

一类求解优化问题的神经网络及其全局吸引性分析

王占山¹, 康云云², 牛海莎¹

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 国网池州供电公司, 安徽 池州 247100)

摘 要: 构造了一个以微分包含形式给出的神经网络模型来求解带有等式约束和不等式约束的非线性最优化问题. 通过在网络模型中引入含有加权矩阵的高阶补偿项, 不仅提高了神经网络优化计算的收敛速度, 而且改进了优化解从不可行域逐步收敛到稳定域的问题. 理论上不仅证明了神经网络的解的全局存在性和唯一性, 也证明了解的有界性以及有限的时间内收敛到最优化问题所确定的最优解集中, 并分析了神经网络的全局吸引性. 通过三个数值例子验证了所提出的神经网络优化的有效性.

关 键 词: 神经网络; 微分包含; 神经计算; 优化; 全局吸引

中图分类号: TP 183 **文献标志码:** A **文章编号:** 1005-3026(2017)02-0153-05

A Class of Neural Networks for Solving Optimization Problems with Global Attractivity

WANG Zhan-shan¹, KANG Yun-yun², NIU Hai-sha¹

(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China; 2. Chizhou Power Supply Company of State Grid, Chizhou 247100, China. Corresponding author: WANG Zhan-shan, E-mail: zhanshan_wang @ 163. com)

Abstract: A recurrent neural network in the form of differential inclusion was proposed for solving a class of nonlinear optimization problems, where the constraints were defined by a class of inequality and equality constraints. A higher-order compensation term was involved in the considered neural model, therefore, the convergence rate of the neural computation was significantly increased and the unstable problem of the optimal solution from infeasible domain to feasible domain was solved. In theory, it is proven that not only the solution of the proposed network exists globally and uniquely, but also the solution of the proposed network is bounded and is convergent to the optimal solution set of the optimization problem. Meanwhile, global attractivity of the neural network was analyzed. Three numerical examples were used to show the effectiveness and good performance of the proposed neural network.

Key words: neural networks; differential inclusion; neural computation; optimization; global attractivity

非线性优化可以广义地理解为带有等式和不等式约束条件的目标函数优化问题^[1-11]. 在各种各样的科学与工程领域中都经常遇到这类最优化问题. 目前, 已经有很多方法可以解决这类问题, 然而, 在这些方法中, 多数需要大量的计算, 而且并不适合于解决实时或近似实时最优化的问题. 神经网络方法为解决约束最优化问题提供了一个有效的发展方向, 通过使用具有高度并行计算能力的神经网络体系结构, 使许多复杂的最优化问

题得到实时解决^[2].
1982 年美国物理学家 Hopfield 教授首次将优化问题与一类神经网络相联系^[3], 并利用能量函数的概念, 使所设计的神经网络能够等解相应的优化问题. 然而, 因为网络中要用到目标函数和约束函数的梯度, 所以大部分神经网络模型只适用于解决光滑凸优化问题; 但在实际应用中, 大部分优化问题都不具备光滑的凸性性质, 所以设计适当的神经网络来解决非光滑凸优化和非光滑非

凸优化问题是很必要的.

文献[4]构造了一个神经网络模型来求解非光滑凸优化问题,文献[5]建立了一个单层递归神经网络模型来求解带有线性等式约束条件的伪凸优化问题;然而文献[4-5]求解的都是只带有线性等式约束的非光滑凸优化问题,没有考虑不等式约束的情况.

文献[6]中不是任意的初始点都可以使神经网络收敛于优化问题的最优解.文献[7]中由于正增益的不同,神经网络的解是不一样的,很可能不收敛.无论待解决的最优化问题是否光滑,大部分的神经网络技术解决此类约束最优化问题都是带惩罚项的,而且只能给出近似解或有可能无法用电路实现.基于此,文献[9]提出了一种带有微分包含形式的神经网络模型来解决非光滑凸优化问题,克服了罚函数法引入的目标函数二阶导数不连续的缺点;但是在优化计算时,仍存在收敛速度慢,需要可行域来确保神经网络的有效性,以及从不可行域逐步收敛到解的不足.

为克服上述文献中的缺点,本文通过在网络模型中引入含有加权矩阵的高阶补偿项,重新构造了一类优化神经网络,不仅对于任意的初始值,神经网络的解都能收敛到优化问题的最优解,而且不用确定收敛域来确保神经网络的有效性,收敛速度也得到了提高.

1 神经网络模型与理论分析

1.1 预备知识

首先给出非线性规划的一般形式:

$$\begin{aligned} \min \quad & f(\mathbf{x}) ; \\ \text{s. t.} \quad & g_i(\mathbf{x}) \leq 0, \quad i=1,2,\cdots,m, \\ & \mathbf{Ax} = \mathbf{b}_0. \end{aligned} \quad (1)$$

式中: $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbf{R}^n$; $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 为目标函数,可以是非光滑非凸函数,在本文中假设 f 是有界的; $g_i (i=1,2,\cdots,m): \mathbf{R}^n \rightarrow \mathbf{R}$ 是不等式约束函数,可以是非光滑非凸函数; $\mathbf{A} \in \mathbf{R}^{m \times n}$ 是一个行满秩矩阵; $\mathbf{b}_0 \in \mathbf{R}^m$.

如果 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 在 \mathbf{R}^n 上是局部 Lipschitz 的,则 f 在 \mathbf{x} 点沿方向 $\mathbf{v} \in \mathbf{R}^n$ 的广义方向导数定义为

$$f^0(\mathbf{x}; \mathbf{v}) = \limsup_{y \rightarrow \mathbf{x}, t \rightarrow 0^+} \frac{f(\mathbf{y} + t\mathbf{v}) - f(\mathbf{y})}{t}.$$

引理 1^[9] 如果 $V: \mathbf{R}^n \rightarrow \mathbf{R}$ 对 $\mathbf{x}(t)$ 是正则函数且 $\mathbf{x}(t)$ 对 t 是可微的,且在 t 附近是 Lipschitz 的,则

$$\begin{aligned} \frac{d}{dt} V(\mathbf{x}(t), t) &= \langle \xi, \dot{\mathbf{x}} \rangle = \xi^T \dot{\mathbf{x}}, \\ \forall \xi &\in \partial V(\mathbf{x}(t)). \end{aligned}$$

1.2 构造神经网络

满足所有约束条件的点集(可行域)可表示为

$$S = \{ \mathbf{x} \in \mathbf{R}^n : g(\mathbf{x}) \leq 0, \mathbf{Ax} = \mathbf{b} \};$$

$$S_1 = \{ \mathbf{x} \in \mathbf{R}^n : g(\mathbf{x}) \leq 0 \}, S_2 = \{ \mathbf{x} \in \mathbf{R}^n : \mathbf{Ax} = \mathbf{b} \}.$$

为解决问题(1),构造如下神经网络模型

$$\begin{aligned} \dot{\mathbf{x}}(t) \in & -(\partial L(\mathbf{x}(t)) + \varepsilon(t)(\mathbf{I} - \mathbf{P})\partial \mathcal{B}(\mathbf{x}(t)) + \\ & \varepsilon^2(t)(\mathbf{I} - \mathbf{P})\partial f(\mathbf{x}(t)) + \varepsilon^3(t)(\mathbf{I} - \mathbf{P})\mathbf{x}(t)). \end{aligned} \quad (2)$$

\mathbf{I} 是单位矩阵, $\mathbf{P} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{A}$.

$$\mathcal{B}(\mathbf{x}) = \begin{cases} 0, & \text{若 } \mathbf{x} \in S_1; \\ \sum_{i \in I_+(\mathbf{x})} g_i(\mathbf{x}), & \text{若 } \mathbf{x} \notin S_1. \end{cases}$$

$$\partial \mathcal{B}(\mathbf{x}) = \begin{cases} \sum_{i \in I_0(\mathbf{x})} [0, 1] \partial g_i(\mathbf{x}), & \mathbf{x} \in \text{bd}(S_1); \\ 0, & \mathbf{x} \in \text{int}(S_1); \\ \sum_{i \in I_+(\mathbf{x})} \partial g_i(\mathbf{x}), & \mathbf{x} \in S_1 \text{ 且 } I_0(\mathbf{x}) = \emptyset; \\ \sum_{i \in I_+(\mathbf{x})} \partial g_i(\mathbf{x}) + \sum_{i \in I_0(\mathbf{x})} [0, 1] \partial g_i(\mathbf{x}), & \mathbf{x} \in S_1 \text{ 且 } I_0(\mathbf{x}) \neq \emptyset. \end{cases} \quad (3)$$

$$I_+(\mathbf{x}) = \{ i \in \{1, 2, \cdots, m\} : g_i(\mathbf{x}) > 0 \},$$

$$I_0(\mathbf{x}) = \{ i \in \{1, 2, \cdots, m\} : g_i(\mathbf{x}) = 0 \}.$$

$$L(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_1,$$

$$\partial L(\mathbf{x}) = \mathbf{A}^T K[\phi_{[-1,1]}(\mathbf{Ax} - \mathbf{b})].$$

式中 $\text{bd}(\cdot)$ 为边界域, $\text{int}(\cdot)$ 为内部域.

1.3 主要结果

下面给出本文的定理及其证明过程.在进行分析时首先要满足一些假设条件.

假设条件 1: 存在 $\bar{\mathbf{x}} \in \mathbf{R}^n$ 和 $r > 0$, 有 $\bar{\mathbf{x}} \in \text{int}(S_1) \cap S_2$ 和 $S \subset \mathcal{B}(\bar{\mathbf{x}}, r)$, 其中 $\mathcal{B}(\bar{\mathbf{x}}, r) = \{ \mathbf{x} \in \mathbf{R}^n : \|\mathbf{x} - \bar{\mathbf{x}}\| \leq r \}$.

假设条件 2: 目标函数 $f(\mathbf{x})$ 在可行域 S 上是凸函数且在 $\mathcal{B}(\bar{\mathbf{x}}, r) \cap S_2$ 上是 Lipschitz 有界, 不等式约束 $g_i(\mathbf{x})$ 在 S_2 上是凸函数.

说明 1: 从假设条件 1 中可知可行域 S 是有界的, 而且 $\text{int}(S_1) \cap S_2$ 是非空的, 而文献[4]~文献[8]中的假设条件不等式约束可行域 S_1 是有界的, 相比这些文献, 本文所提的假设条件更弱, 更具一般性.

定理 1 当假设条件 1 和假设条件 2 满足时, 构造的神经网络(2)的解全局存在且唯一.

证明 由于神经网络(2)是一个非空紧凸值的上半连续集值映射,因此,对任意 $\mathbf{x}_0 \in \mathbf{B}(\hat{\mathbf{x}}, r)$, 神经网络(2)存在一个局部解 $\mathbf{x}: [0, T) \rightarrow \mathbf{R}^n$, 其中 $[0, T)$ 是神经网络(2)的解存在的最大区间. 由于 f 在可行域 S 上是凸函数, $\|\dot{\mathbf{x}}(t)\|$ 具有非增性,所以对神经网络的两个解 $\mathbf{x}(t), \mathbf{y}(t) \in S$ 时,有

$$\begin{aligned} & \frac{d}{dt} \left(\frac{1}{2} \|\mathbf{x}(t) - \mathbf{y}(t)\|^2 \right) = \\ & (\mathbf{x}(t) - \mathbf{y}(t))^T - (\dot{\mathbf{x}}(t) - \dot{\mathbf{y}}(t)) \leq 0, \\ & \text{可推出} \\ & \|\mathbf{x}(t) - \mathbf{y}(t)\| \leq \|\mathbf{x}_0 - \mathbf{y}_0\|, \forall t \in [0, T). \quad (4) \end{aligned}$$

取 $h \in [0, T)$ 使 $\mathbf{y}_0 = \mathbf{x}(h)$, 则

$$\begin{aligned} & \|\mathbf{x}(t)\| \leq \|\mathbf{x}(t-h)\| + \|\mathbf{x}(0) - \mathbf{x}(h)\| \\ & \Rightarrow \lim_{x \rightarrow T} \|\mathbf{x}(t)\| \leq \|\mathbf{x}(t-T)\| + \|\mathbf{x}(0) - \mathbf{x}(h)\|. \end{aligned}$$

$\mathbf{x}(t)$ 在 $[0, T)$ 有界推出它可被扩展,这与 $[0, T)$ 是神经网络(2)存在最大区间的解相矛盾,所以神经网络(2)的解是全局存在,而且通过式(4)可以看出当初始点为 \mathbf{x}_0 时,解是唯一的.

定理 2 当假设条件 1 和假设条件 2 满足时,对任意初始点,神经网络(2)的解有界并且在有限时间内收敛到可行域中.

证明 首先构造 Lyapunov 能量函数 $V(\mathbf{x}, t) = L(\mathbf{x}(t)) + \varepsilon(t)(\mathbf{I} - \mathbf{P})\mathcal{B}(\mathbf{x}(t)) + \varepsilon^2(t)(\mathbf{I} - \mathbf{P})f(\mathbf{x}(t)) + \frac{1}{2}\varepsilon^3(t)(\mathbf{I} - \mathbf{P})\|\mathbf{x}(t)\|^2$. (5)

$$\begin{aligned} & \frac{d}{dt} V(\mathbf{x}(t), t) = \\ & -\|\dot{\mathbf{x}}(t)\|^2 + (\mathbf{I} - \mathbf{P})(\varepsilon(t)\mathcal{B}(\mathbf{x}(t)) + \\ & 2\varepsilon(t)\varepsilon(t)f(\mathbf{x}(t)) + \frac{3}{2}\varepsilon^2(t)\varepsilon(t)\|\mathbf{x}(t)\|^2). \quad (6) \end{aligned}$$

由式(6)和可行域 S 有界可知 $V(\mathbf{x}(t), t)$ 是非增的.

$$\begin{aligned} & \text{记 } \mathcal{M} = \bar{\xi}_2 + (\mathbf{I} - \mathbf{P})(\varepsilon(t)\bar{\xi}_1 + \varepsilon^2(t)\bar{\eta} + \varepsilon^3(t)\mathbf{x}(t)), \\ & \frac{d}{dt} \left(\frac{1}{2} \|\mathbf{x}(t) - \hat{\mathbf{x}}\|^2 \right) = (\mathbf{x}(t) - \hat{\mathbf{x}})^T \dot{\mathbf{x}}(t) = \\ & (\mathbf{x}(t) - \hat{\mathbf{x}})^T (-\bar{\xi}_2 - (\mathbf{I} - \mathbf{P})(\varepsilon(t)\bar{\xi}_1 + \varepsilon^2(t)\bar{\eta} + \varepsilon^3(t)\mathbf{x}(t))) = -(\mathbf{x}(t) - \hat{\mathbf{x}})^T \mathcal{M}. \quad (7) \end{aligned}$$

其中 $\hat{\mathbf{x}}$ 为神经网络的最优解,由于 $V(\mathbf{x}, t)$ 在可行域 S 上为凸函数,所以当 $\mathbf{x} \in S$ 且结合式(7),可以得到

$$\begin{aligned} & V(\mathbf{x}, t) - V(\hat{\mathbf{x}}, t) \leq \langle \mathcal{M}, \mathbf{x} - \hat{\mathbf{x}} \rangle \leq \\ & -\frac{d}{dt} \left(\frac{1}{2} \|\mathbf{x}(t) - \hat{\mathbf{x}}\|^2 \right), \quad (8) \end{aligned}$$

$$\begin{aligned} & 2 \int_0^t (V(\mathbf{x}(s), s) - V(\hat{\mathbf{x}}, s)) ds \leq \|\mathbf{x}_0 - \hat{\mathbf{x}}\|^2 - \\ & \|\mathbf{x}(t) - \hat{\mathbf{x}}\|^2. \quad (9) \end{aligned}$$

由式(9)结合 $V(\mathbf{x}, t)$ 的非增特性得

$$\begin{aligned} & \int_0^t V(\mathbf{x}(s), s) ds \geq tV(\mathbf{x}(t), t) \geq \\ & t[L(\mathbf{x}(t)) + \varepsilon(t)(\mathbf{I} - \mathbf{P})f(\mathbf{x}(t))] , \quad (10) \end{aligned}$$

$$\int_0^t V(\hat{\mathbf{x}}, s) ds = \int_0^t (\varepsilon^2(s)f(\hat{\mathbf{x}}) + \frac{1}{2}\varepsilon^3(s)\|\hat{\mathbf{x}}\|^2) ds , \quad (11)$$

$$\begin{aligned} & t[L(\mathbf{x}(t)) + \varepsilon(t)(\mathbf{I} - \mathbf{P})\mathcal{B}(\mathbf{x}(t)) + \varepsilon^2(t)(\mathbf{I} - \\ & \mathbf{P})f(\mathbf{x}(t))] \leq \frac{1}{2} \|\mathbf{x}_0 - \hat{\mathbf{x}}\|^2 + \end{aligned}$$

$$\int_0^t (\varepsilon^2(s)f(\hat{\mathbf{x}}) + \frac{1}{2}\varepsilon^3(s)\|\hat{\mathbf{x}}\|^2) ds = W. \quad (12)$$

由式(12)可得

$$\begin{aligned} & L(\mathbf{x}(t)) \leq \frac{W}{(\mathbf{I} - \mathbf{P})t} \\ & \Rightarrow \lim_{t \rightarrow +\infty} L(\mathbf{x}(t)) = 0. \quad (13) \end{aligned}$$

$$\begin{aligned} & \mathcal{B}(\mathbf{x}(t)) \leq \frac{W}{(\mathbf{I} - \mathbf{P})\varepsilon(t)t} \\ & \Rightarrow \lim_{t \rightarrow +\infty} \mathcal{B}(\mathbf{x}(t)) = 0. \quad (14) \end{aligned}$$

因为可行域 S 有界, $\mathbf{x}(t)$ 在 $[0, +\infty)$ 是有界的,即 $\|\mathbf{x}(t)\| \leq \gamma$, 并且神经网络(2)的解收敛到式(1)可行域 S , 即 $\lim_{t \rightarrow \infty} \min \|\mathbf{x}(t) - S\| = 0$.

说明 2: 定理 2 的结论对 $\text{int}(S_1) = \emptyset$ 时仍然有效,上面证明了最后收敛到可行域. 下面进一步说明可在有限时间内收敛到可行域.

由于 $\mathbf{A}(\mathbf{I} - \mathbf{P}) = 0$, 所以当 $\mathbf{x} \in \mathbf{R}^n \setminus S_2$ 时, $\frac{d}{dt} L(\mathbf{x}) = -\|\mathbf{A}^T \boldsymbol{\eta}\|_2^2, \forall \boldsymbol{\eta} \in K[\phi_{[-1,1]}(\mathbf{A}\mathbf{x} - \mathbf{b})]$.

又因为 $\|\mathbf{A}^T \boldsymbol{\eta}\| \geq \sqrt{\lambda_{\min}(\mathbf{A}\mathbf{A}^T)}$, 所以

$$\begin{aligned} & \frac{d}{dt} L(\mathbf{x}) = -\lambda_{\min}(\mathbf{A}\mathbf{A}^T) \\ & \Rightarrow L(\mathbf{x}) \leq L(\mathbf{x}(0)) - \lambda_{\min}(\mathbf{A}\mathbf{A}^T)t. \quad (15) \end{aligned}$$

由式(15)知,当 $t = L(\mathbf{x}(0))/\lambda_{\min}(\mathbf{A}\mathbf{A}^T)$ 时, $L(\mathbf{x}) = 0, \mathbf{x}(t)$ 收敛到 S_2 , 其中 λ_{\min} 为矩阵的最小特征值.

当进入到 S_2 时, $P\dot{\mathbf{x}} = 0, \dot{\mathbf{x}} = (\mathbf{I} - \mathbf{P})\dot{\mathbf{x}}$,

$$\begin{aligned} & \frac{d}{dt} \mathcal{B}(\mathbf{x}(t)) = \langle \zeta_2, -(\mathbf{I} - \mathbf{P})(\varepsilon(t)\zeta_2 + \varepsilon^2(t)\boldsymbol{\eta}_2 + \varepsilon^3(t)\mathbf{x}(t)) \rangle \leq -\varepsilon(t)\|(\mathbf{I} - \mathbf{P})\zeta_2\|(\|(\mathbf{I} - \mathbf{P})\zeta_2\| - \varepsilon(t)(l_f + \gamma)). \quad (16) \end{aligned}$$

其中, $\zeta_2 \in \partial \mathcal{B}(\mathbf{x}(t)), \boldsymbol{\eta}_2 \in f(\mathbf{x}(t)), \|\boldsymbol{\eta}_2\| \leq l_f$. l_f 是一个正常数.

由于当 $\mathbf{x}(t) \in S_2 \setminus S_1$ 时, 存在 β 满足 $\|(\mathbf{I} -$

$P)\zeta_2\|\geqslant\frac{\beta}{2\gamma}$,可知当 $\varepsilon(t)\leqslant\frac{\beta}{4\gamma(l_f+\gamma)}$ 时到达可行域 S_1 ,因为 $\lim_{t\rightarrow\infty}\varepsilon(t)=0$,所以一定存在一个 \bar{t} ,当 $t\geqslant\bar{t}$ 时到达可行域 S_1 .

定理 3 当假设条件 1 和假设条件 2 成立时,对任意的初始点,神经网络(2)的解收敛到式(1)的最优解集.

定理 4 当假设条件 1 和假设条件 2 成立时,对任意初始点, $\lim_{t\rightarrow\infty}\|\mathbf{x}(t)-\hat{\mathbf{x}}\|=0$,即神经网络具有全局吸引的性质.

说明 3:定理 1 给出了神经网络解的存在性及唯一性,定理 2 证明了神经网络的解在有限的时间内是收敛到可行域的,定理 3 证明了神经网络的解能收敛到优化问题的最优解集中,最后,定理 4 证明了本文所提出的神经网络具有全局吸引的性质.

2 仿真分析

例 1 光滑非凸优化问题:

$$\begin{aligned} \min \quad & f(\mathbf{x})=x_1-x_2^2; \\ \text{s. t.} \quad & x_1^2+x_2^2\leqslant 5, \\ & x_1+x_2=1. \end{aligned}$$

取 $\varepsilon(t)=1/(t+1)^{1/3}$,从图 1 可以看出神经网络(6)的解均能收敛到优化问题的最优解,即 $\hat{\mathbf{x}}=(-1,2)^T$.这里引用的是文献[6]中的例子,文献[6]中以 $(7/2,-5/2)^T$ 和 $(-3/2,5/2)^T$ 为不同的初始点,网络的轨道分别收敛于 $(2,-1)^T$ 和 $(-1,2)^T$,但只有 $(-1,2)^T$ 是能达到其在可行域上的最小值.但本文中提出的网络轨道能以不同的初始点都收敛到可行域上的最小值,更具优势.

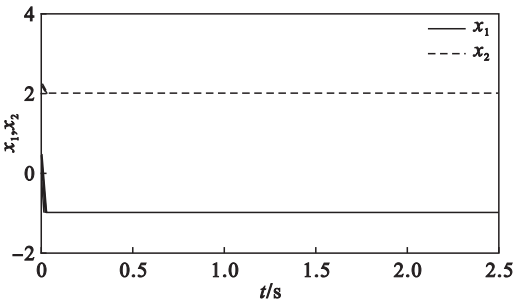


图 1 任意 10 个初始点的状态轨迹仿真图
Fig. 1 Properties of $(x_1,x_2)^T$ with any ten initial points

例 2 非光滑优化问题:

$$\begin{aligned} \min \quad & f(\mathbf{x})=2|x_1-x_2+2|+|x_2-x_3+2x_4|+ \\ & |x_3+x_4-4|+\cos(0.2(x_1-x_4)); \\ \text{s. t.} \quad & |2x_1-x_2|+x_3^2-4\leqslant 0, \end{aligned}$$

$$\begin{aligned} -x_2-x_3-|x_4+2|-2\leqslant 0, \\ x_1-x_4-5\leqslant 0, \\ x_4-x_1-5\leqslant 0, \\ x_1^2+x_2^2+x_3^2+x_4^2-8\leqslant 0, \\ x_1+2x_2=2, \\ -2x_3+x_4=-2. \end{aligned}$$

取 $\varepsilon(t)=4/(t+1)^{1/2}$,从图 2 可以看出,神经网络(6)最终均收敛于优化问题的最优解,即 $\hat{\mathbf{x}}=(-0.9,1.45,0.85,-0.3)^T$.这里引用的是文献[7]中的例子,从仿真图中可以看出,任意初始点时仍能收敛到可行域中,并且不用确定罚参数.但是在文献[7]中,正增益 σ 不同时,神经网络的解不一定是优化问题的最优解.所以,本文提出的神经网络效果更好,更具优势.

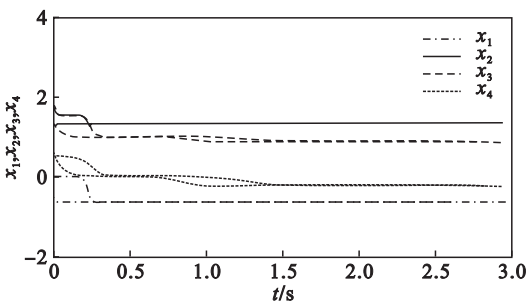


图 2 任意 10 个初始点的状态轨迹仿真图
Fig. 2 Properties of $(x_1,x_2,x_3,x_4)^T$ with any ten initial points

例 3 非光滑凸优化问题:

$$\begin{aligned} \min \quad & f(\mathbf{x})=(x_1-x_2+x_3)^2+|x_1+2x_3|+ \\ & |x_2-1|+20(x_3-2)^2+e^{-x_1-x_2-x_3}; \\ \text{s. t.} \quad & x_1^2+0.5e^{|x_1-1|}-6x_2-4x_3\leqslant -2, \\ & x_3^2-2\leqslant 0, \\ & x_1+3x_2+0.2x_3=1, \\ & x_i\geqslant 0, i=1,2,3. \end{aligned}$$

取 $\varepsilon(t)=1/(t+1)^{1/3}$,图 3 是神经网络(6)以 $(0,0,0)^T$ 为初始点的状态轨迹仿真图,最终收敛于最优解 $\hat{\mathbf{x}}=(0,0.228\ 8,1.569\ 0)^T$.

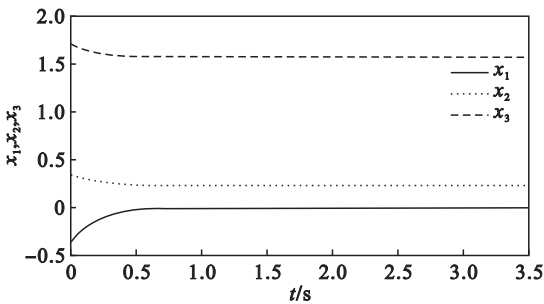


图 3 以 $(0,0,0)^T$ 为初始点的状态轨迹仿真图
Fig. 3 Properties of $(x_1,x_2,x_3)^T$ with zero initial points

图 4 是神经网络(6)以 $(0,0,0)^T$ 为初始点的目标函数和约束条件的状态轨迹仿真图. 图 5 是任意 10 个初始点的目标函数的状态轨迹仿真图.

将本文结果与文献[9]的结果作了对比,从仿真结果可以看出,本文所提出的神经网络的收敛速度更快,并且增加了任意初始点时目标函数 f 的仿真图.

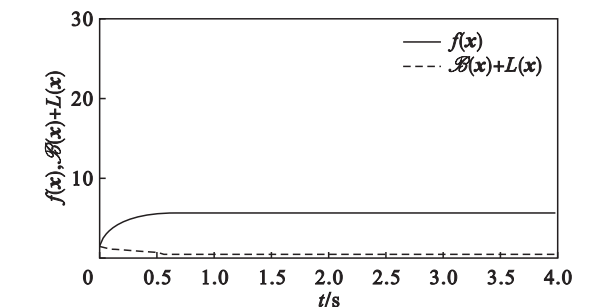


图 4 以 $(0,0,0)^T$ 为初始点的目标函数和约束条件的状态轨迹仿真图
Fig. 4 Properties of $f(x)$ and $B(x) + L(x)$ with initial point $(0,0,0)^T$

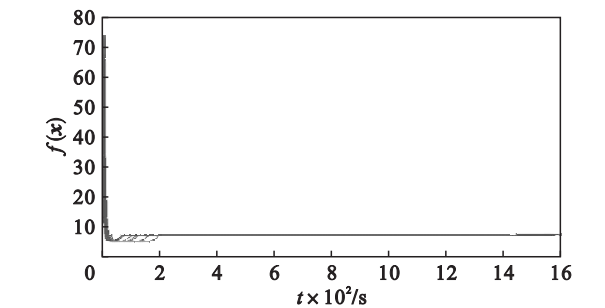


图 5 任意 10 个初始点的目标函数的状态轨迹仿真图
Fig. 5 Properties of f with any ten initial points

3 结 语

本文构造了一类新的具有微分包含的神经网络模型来解决一些非线性优化问题,通过引入 $\varepsilon(t)$ 的高阶补偿项及相应的加权矩阵,使得神经网络具有全局吸引的性质,并且不用确定罚函数,这使得提出的神经网络模型对解决非线性优化问题更有效. 与现有的文献进行对比,本文所提出的神经网络可以用电路实现,对任意的初值,神经网络的解都能使目标函数取得最小值,并且本文不需要先找到可行域就能确保神经网络模型的有效性. 通过三个数值例子验证了本文方法的有效性.

参考文献:

[1] Bazaraa M S, Sherali H D, Shetty C M. Nonlinear programming: theory and algorithms [M]. Hoboken: John Wiley & Sons, 2013: 1–21.

[2] Cichocki A, Unbehauen R. Neural networks for optimization and signal processing [M]. Hoboken: John Wiley & Sons, 1993: 1–20.

[3] 季策, 张化光. 一类具有时滞的广义 Hopfield 神经网络的动态分析[J]. 东北大学学报(自然科学版), 2004, 25(3): 205–208.
(Ji Ce, Zhang Hua-guang. Dynamic analysis of a class of generalized Hopfield neural networks with hysteresis [J]. Journal of Northeastern University (Natural Science), 2004, 25(3): 205–208.)

[4] Liu Q S, Wang J. A one-layer recurrent neural network for non-smooth convex optimization subject to linear equality constraints [M]// Advances in Neuro-Information Processing. Heidelberg: Springer Berlin, 2009: 2–30.

[5] Guo Z S, Liu Q S, Wang J. A one-layer recurrent neural network for pseudoconvex optimization subject to linear equality constraints [J]. IEEE Transactions on Neural Networks, 2011, 22(12): 1892–1900.

[6] Bian W, Xue X P. Subgradient-based neural networks for nonsmooth nonconvex optimization problems [J]. IEEE Transactions on Neural Networks, 2009, 20(6): 1024–1038.

[7] Liu Q S, Wang J. A one-layer recurrent neural network for constrained nonsmooth optimization [J]. IEEE Transactions on Systems, Man, and Cybernetics, 2011, 41(5): 1323–1333.

[8] Li G C, Song S J, Wu C. Generalized gradient projection neural networks for nonsmooth optimization problems [J]. Science China: Information Sciences, 2010, 53(5): 990–1005.

[9] Bian W, Xue X P. Neural network for solving constrained convex optimization problems with global attractivity [J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2013, 60(3): 710–723.

[10] 耿平, 刘静, 曾梅光. 多变量非线性复杂系统的优化模拟退火算法[J]. 东北大学学报(自然科学版), 2002, 23(3): 270–272.
(Geng Ping, Liu Jing, Zeng Mei-guang. Multivariate nonlinear complex system optimization and simulated annealing algorithm [J]. Journal of Northeastern University (Natural Science), 2002, 23(3): 270–272.)

[11] 王占山. 复杂神经动力网络的稳定性和同步性[M]. 北京: 科学出版社, 2014.
(Wang Zhan-shan. Stability and synchronization of complex neural dynamical networks [M]. Beijing: Science Press, 2014.)