots as many as possible. In addition, the method is widely used in application without calculation of Jacobi matrix.

Key Words: nonlinear equations; unconstrained optimization; gradient descend neural network; Jacobi matrix

(责任编辑:熊文涛)