

引用格式: 雍龙泉. 神经网络方法求解绝对值方程及线性互补[J]. 陕西理工大学学报(自然科学版), 2020, 36(5): 72-81.

# 神经网络方法求解绝对值方程及线性互补

雍龙泉<sup>1,2</sup>

(1. 陕西理工大学 数学与计算机科学学院, 陕西 汉中 723000;

2. 陕西省工业自动化重点实验室, 陕西 汉中 723000)

**摘 要:** 首先给出了绝对值函数的 3 个光滑逼近函数, 分析了这些光滑逼近函数的性质; 然后选取性质较好的光滑函数来处理绝对值方程, 得到一个可微的无约束优化问题; 建立了求解无约束优化问题的梯度下降神经网络模型。通过求解唯一解、多个解的绝对值方程, 结果表明该方法不依赖初始点, 且具有收敛快等优点。最后把该方法应用于求解线性互补问题。

**关键词:** 绝对值函数; 光滑逼近函数; 绝对值方程; 无约束优化; 梯度下降神经网络; 线性互补

**中图分类号:** O221; TP183 **文献标识码:** A **文章编号:** 2096-3998(2020)05-0072-10

绝对值方程(Absolute Value Equations, AVE)是指:

$$Ax - |x| = b, \quad (1)$$

其中  $A \in \mathbf{R}^{n \times n}$ ,  $x, b \in \mathbf{R}^n$ ,  $|x|$  表示对  $x$  的各个分量取绝对值。绝对值方程的研究来源于两个方面, 一个是区间线性方程, 另一个是线性互补问题。由于任意的线性互补都能转化为绝对值方程, 而绝对值方程具有简单的结构, 因此关于绝对值方程的研究引起了众多学者的关注<sup>[1-2]</sup>。其中 Rohn<sup>[3]</sup> 在理论方面做出了奠基性的工作, Mangasarian<sup>[4-7]</sup> 在理论和算法方面做出了巨大的贡献, 指出绝对值方程等价于双线性规划, 并给出了相应的求解算法。

近年来, 针对唯一解的绝对值方程, 求解算法主要有 5 类: (1) 通过把绝对值方程转化为线性互补问题进行求解<sup>[1]</sup>; (2) 把绝对值方程转化为一个非光滑方程, 采用广义牛顿法或非光滑牛顿法进行求解<sup>[7-10]</sup>; (3) 通过光滑处理, 绝对值方程转化为一个光滑优化问题, 采用光滑牛顿法求解<sup>[11-18]</sup>; (4) 采用矩阵分裂技术, 求解某些具有特殊结构的绝对值方程<sup>[19-24]</sup>; (5) 采用神经网络方法进行求解<sup>[25-30]</sup>。此外, 智能优化算法近年来也用于求解绝对值方程<sup>[31]</sup>。针对多个解的绝对值方程, 可以借助聚类技术, 找到尽可能多的解, 该方面研究较多的是具有  $2^n$  个解的绝对值方程<sup>[32-34]</sup>。此外, 针对多个解的绝对值方程, 稀疏解的研究目前也成为了一个热点<sup>[35-37]</sup>。

本文选取逼近性质较好且不易发生溢出的一致光滑逼近函数来处理绝对值方程, 建立求解无约束优化问题的梯度下降神经网络模型, 并应用于求解绝对值方程和线性互补问题。

文中  $I$  表示单位矩阵,  $\|\cdot\|$  表示 2 范数。

## 1 绝对值函数的一致光滑逼近函数

**定义 1(光滑逼近函数)** 给定非光滑函数  $f(t): \mathbf{R} \rightarrow \mathbf{R}$ , 我们称光滑函数  $f_\mu(t)$ ,  $\mu > 0$  为  $f(t)$  的光滑

收稿日期: 2020-04-16 修回日期: 2020-05-03

基金项目: 国家自然科学基金资助项目(11401357); 陕西省青年科技新星项目(2016KJXX-95); 陕西省教育厅科研基金资助项目(20JS021); 陕西理工大学科研基金资助项目(SLGYQZX2002)

作者简介: 雍龙泉(1980—), 男, 陕西省洋县人, 博士, 教授, 主要研究方向为最优化理论与算法、智能优化算法。

逼近函数, 如果  $\forall t \in \mathbf{R}$ , 存在  $\kappa > 0$ , 使得

$$|f_{\mu}(t) - f(t)| \leq \kappa\mu, \quad \forall \mu > 0,$$

如果  $\kappa$  不依赖于  $t$ , 则称  $f_{\mu}(t)$  为  $f(t)$  的一致光滑逼近函数<sup>[18-38-40]</sup>。

一致光滑逼近函数可以分为从上方一致逼近和从下方一致逼近。文献[38-39]给出了绝对值函数  $\phi(t) = |t|$  的多个一致光滑逼近函数, 为了便于研究, 本文选取如下3个光滑函数:

$$\phi_{\mu}^1(t) = \mu [\ln(1 + e^{-\frac{t}{\mu}}) + \ln(1 + e^{\frac{t}{\mu}})],$$

$$\phi_{\mu}^2(t) = \sqrt{t^2 + \mu^2},$$

$$\phi_{\mu}^3(t) = \mu \ln(e^{\frac{t}{\mu}} + e^{-\frac{t}{\mu}}).$$

下面给出这3个一致光滑逼近函数的性质<sup>[38-39]</sup>。

定理1  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 分别按照上面的定义, 则  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 满足:

(1)  $0 < \phi_{\mu}^1(t) - \phi(t) \leq \mu \ln 4$ ,  $0 < \phi_{\mu}^2(t) - \phi(t) \leq \mu$ ,  $0 < \phi_{\mu}^3(t) - \phi(t) \leq \mu \ln 2$ , 因此  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 在  $\mu > 0, t \in \mathbf{R}$  上是  $\phi(t)$  的上方一致光滑逼近函数;

(2)  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 在  $\mu > 0, t \in \mathbf{R}$  上是连续可微的, 且都满足

$$\left| \frac{d\phi_{\mu}^i(t)}{dt} \right| < 1, \quad \left. \frac{d\phi_{\mu}^i(t)}{dt} \right|_{t=0} = 0, \quad i=1, 2, 3;$$

(3)  $\forall t \in \mathbf{R}, \lim_{\mu \rightarrow 0^+} \phi_{\mu}^i(t) = \phi(t) = |t|$ ,  $i=1, 2, 3$ 。

图1给出了  $\mu=0.4, 0.2$  时  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 与  $\phi(t) = |t|$  的图像。

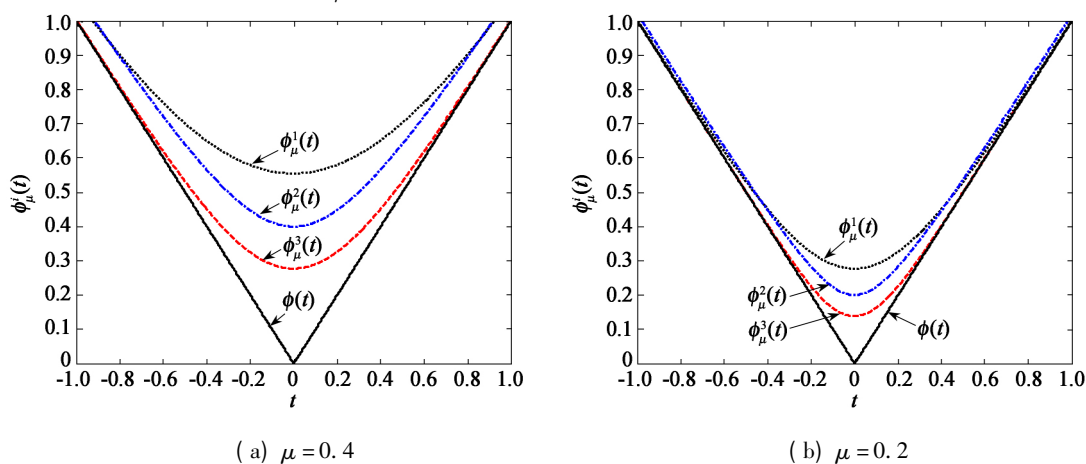


图1  $\phi_{\mu}^i(t)$  ( $i=1, 2, 3$ ) 与  $\phi(t) = |t|$  的图像

结合定理1可知, 在理论上  $\phi_{\mu}^3(t)$  的逼近程度优于  $\phi_{\mu}^2(t)$  与  $\phi_{\mu}^1(t)$ 。因此文献[25-27]利用  $\phi_{\mu}^3(t)$  光滑处理绝对值方程, 之后采用梯度下降神经网络方法求解, 得到结论: 选取  $\phi_{\mu}^3(t)$  作为光滑函数时, 计算时间和精度都比选取  $\phi_{\mu}^2(t)$  作为光滑函数好。

但实际计算时  $\phi_{\mu}^1(t)$  与  $\phi_{\mu}^3(t)$  中含有指数项, 因此采用 MATLAB 实现算法时, 易发生溢出现象(任何软件都会产生溢出)。比如取  $t=8, \mu=0.01$ , 则  $\phi_{\mu}^1(t)$  与  $\phi_{\mu}^3(t)$  中  $e^{\frac{t}{\mu}}$  就产生溢出。事实上, MATLAB 软件中双精度实数的范围为  $[-1.7977 \times 10^{308}, 1.7977 \times 10^{308}]$ , 而  $e^{709} = 8.2184 \times 10^{307}$ , 故  $e^{710}, e^{800}$  都会溢出(输出 inf)。为了防止溢出现象的发生, 本文还是选取函数  $\phi_{\mu}^2(t)$  来光滑化处理绝对值方程。为了书写方便, 记  $\phi_{\mu}^2(t)$  为  $\phi_{\mu}(t)$ 。

## 2 绝对值方程的光滑逼近处理

引理1<sup>[1]</sup> 若  $A \in \mathbf{R}^{n \times n}$ , 且矩阵  $A$  的所有奇异值都大于1, 则矩阵  $A$  可逆, 且  $\|A^{-1}\| < 1$ 。

引理2<sup>[1]</sup> (i) 若  $A \in \mathbf{R}^{n \times n}$ , 且矩阵  $A$  的所有奇异值都大于1, 则  $\forall b \in \mathbf{R}^n$ , AVE 存在唯一解;

(ii) 若  $A \in \mathbf{R}^{n \times n}$ , 且矩阵  $A$  满足  $\|A^{-1}\| < 1$ , 则  $\forall b \in \mathbf{R}^n$ , AVE 存在唯一解。

引理3<sup>[16]</sup> 若  $A \in \mathbf{R}^{n \times n}$  且矩阵  $A$  满足  $\|A^{-1}\| < 1$  矩阵  $D = \text{diag}(d_1, d_2, \dots, d_n)$  这里  $|d_i| \leq 1, i = 1, 2, \dots, n$  则  $A \pm D$  非奇异。

问题(1)等价于非线性方程组

$$H(x) = 0, \quad (2)$$

其中  $H(x) := Ax - |x| - b$ 。由于  $H: \mathbf{R}^n \rightarrow \mathbf{R}^n$  是非光滑函数,以下构造一个光滑函数  $H_\mu(x)$  来逼近非光滑函数  $H(x)$ 。

定义2<sup>[17]</sup> 给定非光滑函数  $H: \mathbf{R}^n \rightarrow \mathbf{R}^n$  称光滑函数  $H_\mu: \mathbf{R}^n \rightarrow \mathbf{R}^n (\mu > 0)$  为  $H$  的光滑逼近函数 如果  $\forall x \in \mathbf{R}^n$  存在  $\kappa > 0$  使得

$$\|H_\mu(x) - H(x)\| \leq \kappa\mu, \quad \forall \mu > 0,$$

如果  $\kappa$  不依赖于  $x$  则称  $H_\mu$  为  $H$  的一致光滑逼近函数。

记向量绝对值函数  $\phi(x) = |x| = (|x_1|, |x_2|, \dots, |x_n|)^T$  的每一个分量为

$$\phi(x_i) := |x_i|, \quad i = 1, 2, \dots, n,$$

对每一个  $\phi(x_i)$  采用定理1中的光滑函数

$$\phi_\mu(x_i) = \sqrt{x_i^2 + \mu^2}, \quad i = 1, 2, \dots, n,$$

进行光滑化处理 就得到向量绝对值函数  $\phi(x)$  的光滑逼近函数为

$$\phi_\mu(x) = (\phi_\mu(x_1), \phi_\mu(x_2), \dots, \phi_\mu(x_n))^T,$$

且每一个分量  $\phi_\mu(x_i)$  满足

$$0 < \phi_\mu(x_i) - \phi(x_i) < \mu, \quad i = 1, 2, \dots, n.$$

从而有

$$\|\phi_\mu(x) - \phi(x)\| = \sqrt{\sum_{i=1}^n |\phi_\mu(x_i) - \phi(x_i)|^2} \leq \sqrt{n}\mu,$$

于是当  $\mu \rightarrow 0$  时  $\phi_\mu(x)$  一致光滑逼近绝对值函数  $\phi(x)$ 。

通过上面的光滑处理,问题(2)转化为如下光滑方程组:

$$H_\mu(x) = 0, \quad (3)$$

其中  $H_\mu(x) = Ax - \phi_\mu(x) - b$   $\phi_\mu(x) = (\phi_\mu(x_1), \phi_\mu(x_2), \dots, \phi_\mu(x_n))^T$ 。

下面研究光滑函数  $H_\mu(x)$  的一些性质。

定理2  $\forall \mu > 0$   $H_\mu(x)$  一致光滑逼近  $H(x)$ 。

证明  $\forall \mu > 0$  由于  $H_\mu(x)$  可微 且满足

$$\|H_\mu(x) - H(x)\| = \|\phi_\mu(x) - \phi(x)\| = \sqrt{\sum_{i=1}^n |\phi_\mu(x_i) - \phi(x_i)|^2} \leq \sqrt{n}\mu,$$

则  $H_\mu(x)$  一致光滑逼近  $H(x)$ 。

记  $H_\mu(x)$  的 Jacobian 矩阵为  $H'_\mu(x)$  简单运算可得

$$H'_\mu(x) = A - D_\mu(x),$$

其中  $D_\mu(x) = \text{diag}\left(\frac{x_1}{\sqrt{x_1^2 + \mu^2}}, \frac{x_2}{\sqrt{x_2^2 + \mu^2}}, \dots, \frac{x_n}{\sqrt{x_n^2 + \mu^2}}\right)$ 。

定理3  $\forall \mu > 0$   $A \in \mathbf{R}^{n \times n}$  若  $\|A^{-1}\| < 1$  则  $H_\mu(x)$  的 Jacobian 矩阵  $H'_\mu(x)$  非奇异。

证明 利用定理1中  $D_\mu(x_i) = \left|\frac{x_i}{\sqrt{x_i^2 + \mu^2}}\right| < 1, i = 1, 2, \dots, n$  结合引理3可知  $H'_\mu(x)$  非奇异。

### 3 梯度下降神经网络模型

定义函数

$$E_\mu(x) = \frac{1}{2} \|H_\mu(x)\|^2, \quad (4)$$

称  $E_\mu(x)$  为能量函数。到此, 求解绝对值方程的近似解已经转化为求解连续可微的优化问题  $\min E_\mu(x)$ 。

定理4  $\forall \mu > 0$ , 矩阵  $A$  满足  $\|A^{-1}\| < 1$ , 若  $\nabla E_\mu(x) = 0$ , 则  $E_\mu(x) = 0$ 。

证明  $\forall \mu > 0$ , 由于

$$\nabla E_\mu(x) = [H'_\mu(x)]^T H_\mu(x),$$

利用定理3可知  $H'_\mu(x)$  非奇异, 因此若  $\nabla E_\mu(x) = 0$ , 则  $H_\mu(x) = 0$ , 从而就有  $E_\mu(x) = 0$ 。

梯度神经网络是基于求解问题(4)的最速下降模型的连续化形式, 近年来已广泛应用于求解非线性互补、二阶锥规划等问题<sup>[41-45]</sup>。

下面给出求解光滑优化问题(4)的梯度下降神经网络模型:

$$\frac{dx}{dt} = -\tau \nabla E_\mu(x) = -\tau [H'_\mu(x)]^T H_\mu(x), \quad \tau > 0, \quad (5)$$

这里参数  $\tau$  表示梯度下降算法的步长。本文直接给出该神经网络的收敛性结果, 详细的证明见文献[25-27]。

定理5  $\forall \mu > 0$ , 矩阵  $A$  满足  $\|A^{-1}\| < 1$ , 神经网络(5)的平衡点是绝对值方程的解, 反之, 绝对值方程的解是神经网络(5)的平衡点。

下面采用模型(5)来计算一些常见的绝对值方程, 以验证算法的有效性。

## 4 数值实验

给出3个绝对值方程(前2个具有唯一解, 最后一个有4个解), 通过将其转化为神经网络模型(5), 采用四阶 Runge-Kutta 法求解微分方程组(对应 MATLAB 中的函数 ode45), 程序采用 MATLAB R2009a 编写。设置  $\mu = 0.01$ , 终止条件  $E_\mu(x) \leq 1.0 \times 10^{-10}$ 。

初始点取为  $x^{(0)} = 100 \times \text{rand}(n, 1) - 100 \times \text{rand}(n, 1) \in [-100, 100]^n$ , 即初始点在任一象(卦)限随机选取; 这对于具有多个解的绝对值方程, 多次运行就有可能找到尽可能多的解。由于  $x_i \in [-100, 100]$ , 故  $x_i/\mu \in [-1.0 \times 10^4, 1.0 \times 10^4]$ ,  $i = 1, 2, \dots, n$ , 如果光滑函数采用  $\phi_\mu^1(t)$  与  $\phi_\mu^3(t)$ , 则会发生溢出现象而无法运行。采用光滑函数  $\phi_\mu^2(t)$  就避免了溢出现象的发生。

算例4.1 考虑绝对值方程(1), 其中

$$A = \begin{bmatrix} 10 & 1 & 2 & 0 \\ 1 & 11 & 3 & 1 \\ 0 & 2 & 12 & 1 \\ 1 & 7 & 0 & 13 \end{bmatrix}, \quad b = \begin{bmatrix} 12 \\ 15 \\ 14 \\ 20 \end{bmatrix},$$

由于矩阵  $A$  的奇异值  $(\text{SVD}(A) = [17.4349, 12.2629, 9.6389, 7.5984])$  大于1, 因此该 AVE 问题存在唯一解  $x^* = (1, 1, 1, 1)^T$ 。表1给出了参数  $\tau$  取不同值的计算结果; 图2和图3分别给出了  $\tau = 100$  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。

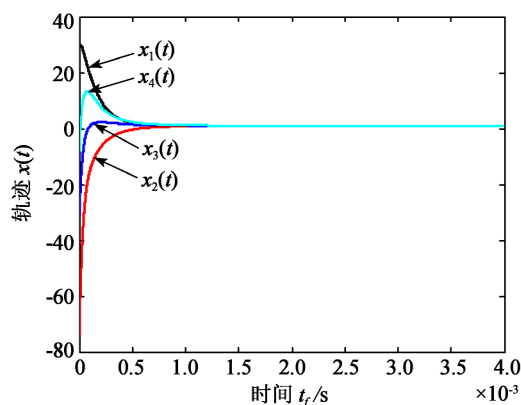


图2 近似解随时间的变化曲线

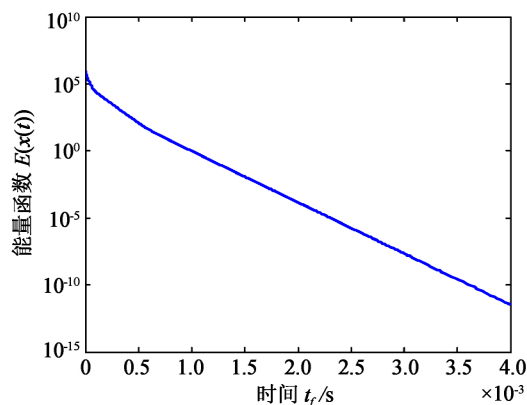


图3 能量函数随时间的变化曲线

算例 4.2 绝对值方程 (1) 中所有奇异值大于 1, 矩阵  $A$  由如下 MATLAB 命令产生

```
rand('state', 0);
R = rand(n, n);
A = R' * R + n * eye(n);
b = (A - eye(n)) * ones(n, 1)。
```

说明: 给出矩阵维数  $n$  后, 可以用上述代码产生和本文相同的数据, 该 AVE 具有唯一解  $x^* = (1, 1, \dots, 1)^T$ 。取不同的维数  $n$ , 表 1 给出了参数  $\tau$  取不同值的计算结果。图 4 给出了  $\tau = 100$  时维数  $n$  取 10 和 50 的能量函数随时间的变化曲线。

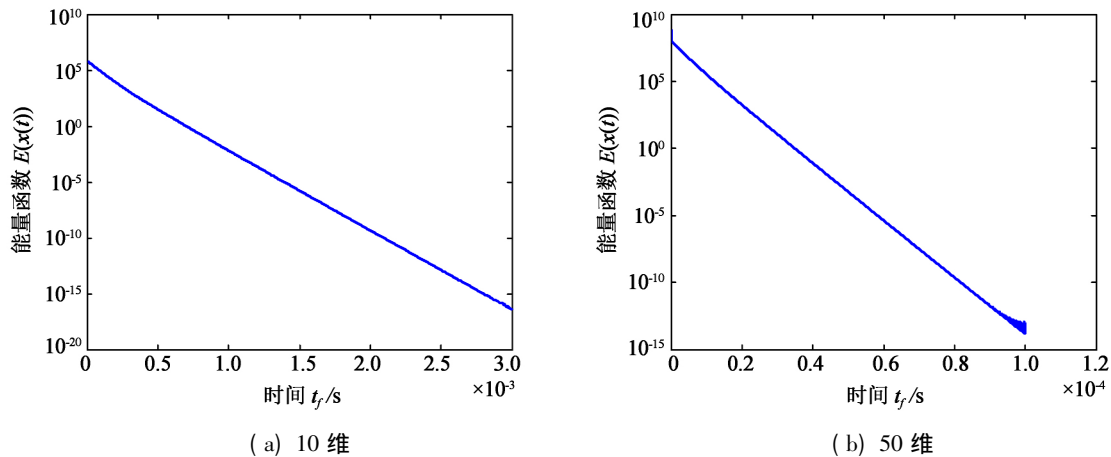


图 4 能量函数随时间的变化曲线

表 1 参数  $\tau$  取不同值的计算结果

	参数 $\tau$ 值	$t_f$	$\ x(t_f) - x^*\ $	$E(x(t_f))$
算例 4.1 (唯一解)	$\tau = 1$	0.36	$7.241 \times 10^{-6}$	$7.132 \times 10^{-11}$
	$\tau = 10$	0.04	$6.876 \times 10^{-6}$	$1.425 \times 10^{-11}$
	$\tau = 100$	0.004	$6.875 \times 10^{-6}$	$1.025 \times 10^{-12}$
算例 4.2 ( $n=10$ ) (唯一解)	$\tau = 1$	0.21	$5.157 \times 10^{-6}$	$8.938 \times 10^{-11}$
	$\tau = 10$	0.03	$5.187 \times 10^{-6}$	$3.834 \times 10^{-17}$
	$\tau = 100$	0.003	$5.187 \times 10^{-6}$	$3.828 \times 10^{-17}$
算例 4.2 ( $n=50$ ) (唯一解)	$\tau = 1$	0.009	$7.221 \times 10^{-7}$	$1.957 \times 10^{-12}$
	$\tau = 10$	0.001	$7.276 \times 10^{-7}$	$1.820 \times 10^{-14}$
	$\tau = 100$	0.000 1	$7.276 \times 10^{-7}$	$1.826 \times 10^{-14}$

算例 4.3 考虑绝对值方程 (1) 其中

$$A = \begin{bmatrix} 0.1 & 0.02 \\ 0.2 & 0.01 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ -2 \end{bmatrix},$$

由于  $\gamma = \frac{\min_i |b_i|}{\max_i |b_i|} = 0.5$ ,  $\|A\|_\infty = 0.21 < \frac{\gamma}{2}$ , 所以该 AVE 具有  $2^2$  个解, 这  $2^2$  个解为  $x^i = D_i(AD_i - I)^{-1}b$ ,  $i = 1, 2, 3, 4$ 。其中

$$D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad D_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_4 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}。$$

简单计算可得

$$x^1 = \begin{bmatrix} 1.1612 \\ 2.2548 \end{bmatrix}, \quad x^2 = \begin{bmatrix} 1.0624 \\ -2.1906 \end{bmatrix}, \quad x^3 = \begin{bmatrix} -0.9424 \\ 1.8298 \end{bmatrix}, \quad x^4 = \begin{bmatrix} -0.8762 \\ -1.8067 \end{bmatrix}。$$

图 5 给出了  $\tau = 100$  时, 初始点在不同象限时的近似解随时间的变化曲线。

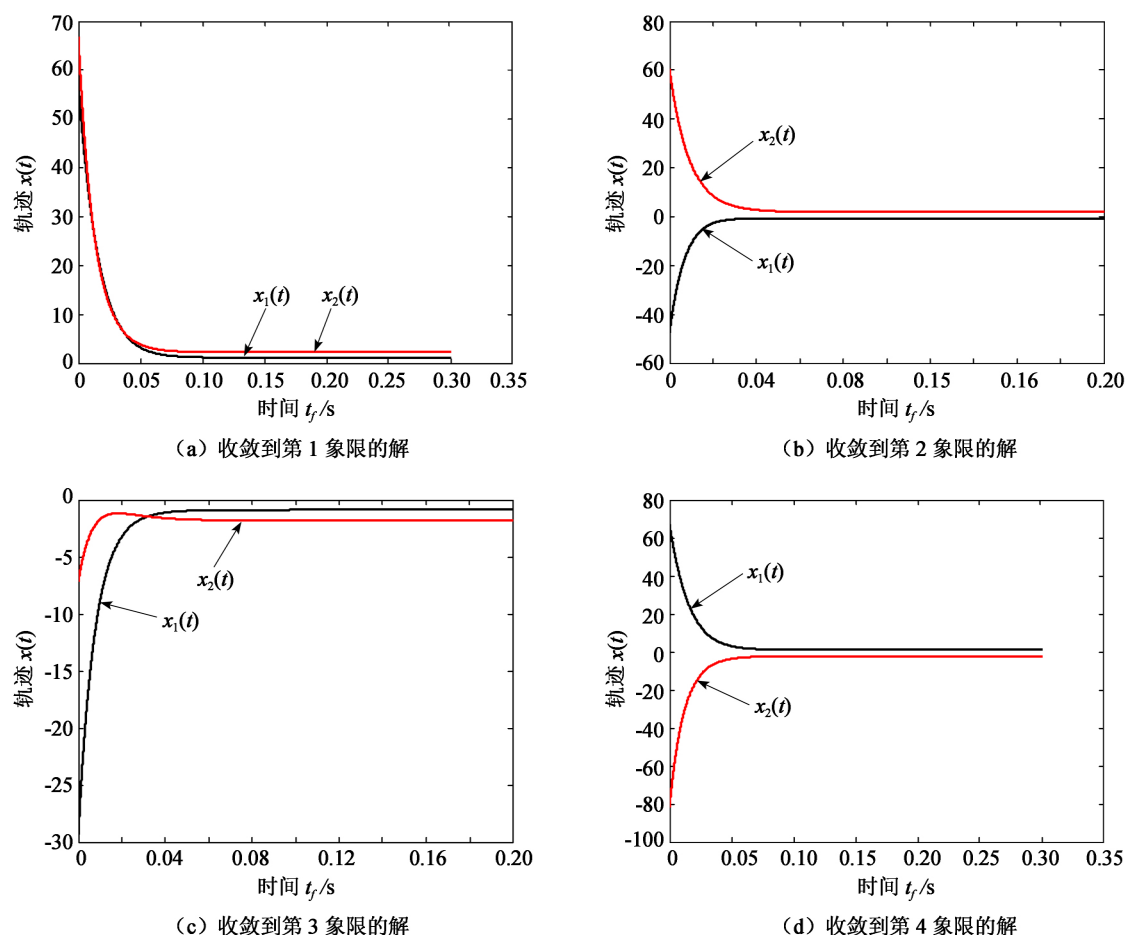


图5 初始点在不同象限时的收敛曲线

从图5可以看出,由于该问题的解分布在4个象限,初始点在哪个象限,则算法收敛到相应象限的解。

## 5 应用于线性互补问题

线性互补问题就是求一个向量  $z \in \mathbf{R}^n$ ,使得  $z^T(Mz + q) = 0$ ,  $z \geq 0$ ,  $Mz + q \geq 0$ 。这里  $M \in \mathbf{R}^{n \times n}$ ,  $q \in \mathbf{R}^n$ 。线性互补问题简记为  $LCP(M, q)$ 。文献[2]中详细研究了 AVE 与  $LCP(M, q)$  的转化,指出若 1 不是矩阵  $M$  的特征值,则  $LCP(M, q)$  等价于绝对值方程  $(M + I)(M - I)^{-1}x - |x| = ((M + I)(M - I)^{-1} - I)q$ ,这里  $x = (M - I)z + q$ 。

下面把上述方法应用于求解线性互补问题,为了便于计算,取初始点  $x^{(0)} = (0, 0, \dots, 0)^T$ ,设置  $\mu = 0.01$ ,终止条件  $E_\mu(x) \leq 1.0 \times 10^{-10}$ 。

算例 5.1 考虑线性互补问题  $LCP(M, q)$ ,其中

$$M = \begin{bmatrix} 1 & -4 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} -5 \\ -5 \\ 1 \\ 1 \end{bmatrix},$$

$\forall z$ , 这里  $z = (z_1, z_2, z_3, z_4)^T$ 。由于  $z^T M z = (z_1 - 2z_2)^2 - 3z_2^2$ ,令  $\tilde{z} = (2, 1, 0, 0)^T$ ,则有  $\tilde{z}^T M \tilde{z} = -3 < 0$ 。虽然矩阵  $M$  不是正定矩阵<sup>[46]</sup>,但是该问题具有唯一解  $z^* = (1, 1, 8, 4)^T$ 。

由于 1 不是矩阵  $M$  特征值,因此线性互补可以转化为绝对值方程  $Ax - |x| = b$ ,这里  $A = (M + I)(M - I)^{-1}$ ,  $b = ((M + I)(M - I)^{-1} - I)q$ ;简单计算可得

$$A = \begin{bmatrix} -1 & 8 & -2 & 8 \\ 0 & -1 & 0 & -2 \\ 2 & -8 & 1 & -8 \\ 0 & 2 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -24 \\ 8 \\ 22 \\ -10 \end{bmatrix}, \quad x^* = \begin{bmatrix} -1 \\ -1 \\ -8 \\ -4 \end{bmatrix}.$$

图6和图7给出了 $\tau=100$ 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。

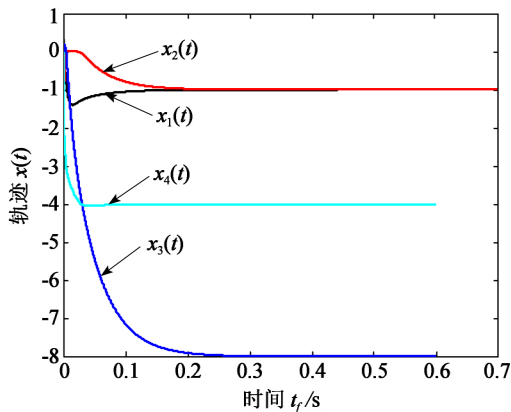


图6 近似解随时间的变化曲线

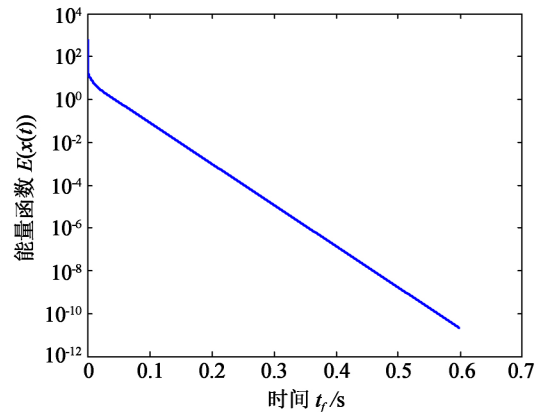


图7 能量函数随时间的变化曲线

算例 5.2 考虑线性互补问题  $LCP(M, q)$  其中

$$M = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & -1 & -2 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} -8 \\ -6 \\ -4 \\ 3 \end{bmatrix},$$

$\forall z$  这里  $z = (z_1, z_2, z_3, z_4)^T$ 。由于

$$z^T M z = 2z_1^2 + 2z_1z_2 + 2z_1z_3 + 2z_2^2 + z_3^2 = (z_1 + z_2)^2 + z_2^2 + (z_1 + z_3)^2 \geq 0,$$

这里  $z_4$  前面的系数为 0, 因此矩阵  $M$  是半正定矩阵<sup>[46]</sup>; 且该线性互补问题具有唯一解  $z^* = (2.5, 0.5, 0.25)^T$ 。由于 1 是矩阵  $M$  的特征值, 因此令  $\lambda = 3$ , 则  $\lambda M$  的特征值就不为 1, 且  $LCP(\lambda M, \lambda q)$  与  $LCP(M, q)$  具有共同最优解  $z^*$ , 而  $LCP(\lambda M, \lambda q)$  就可以转化为绝对值方程  $Ax - |x| = b$ , 这里  $A = (M + I)(M - I)^{-1}$ ,  $b = ((M + I)(M - I)^{-1} - I)q$ ; 简单计算可得

$$A = \begin{bmatrix} 2.0330 & -0.3626 & -0.2637 & 0.4286 \\ -0.3626 & 1.4890 & -0.0989 & -0.2143 \\ -0.2637 & -0.0989 & 1.1099 & -0.4286 \\ -0.4286 & 0.2143 & 0.4286 & 0.9286 \end{bmatrix}, \quad b = \begin{bmatrix} -11.2418 \\ -0.8407 \\ 2.9341 \\ 0.6429 \end{bmatrix}, \quad x^* = \begin{bmatrix} -2.5000 \\ -0.5000 \\ 10.5000 \\ -2.5000 \end{bmatrix}.$$

图8和图9给出了 $\tau=100$ 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。

算例 5.3 考虑线性互补问题  $LCP(M, q)$  其中

$$M = \begin{bmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad q = \begin{bmatrix} -1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix},$$

矩阵  $M$  是半正定矩阵<sup>[47]</sup>, 且该线性互补问题具有唯一解  $z^* = (0, 0, \dots, 0, 1)^T$ 。

取  $n=4$ , 由于 1 是矩阵  $M$  的特征值, 因此令  $\lambda = 3$ , 则  $\lambda M$  的特征值就不为 1, 且  $LCP(\lambda M, \lambda q)$  与  $LCP(M, q)$  具有共同最优解  $z^*$ , 而  $LCP(\lambda M, \lambda q)$  就可以转化为绝对值方程  $Ax - |x| = b$ , 这里  $A = (M + I)(M - I)^{-1}$ ,  $b = ((M + I)(M - I)^{-1} - I)q$ ; 简单计算可得

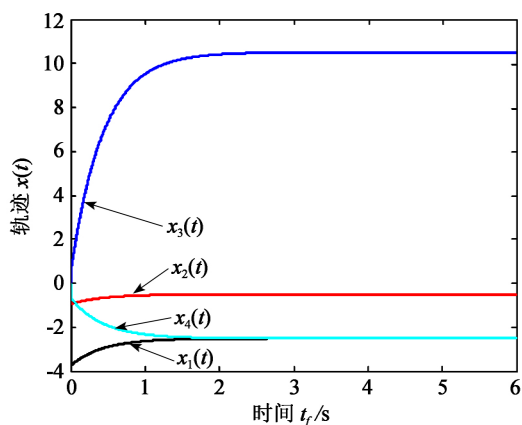


图8 近似解随时间的变化曲线

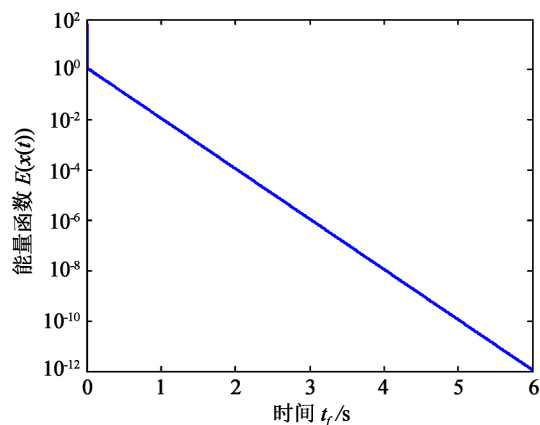


图9 能量函数随时间的变化曲线

$$A = \begin{bmatrix} 2 & -3 & 6 & -12 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 24 \\ -12 \\ 6 \\ -3 \end{bmatrix}, \quad x^* = \begin{bmatrix} 3 \\ 3 \\ 3 \\ -1 \end{bmatrix}.$$

图10和图11给出了 $\tau=100$ 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。

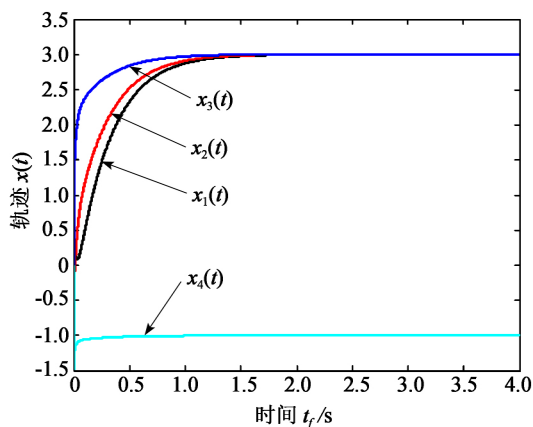


图10 近似解随时间的变化曲线

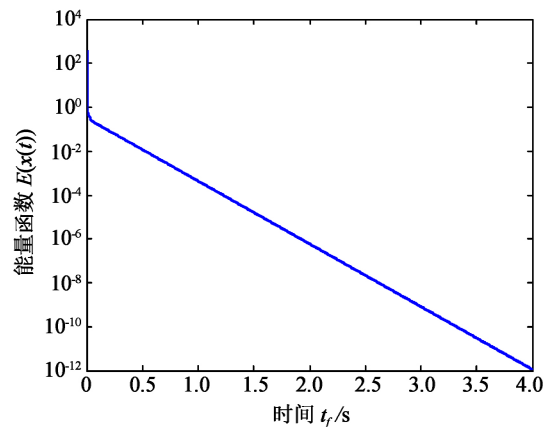


图11 能量函数随时间的变化曲线

## 6 结束语

本文选取逼近性质较好且不易发生溢出的光滑逼近函数来处理绝对值方程,通过构造梯度下降神经网络模型求解无约束优化问题而得到绝对值方程的近似解。结果表明该方法不依赖初始点,且具有收敛快等优点。文献[39]中系统地给出了绝对值函数的上方和下方一致光滑逼近函数,下一步可以通过改变光滑逼近函数,以期获得更好的计算结果。

## [参考文献]

- [1] 雍龙泉. 绝对值方程的算法研究[D]. 西安: 西安电子科技大学, 2014.
- [2] 雍龙泉, 刘三阳, 拓守恒, 等. 线性互补问题与绝对值方程的转化[J]. 吉林大学学报(理学版), 2014, 52(4): 682-686.
- [3] ROHN J. A Theorem of the Alternatives for the Equation  $Ax + B|x| = b$  [J]. Linear and Multilinear Algebra, 2004, 52(6): 421-426.
- [4] MANGASARIAN O L, MEYER R R. Absolute value equations [J]. Linear Algebra and its Applications, 2006(5): 359-367.
- [5] MANGASARIAN O L. Absolute value programming [J]. Computational Optimization and Applications, 2007, 36(1): 43-53.



- [6] MANGASARIAN O L. Absolute value equation solution via concave minimization[J]. Optimization Letters 2007 1(1):3-8.
- [7] MANGASARIAN O L. A generalized Newton method for absolute value equations[J]. Optimization Letters 2009 3(3):101-108.
- [8] 魏庆举. 绝对值方程的广义牛顿算法及其收敛性[D]. 北京: 北京交通大学 2009.
- [9] LI Cui-xia. A Modified Generalized Newton Method for Absolute Value Equations[J]. Journal of Optimization Theory and Applications 2016(3):1055-1059.
- [10] BELLO CRUZ J Y, FERREIRA O P, PRUDENTE L F. On the global convergence of the inexact semi-smooth Newton method for absolute value equation[J]. Computational Optimization and Applications 2016 65(1):93-108.
- [11] 高竹峰. 求解绝对值方程组的光滑牛顿法[D]. 天津: 天津大学 2009.
- [12] CACCETTA L, QU Biao, ZHOU Guang-lu. A globally and quadratically convergent method for absolute value equations[J]. Computational Optimization and Applications 2011 48:45-58.
- [13] 雍龙泉, 拓守恒. 基于凝聚函数的拟牛顿算法求解绝对值方程[J]. 系统科学与数学 2012 32(11):1427-1436.
- [14] JIANG Xiao-qin, ZHANG Ying. A smoothing-type algorithms for absolute value equations[J]. Journal of Industrial and Management Optimization 2013 9(4):789-798.
- [15] ESMAEILI H, MAHMOODABADI E, AHMADI M. A uniform approximation method to solve absolute value equation[J]. Bulletin of the Iranian Mathematical Society 2015 41(5):1259-1269.
- [16] 雍龙泉. 基于调节熵函数的光滑牛顿法求解绝对值方程[J]. 兰州大学学报(自然科学版) 2016 52(4):540-544.
- [17] SAHEYA B, YU Cheng-he, CHEN Jin-shan. Numerical comparisons based on four smoothing functions for absolute value equation[J]. Journal of Applied Mathematics and Computing 2018 56(1/2):131-149.
- [18] 雍龙泉. 上方一致光滑逼近函数及其在绝对值方程中的应用[J]. 数学的实践与认识 2018 48(13):236-241.
- [19] 雍龙泉. 迭代法求解实对称矩阵绝对值方程[J]. 西南大学学报(自然科学版) 2012 34(5):32-37.
- [20] SALKUYEH D K. The Picard-HSS iteration method for absolute value equations[J]. Optimization Letters 2014 8(8):2191-2202.
- [21] ROHN J, HOOSHYARBAKHS V, FARHADSEFAT R. An iterative method for solving absolute value equations and sufficient conditions for unique solvability[J]. Optimization Letters 2014 8(1):35-44.
- [22] HAGHANI K F. On Generalized Traub's Method for Absolute Value Equations[J]. Journal of Optimization Theory and Applications 2015(2):619-625.
- [23] WU Shi-Liang, GUO Peng. On the Unique Solvability of the Absolute Value Equation[J]. Journal of Optimization Theory and Applications 2016(2):705-712.
- [24] GU Xian-ming, HUANG Ting-zhu, WANG Sheng-feng et al. Two CSCS-based iteration methods for solving absolute value equations[J]. Mathematics 2017 7(4):1336-1356.
- [25] WANG Fei-ran, YU Zhen-sheng, GAO Chang. A Smoothing Neural Network Algorithm for Absolute Value Equations[J]. Engineering 2015 7(9):567-576.
- [26] 王斐然, 耿秀荣. 最速下降神经网络算法求解广义绝对值方程[J]. 桂林航天工业学院学报 2017(1):48-55.
- [27] 张晓敏. 神经网络算法求解绝对值方程的解及其稀疏解[D]. 郑州: 郑州大学 2017.
- [28] HUANG Xin-jian, CUI Bao-tong. Neural Network-Based Methodology for Solving Absolute Value Equations[J]. ICIC Express Letters 2017 11(4):853-861.
- [29] MANSOORI A, ESHAGHNEZHAD M, EFFATI S. An Efficient Neural Network Model for Solving the Absolute Value Equations[J]. IEEE Transactions on Circuits and Systems II: Express Briefs 2018 65(3):391-395.
- [30] SAHEYA B, NGUYEN C T, CHEN Jin-shan. Neural network based on systematically generated smoothing functions for absolute value equation[J]. Journal of Applied Mathematics and Computing 2019 61:533-558.
- [31] 雍龙泉. 一种全局和声搜索算法求解绝对值方程[J]. 计算机应用研究 2013 30(11):3276-3279.
- [32] 雍龙泉, 刘三阳, 拓守恒, 等. 具有  $2^n$  个解的绝对值方程问题[J]. 吉林大学学报(理学版) 2013 51(3):383-388.
- [33] 雍龙泉, 刘三阳, 拓守恒, 等. 全局和声搜索算法求解具有  $2^n$  个解的绝对值方程[J]. 小型微型计算机系统 2014, 35(8):1861-1864.
- [34] YONG Long-quan, LIU San-yang, TUO Shou-heng. A New Method for Absolute Value Equation with  $2^n$  Distinct Solutions[J]. ICIC Express Letters 2014 8(7):1917-1922.
- [35] 雍龙泉. 绝对值方程算例分析[J]. 数学的实践与认识 2016 46(5):286-290.

- [36] LIU Xiao-hong ,FAN Jie ,LI Wen-juan. Concave minimization for sparse solutions of absolute value equations [J]. Transactions of Tianjin University 2016 22( 1) : 89-94.
- [37] YONG Long-quan ,TUO Shou-heng ,SHI Jia-rong. Sparse solution of some special optimization problems [J]. Journal of Interdisciplinary Mathematics 2017 20( 2) : 595-602.
- [38] 雍龙泉. 绝对值函数的一致光滑逼近函数 [J]. 数学的实践与认识 2015 45( 20) : 250-255.
- [39] 雍龙泉. 一致光滑逼近函数及其性质 [J]. 陕西理工大学学报( 自然科学版) 2018 34( 1) : 74-79.
- [40] 雍龙泉. 基于上方一致光滑逼近函数的高阶牛顿法求解线性规划 [J]. 吉林大学学报( 理学版) 2019 57( 2) : 265-270.
- [41] CHEN Jein-shan ,KO Chun-hsu ,PAN Shao-hua. A neural network based on generalized Fischer-Burmeister function for nonlinear complementarity problems [J]. Information Sciences 2010 180: 697-711.
- [42] KO Chun-hsu ,CHEN Jein-shan ,YANG Ching-yu. Recurrent neural networks for solving second-order cone programs [J]. Neurocomputing 2011 74( 17) : 3646-3653.
- [43] MIAO Xin-he ,CHEN Jein-shan ,KO Chun-hsu. A neural network based on the generalized FB function for nonlinear convex programs with second-order cone constraints [J]. Neurocomputing 2016 203: 62-72.
- [44] SUN Ju-he ,WU Xiao-ren ,SAHEYA B. Neural network for solving SOCQP and SOCCVI based on two discrete-type classes of SOC complementarity functions [J]. Mathematical Problems in Engineering 2019( PT. 4) : 4545064. 1-4545064. 18.
- [45] ALCANTARA J H ,CHEN Jein-shan. Neural networks based on three classes of NCP-functions for solving nonlinear complementarity problems [J]. Neurocomputing 2019( 24) : 102-113.
- [46] 雍龙泉, 刘三阳, 史加荣, 等. 几类特殊矩阵及性质 [J]. 兰州大学学报( 自然科学版) 2016 52( 3) : 385-389.
- [47] 雍龙泉. 一类高阶牛顿迭代法及其在线性互补问题中的应用 [J]. 数学的实践与认识 2019 49( 14) : 160-167.

[责任编辑: 魏 强]

## Neural network method for absolute value equation and linear complementarity problem

YONG Long-quan<sup>1 2</sup>

( 1. School of Mathematics and Computer Science , Shaanxi University of Technology ,  
Hanzhong 723000 , China

2. Shaanxi Key Laboratory of Industrial Automation , Hanzhong 723000 , China)

**Abstract:** Three smooth approximating functions of absolute value function are listed , and their properties are analyzed. After selecting smooth function with better properties to deal with the absolute value equation , a differentiable unconstrained optimization problem is reached , and the problem can be solved by neural network model based on gradient descend. By solving some absolute value equations with unique solution or multiple solutions , the results show that this method is not dependent on the initial point , and has the advantages of fast convergence. Finally we apply the method to solve the linear complementary problem.

**Key words:** absolute value function; smooth approximation function; absolute value equation; unconstrained optimization; neural network model based on gradient descend; linear complementarity problem