第 23 卷第 4 期 控 制 理 论 与 应 用 Vol. 23 No. 4
2006 年 8 月 Control Theory & Applications Aug. 2006

文章编号：1000−8152(2006)04−0645−04

求解广义特征根问题的反馈神经网络方法

冯芙叶[1,3], 魏正红[2], 刘付显[1]

(1. 空军工程大学 导弹学院, 陕西 三原 710038; 2. 深圳大学 师范学院 数学系, 广东 深圳 518060;

3. 东莞理工学院 软件学院, 广东 东莞 523000)

摘要: 研究了广义特征根问题求解的神经网络方法, 给出了求解该问题的一个时间连续性反馈网络模型, 利用LaSalle不变原理分析并证明了该网络的拟全局收敛性, 这是网络能够确切的求解广义特征根问题的保证. 同时, 该网络解决了已有的基于罚函数方法构造的特征根问题的神经网络存在的一些基本缺陷：其一, 基于罚函数的网络模型所得到的解可能不是真解, 甚至可能都不是可行解；其二, 它们的共同缺陷是有一个需要调节的参数, 但是参数的选择并没有一个可供参考的准则；其三, 这些模型的稳定性无法保证. 本文所提出的网络模型解决了这些问题, 并且, 此网络具有一个很好的特征就是在初始点选定在问题的可行解集的话, 网络轨线将永远是可行的并收敛到一个广义特征向量. 最后, 数值模拟也表明这里所提出的网络的可靠性能, 进一步证明了此网络可以很好地求解广义特征根问题.

关键词: 广义特征根问题; 反馈神经网络; 拟全局稳定性

中图分类号: TP273 **文献标识码**: A

Recurrent neural network for solving the generalized eigenvalue problem

FENG Fu-ye[1,3], WEI Zheng-hong[2], LIU Fu-xian[1]

(1. Missile College, Engineering University of Air Force, Sanyuan Shaanxi 713800, China;

2. Mathematics Department of Normal Insitute, Shenzhen University, Shenzhen Guangdong 518060, China;

3. Software College, Dongguan Polytechnic University, Dongguan Guangdong 523000, China)

Abstract: Neural network method for solving generalized eigenvalue and eigenvector problems is studied in this paper and a continuous-time recurrent neural network model is presented. By using LaSalle's invariant principle, it is shown that the proposed network is globally quasi-convergent which guarantees an exact generalized eigenvector that can be found by the new model. Furthermore, this new model overcomes the following three defects in the existing neural network models based on penalty function method. 1) False solutions may be found by using penalty function method, sometimes, even unfeasible solutions could be found. 2) There is a parameter to be tuned in the process, but no definite rule available for the tuning. 3) No any stability result could be ensured for the existing models. The new model proposed here solves all these problems and, moreover, it has a good characteristic that it's trajectories can never escape from the feasible region, but will converge to the set of the generalized eigenvectors for any initial point in the feasible set. Finally, good performance of the proposed model for finding a generalized eigenvector is also demonstrated by numerical simulation results.

Key words: generalized eigenvalue problem; recurrent neural network; global quasi-convergence

1 引言(Introduction)

计算矩阵$A \in \mathbb{R}^{n\times n}$ 的特征根问题与其相对应的特征向量问题在许多科学计算以及工程应用中是一个必须的步骤, 比如信号处理, 控制理论, 以及地球物理等. 当然, 寻找新的求解方法是数值线性代数研究的一个重要问题.关于这方面的经典方法以及最新进展可以在 Golub 的书[1] 以及他的一篇综述性文章[2]中找到. 传统方法由于受到时间的限制, 对大规模的科学计算问题已经很不适用. 由于神经网络算法具有实时计算以及易于电路模拟实现的两大强大功效, 自从Hopfield与Tank[3,4]的开创性工作开始, 最近20余年已有许多文献研究了最优化问题的神经网络求解, 可参看文献[3∼6]. 有关最优化问题求解的神经网络的一些基本模型、方法表述等可以参看

收稿日期: 2005−04−12; 收修改稿日期: 2005−08−24.

专著[7].

特征根问题的神经网络求解也有文献涉及, 例如文献[7,8] 就提出了相应的网络模型, 他们的网络模型是基于罚函数方法的. 但是, 基于罚函数方法所寻找到的解很可能不是可行解, 更别说是最优解了, 可参看文献[9]中的论述. 另外, 网络的收敛性在这种方法中更是没法证明. 因此, 这种方法所得到的网络模型是不可靠的, 其有效性与可靠性都有极大的缺陷. 这里, 本文提出求解广义特征根问题的神经网络模型, 与已有的网络相比较, 这里的模型一直保持轨道是可行的, 而且全局拟收敛到广义特征向量. 当然, 所提出的网络模型具有实时计算的强大功效, 而且可以电路模拟实现, 从这点来讲, 这个方法可以用于解决大规模的计算问题.因此, 这里的方法明显优于已有的传统方法.

2 广义特征根问题(Generalized eigenvalue problems)

广义特征根问题就是寻找一个数λ, 参看文献[1,2], 以及与之对应的特征向量$v = [v_1, \cdots, v_n]^{\mathrm{T}} \neq 0 \in \mathbb{R}^n$使得对两个实矩阵$A, B \in \mathbb{R}^{n\times n}$ 满足下面的代数方程组:

$$Av = \lambda Bv, \tag{1}$$

通常矩阵对 (A, B) 称为一个盒子. 显然, 当 v 是一个广义特征向量时, 它的任何一个非零因子 α 的乘积也是一个特征向量: $A(\alpha v) = \lambda B(\alpha v)$. 因此, 通常在正则化方面要求特征向量是关于矩阵B的模为单位向量, 即要求

$$v^{\mathrm{T}}Bv = 1. \tag{2}$$

显然, 当 $B = I$, 这里 I 是单位阵, 广义特征根问题就简化为普通的特征根问题.

基于罚函数方法, 文献[7,8] 给出了求解普通的特征根问题的多层神经网络模型. 但是, 基于罚函数的方法有3个基本缺陷: 1) 没有一个调节这个罚参数的准则可以利用; 2) 所得到的解可能不是真解; 3) 这些模型的稳定性通常无法证明[6~8]. 有鉴于此, 罚函数方法在传统的优化计算中几乎没有用处. 因此, 寻找求解广义特征根问题的新的神经网络模型就显得非常迫切.

3 网络模型与稳定性分析(Neural network model and its stability)

考虑由下面的微分方程组所刻画的神经网络系统

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -(\|Bx\|)^2 Ax + (x^{\mathrm{T}}BAx)Bx. \tag{3}$$

其中: $x \in \mathbb{R}^n$, 以及$A, B \in \mathbb{R}^{n\times n}$分别是实对称以及实对称正定矩阵. 这是一个由自治微分系统所表达的神经网络系统, 其块状流程如图 1 所示.

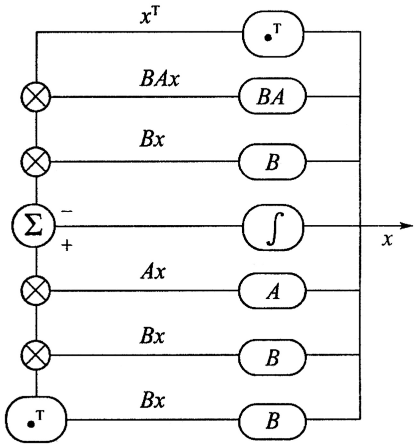

图 1 网络3的块状流程

Fig. 1 Block diagram of the neural network 3

该网络是由一些基本的线性变换器A, B, 积分器$\int$, 乘法器$\times$, 以及加法器Σ构成, 这种基本单元的电路硬件实现在文献[7] 中有详细的论述, 它们都是可以很好实现的电路网络单元.

笔者注意到网络(3)的任何非零平衡点x正好是盒子(A, B)的一个广义特征向量, 而对应的特征值是$\lambda = (x^{\mathrm{T}}BAx)/(\|Bx\|^2)$; 反之, 如果非零向量$x$是盒子$(A, B)$对应于特征值$\lambda$的一个特征向量, 则从$Ax = \lambda Bx$以及$\lambda = (x^{\mathrm{T}}BAx)/(\|Bx\|^2)$可知$x$必是网络(3)的一个平衡点. 因此, 寻求广义特征向量的问题就转化为寻找网络（3）的非平凡平衡点的问题. 网络能否收敛到平衡点, 也就是网络的稳定性问题, 就是首先需要讨论的问题.

笔者给出与一个一般的自治网络系统$\dot{x} = F(x)$稳定性相关的一些基本概念.

定义 1 设$x(t)$是非线性自治网络系统$\dot{x} = F(x)$的一个解. 这里称系统相对于集合W全局收敛到集合X, 如果每个从集合W中出发的解$x(t)$满足条件

$$\rho(x(t), X) \to 0,\ t \to \infty, \tag{4}$$

这里: $\rho(x(t), X) = \inf\limits_{y\in X} \|x - y\|$, $x(0) = x_0 \in W$.

定义 2 设$x(t)$是非线性自治网络系统$\dot{x} = F(x)$ 的一个解. 这里称这个系统对集合W是初始的, 如果每个从集合W中出发的解$x(t)$对所有的$t \geqslant 0$将停留在此W内. 如果网络系统对某个数学规划的可行解集合W 是初始的并且收敛到网络的平衡点集合, 称此网络系统是关于W拟全局收敛的.

下面可以叙述网络(3)的全局收敛性了.

定理 1 网络(3)对可行解集合$W = \{x|x^{\mathrm{T}}Bx = 1\}$是初始的, 而且每个解收敛到盒子$(A,B)$的特征向量集合, 而它就是网络的平衡点集合, 即就是说网络(3)是拟全局收敛的.

证 从式(3)可得

$$\frac{\mathrm{d}x^{\mathrm{T}}Bx}{\mathrm{d}t} = 2x^{\mathrm{T}}B\frac{\mathrm{d}x}{\mathrm{d}t} = \tag{5}$$

$$2[-x^{\mathrm{T}}B\|Bx\|^2Ax + (x^{\mathrm{T}}BBx)x^{\mathrm{T}}BAx] = \tag{6}$$

$$2[-\|Bx\|^2x^{\mathrm{T}}BAx + \|Bx\|^2x^{\mathrm{T}}BAx] = \tag{7}$$

$$0, \tag{8}$$

即就是$x^{\mathrm{T}}Bx$沿着式(3)的轨线为常数, 从而

$$x(t)^{\mathrm{T}}Bx(t) = x_0^{\mathrm{T}}Bx_0 = 1, \tag{9}$$

此即说明系统(3)对W是初始的.

考虑能量函数$V(x) = \frac{1}{2}x^{\mathrm{T}}Ax$, 计算它沿着网络系统的能量变化, 即求全导数

$$\frac{\mathrm{d}V}{\mathrm{d}t} = x^{\mathrm{T}}A\frac{\mathrm{d}x}{\mathrm{d}t} = \tag{10}$$

$$-\|Bx\|^2x^{\mathrm{T}}AAx + (x^{\mathrm{T}}ABx)^2 = \tag{11}$$

$$-\|Bx\|^2\|Ax\|^2 + (x^{\mathrm{T}}ABx)^2 \leqslant \tag{12}$$

$$-\|Bx\|^2\|Ax\|^2 + \|Bx\|^2\|Ax\|^2 = 0, \tag{13}$$

上面的不等式是由Cauchy-Schwartz不等式得来的. 根据式(9) 得到

$$\|B^{\frac{1}{2}}x(t)\|^2 = 1, \tag{14}$$

由此得

$$\|x(t)\| = \|B^{-\frac{1}{2}}B^{\frac{1}{2}}x(t)\| \leqslant \tag{15}$$

$$\|B^{-\frac{1}{2}}\| \cdot \|B^{\frac{1}{2}}x(t)\| \leqslant \tag{16}$$

$$\|B^{-\frac{1}{2}}\|. \tag{17}$$

式(13)说明V沿着网络轨线是能量下降的, 而(17)说明它是有界的, 从而$V(x)$构成了系统(3)的一个Lyapunov函数, 而根据LaSalle不变原理知[10,11], 所有式(3) 的轨线将趋向于下面的集合的最大不变集$\varSigma$:

$$\varSigma \subseteq \{x \mid \frac{\mathrm{d}V}{\mathrm{d}t} = 0\}, \tag{18}$$

但是, Cauchy-Schwartz 等式成立的条件是存在λ使得$Ax = \lambda Bx$, 注意到$x(t)$对集合W是初始的, 从而$x(t)$将趋向于矩阵盒子(A,B)的广义特征向量集合.　　证毕.

4 数值模拟试验(Simulation results)

用下面的例子作数值模拟试验, 设矩阵A是

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \tag{19}$$

矩阵$B = I$, 问题为普通的特征根问题, 此问题有两个特征根$\lambda_1 = 0, \lambda_2 = 5$, 对应的特征向量$\xi_1 = [0.8944 \ \ 0.4472]^{\mathrm{T}}$及$\xi_2 = [-0.4472 \ \ 0.8944]^{\mathrm{T}}$. 此问题的网络模型为

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \begin{pmatrix} -2x_1^2 + 3x_1x_2^2 + 2x_2^3 \\ 2x_1^3 - 3x_1^2x_2 - 2x_1x_2^2 \end{pmatrix}. \tag{20}$$

用ODE23在MATLAB 7.0下求解此问题, 取初值$(0,1)$, 图2给出了寻求ξ_1的网络轨线走势图.

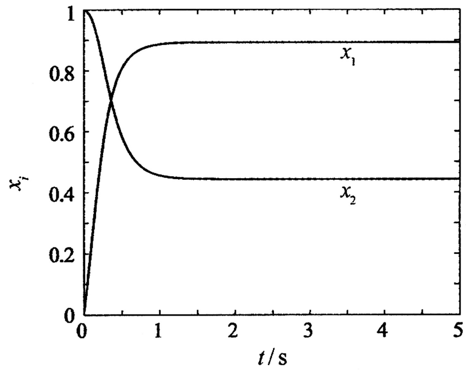

图 2 网络轨线x_1, x_2

Fig. 2 Neural trajectories x_1, x_2

5 结论(Conclusion)

这里给出了求解广义特征根问题的神经网络方法, 证明了网络对其相关的正则化约束集是全局收敛的. 这里的网络模型克服了已有的网络模型的许多缺陷, 包括稳定性问题, 罚函数本身缺陷等.而且网络的全局收敛性也就保证了这个模型是一个稳定运行的、可靠的网络模型. 最后, 数值试验结果也表明这里的网络模型可以很好地求解广义特征根问题.

参考文献(References):

[1] GOLUB G H, VANLOAN C F. *Matrix Computations*[M]. Baltimore: The Johns Hopkins University Press, 1989.

[2] GOLUB G H. Some modified matrix eigenvalue problems[J]. *SIAM Rev* , 1973, 15(1): 318 – 334.

[3] HOPFIELD J J, TANK D W. Neural computation of decisions in optimization problems[J]. *Biolog Cybernetics*, 1985, 52(1): 141 – 152.

[4] TANK D W, HOPFIELD J J. Simple neural optimization networks: an A/D converter, signal decision network, and a linear programming circuit[J]. *IEEE Trans on Circuits and Systems*, 1986, 33(2): 533 – 541.

[5] BOUZERDORM A, PATTISON T R. Neural network for quadratic optimization with bound constraints[J]. *IEEE Trans on Neural Networks*, 1993, 4(2): 293 – 304.

[6] KENNEDY M P, CHAU L O. Neural networks for nonlinear programming[J]. *IEEE Trans on Circuits and Systems*, 1988, 35(6): 554 – 562.

[7] CICHOCKI A, UNBEHAUEN R. *Neural Networks for Optimization and Signal Processing*[M]. New York: John Wiley & Sons, 1993.

[8] CICHOCKI A, UNBEHAUEN R. Neural networks for computing eigenvalues and eigenvectors[J]. *Biolog Cybernetics*, 1992, 68(1): 155 – 164.

[9] BAZARAA M S, SHETTY C M. *Nonlinear Programming, Theory and Algorithms*[M]. New York: John Wiley and Sons, 1979.

[10] LASALLE J. *The Stability of Dynamical Systems*[M]. Philadelphia, PA: SIAM, 1976.

[11] LASALLE J. The stability theory for ordinary differential equations[J]. *J of Differential Equations*, 1983, 4(1): 57 – 65.

作者简介:

冯芙叶 (1963—), 女, 副教授, 空军工程大学博士生, 主要从事运筹学、神经网络等研究, E-mail:zhangquanju@sina.com.cn ;

魏正红 (1964—) , 女, 副教授, 香港浸会大学博士生, 主要从事运筹学、金融数学等研究;

刘付显 (1962—) 男, 教授, 博士生导师, 主要从事军事运筹学、神经网络等研究.

(上接第639页)

[8] 高会军, 王常虹. 不确定离散多时滞系统的时滞相关鲁棒镇定[J]. 自动化学报, 2004, 30(5): 789 – 795.
(GAO Huijun, WANG Changhong. Delay-dependent robust stabilization for uncertain discrete-time systems with multiple state delays[J]. *Acta Automatica Sinica,* 2004, 30(5): 789 – 795.)

[9] de OLIVEIRA M C, BERNUSSOU J, GEROMEL J C. A new discrete-time robust stability condition[J]. *Systems & Control Letters,* 1999, 37(3): 261 – 265.

作者简介:

张文安 (1982—), 男, 研究领域为时滞系统分析与控制、网络控制等, E-mail:zwa_auto@hotmail.com;

俞 立 (1961—), 男, 浙江工业大学教授, 博士生导师, 研究领域为鲁棒控制、网络控制等, E-mail:lyu@zjut.edu.cn;

张贵军 (1973—), 男, 博士，浙江工业大学讲师, 研究领域为优化算法、鲁棒控制等.

(上接第644页)

[10] STECK J K and BALKRISHNAN S N. Use of Hopfield Neural Net works in Optimal Guidance[J]. *IEEE Trans on Aerospace and Electronic Systems,* 1994, 30(1): 287 – 293.

[11] RUAN Xiaogang. Linear Quadratic Dynamic Optimization With Hopfield Network for Discrete-time Systems[C]// *Proc of the 2nd World Congress on Intelligent Control and Automation,* Xian, China, 1997: 1880 – 1883.

[12] 李树荣, 李峰. 基于Hopfield网络的多变量动态矩阵控制及其在精馏塔系统中的应用[J]. 控制理论与应用, 2001, 18(2): 304 – 306.
(LI Shurong, LI Feng. The dynamic matrix control based on Hopfield neural network and its application in distillation colum[J]. *Control Theory and Applications,* 2001, 18(2): 304 – 306.)

[13] 李明爱, 阮晓钢. 基于连续Hopfield网络的最优控制方法[J]. 北京工业大学学报, 2004, 30(1): 27 – 30.
(LI Mingai, RUAN Xiaogang. Optimal Control Based on Continuous Hopfield Neural Nwtwork[J]. *Journal of Beijing University of Technology,* 2004, 30(1): 27 – 30.)

[14] KIM K B. Implementation of stabilizing receding horizon controls for time-varying systems[J].*Automatica,* 2002, 38: 1705 – 1711.

[15] 王永骥, 涂健. 神经元网络控制[M]. 北京：机械工业出版社,1999.
(WANG Yongji, TU Jian. *Neural Network Control*[M]. Beijing: Machinery Industry Press, 1999.)

作者简介:

李明爱 (1966—), 女, 博士, 北京工业大学电子信息与控制工程学院副教授, 主要研究方向为智能控制理论、神经网络控制等, E-mail:limingai@bjut.edu.cn;

乔俊飞 (1968—), 男, 博士, 北京工业大学电子信息与控制工程学院教授, 主要研究方向为非线性系统理论、智能控制等, E-mail:junfeiq@bjut.edu.cn;

阮晓钢 (1958—), 男, 博士生导师,北京工业大学电子信息与控制工程学院教授, 本文通讯作者, 主要研究方向为智能控制理论、人工智能、智能信息处理等, E-mail:adrxg@bjut.edu.cn.