Advances in Applied Mathematics 应用数学进展, 2026, 15(1), 336-350
Published Online January 2026 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2026.151033

Hans汉斯

一种用于求解等式约束鞍点问题的递归神经网络

叶洪志，李小兵

重庆交通大学数学与统计学院，重庆

收稿日期：2025年12月21日；录用日期：2026年1月16日；发布日期：2026年1月21日

摘　要

鞍点问题在工程优化、经济决策、信号处理、智能控制等诸多领域中具有广泛的应用背景，本文针对等式约束下的鞍点问题，提出了一种新型递归神经网络模型。通过构建相应的动力学方程，分析了该神经网络在Lyapunov意义下的稳定性，并证明了其平衡点与原问题最优解的等价性。

关键词

递归神经网络，鞍点问题，等式约束，全局收敛，Lyapunov稳定性

A Recurrent Neural Network for Solving Equality-Constrained Saddle Point Problems

Hongzhi Ye, Xiaobing Li

School of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing

Received: December 21, 2025; accepted: January 16, 2026; published: January 21, 2026

Abstract

Saddle point problems have broad applications in various fields such as engineering optimization, economic decision-making, signal processing, and intelligent control. This paper proposes a novel recurrent neural network model for solving equality-constrained saddle point problems. By constructing the corresponding dynamic equations, the stability of the neural network in the Lyapunov sense is analyzed, and the equivalence between its equilibrium point and the optimal solution of the original problem is proven.

文章引用: 叶洪志, 李小兵. 一种用于求解等式约束鞍点问题的递归神经网络[J]. 应用数学进展, 2026, 15(1): 336-350.
DOI: 10.12677/aam.2026.151033

Keywords

Recurrent Neural Network, Saddle Point Problem, Equality Constraints, Global Convergence, Lyapunov Stability

1. 引言

鞍点问题是数学规划与计算数学中的一个核心课题，其一般形式可表述为：$\min_x \max_y f(x,y)$，也称为 min-max 问题，该问题的解(x^*, y^*)被称为鞍点。这类问题在众多科学与工程领域有着广泛的应用背景。例如，在鲁棒优化中，为了应对数据的不确定性，决策者在最坏情况下做出最优决策，这直接导出了一个 min-max 问题[1]；在现代机器学习中，生成对抗网络(GANs)的训练过程本质上是一个鞍点优化问题[2]，而分布式机器学习中的对偶优化方法[3]、公平分类[4]以及对抗性训练[5]等核心问题，也都可归结为求解鞍点问题；在经济学与博弈论中，市场均衡分析中的 Nash 均衡点也常表现为鞍点形式[6]。再加上各种约束，特别是等式约束的引入，使得这类带约束的鞍点问题具有更加重要的研究价值。

在传统优化方法中，交替梯度下降－上升法[7]、乐观梯度下降法[8]和交替方向乘子法[3] [9]等方法被广泛用于处理鞍点问题。然而，这些传统方法的收敛速度与迭代步长的选择紧密相关，且通常需要进行多次迭代才能获得高精度解。这在需要在线或实时求解的应用场景(如实时优化决策、机器人在线路径规划等)之中可能并不适用。

为解决上述问题，众多学者致力于基于神经网络的优化问题求解方法的研究，并取得了丰硕的成果[10]-[26]。早期工作中，Tank 与 Hopfield [21]提出了用于线性规划 Hopfield 网络，开创了神经网络求解优化问题的先河。随后，Kennedy 与 Chua [22]提出了用于非线性规划的神经网络模型，进一步推动了该领域的发展。在约束处理方面，Zhang 与 Constantinides [23]提出了拉格朗日神经网络，Xia 与 Wang [24]提出了投影神经网络，分别用于处理等式与不等式约束问题。

特别地，基于动力系统的递归神经网络(RNN)方法因其能够并行处理的能力与硬件的可实现性，近年来受到了广泛学者的关注[17] [18]。与传统的迭代算法相比，RNN 通过构建连续时间动力学系统，能够实现高效、实时的优化求解，特别适用于高维、大规模问题[19] [20] [25]。例如，针对以下带等式约束的优化问题(1)，Qin 等人[25]提出了一种单层递归神经网络，证明了其平衡点与原问题最优解的等价性，并建立了系统的 Lyapunov 稳定性与全局收敛性。

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{1}$$

然而，现研究的许多神经网络模型主要针对的是带等式或不等式约束的单目标优化问题，以及不等式约束下的 min-max 问题，很多都未涉及等式约束下的鞍点问题。鉴于此，本文在 Qin 等人[25]的工作基础上，将其研究的单目标等式约束优化问题拓展至等式约束下的 min-max 问题，并构建了一种新型递归神经网络模型，用于求解带有线性等式约束的鞍点优化问题。

本文的主要贡献包括：(1) 建立了一种用于求解等式约束鞍点问题的新型递归神经网络动力学方程，并证明了其平衡点与鞍点问题最优解的等价性；(2) 分析了神经网络 Lyapunov 意义下的稳定性，证明了

当初始点不在可行域内时其状态向量将指数收敛于可行域，并给出了两个神经网络在 Lyapunov 意义下全局收敛到鞍点问题最优解的充分条件；(3) 通过数值实验验证了所提神经网络在高维复杂问题中的有效性与快速收敛性。

本文的主要结构如下：第 2 节介绍相关预备知识；第 3 节建立问题模型与神经网络动力学方程，介绍与问题模型最优解的等价条件；第 4 节进行所提神经网络的稳定性和收敛性分析；第 5 节通过两个数值算例来验证所提理论的有效性和可行性；第 6 节总结全文并提出未来的研究方向。

2. 预备知识

在本文中，设 R^n 为 n 维实欧几里得空间。$\langle\cdot,\cdot\rangle$ 表示 R^n 的内积，$\|\cdot\|$ 表示其对的范数，对于任意的 $x\in R^n$，用 T 表示向量的转置，有 $\|x\|=\sqrt{\langle x,x\rangle}=\sqrt{x^{\mathrm{T}}x}$。函数 f 在 (x_0,y_0) 处关于 x 和 y 的梯度分别记为 $\nabla_x f(x_0,y_0)$ 和 $\nabla_y f(x_0,y_0)$。

Lyapunov 稳定性和收敛性的相关定义请参考文献[27]。

定义 2.1 [28]：设 $S\subset R^n$ 是非空凸集，$\alpha\in(0,1)$，f 是定义在 S 上的函数，如果对任意 x_1，$x_2\in S$，有

$$f(\alpha x_1+(1-\alpha)x_2)\le+\alpha f(x_1)+(1-\alpha)f(x_2),$$

则称函数 f 是 S 上的凸函数。

定义 2.2 [29]：设 X、Y 是两个集合，$T:X\to Y$ 是一种对应法则，如果对任意 $x\in X$，通过 T 有 Y 中的一个子集 $T(x)$ 与之对应，则称 T 是 X 到 Y 的一个集值映射。

定义 2.3 [29]：设 E 是实 Banach 空间，E^* 是 E 的共轭空间，$D\subset E$，$T:D\to E^*$，$\forall x$，$y\in D$。若 T 满足

$$\langle Tx-Ty,x-y\rangle\ge 0,$$

则称 T 是单调算子。

3. 问题描述和模型建立

在本节中，本文将单目标优化问题(1)拓展成一个鞍点问题，并讨论其等价表述。

考虑以下具有等式约束的鞍点问题：

$$\begin{aligned}\min_x\max_y\quad & f(x,y),\\ \text{s.t.}\quad & Ax=b,\\ & Cy=d,\end{aligned}\tag{2}$$

其中 $x=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}\in R^m$，$y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}\in R^n$，$f:R^m\times R^n\to R$ 是双变量实值函数，$A\in R^{p\times m}$，$C\in R^{q\times n}$，$b\in R^p$，$d\in R^q$。在本文中，记可行域 $E=\left\{(x^{\mathrm{T}},y^{\mathrm{T}})^{\mathrm{T}}\mid Ax=b,Cy=d\right\}$，若点 $\left((x^*)^{\mathrm{T}},(y^*)^{\mathrm{T}}\right)^{\mathrm{T}}\in E$ 满足对任意 $(x^{\mathrm{T}},y^{\mathrm{T}})^{\mathrm{T}}\in E$，都有

$$f(x^*,y)\le f(x,y^*),$$

则称 $\left((x^*)^{\mathrm{T}},(y^*)^{\mathrm{T}}\right)^{\mathrm{T}}$ 是函数 f 在可行域 E 上的鞍点。

若 $\left((x^*)^{\mathrm{T}},(y^*)^{\mathrm{T}}\right)^{\mathrm{T}}\in E$ 是问题(2)的鞍点，根据上述定义，易知，对于任意的 $x\in X(y^*)$，$y\in Y(x^*)$，有

$$f(x^*,y)\le f(x^*,y^*)\le f(x,y^*),$$

其中 $X(y^*)=\{x\mid Ax=b\}$，$Y(x^*)=\{y\mid Cy=d\}$。

假设 1：(1)f在可行域E上是连续可微的，且在$X\left(y^*\right)$上是关于x的凸函数，在$Y\left(x^*\right)$是关于y的凹函数；

(2) A，C均为行满秩矩阵，即$rank\left(A\right)=p\le m$，$rank\left(C\right)=q\le n$。

定理 3.1：若假设 1 成立，$\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}$是问题(2)的最优解，即满足$\forall\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}\in E$，有$f\left(x^*,y\right)\le f\left(x^*,y^*\right)\le f\left(x,y^*\right)$，当且仅当存在$\lambda^*\in R^p$，$\mu\in R^q$，使得$\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}$满足下列方程：

$$\begin{cases}\nabla_x f\left(x^*,y^*\right)+A^{\mathrm{T}}\lambda^*=0,\\ -\nabla_y f\left(x^*,y^*\right)+C^{\mathrm{T}}\mu^*=0,\\ Ax^*=b,\\ Cy^*=d.\end{cases}\tag{3}$$

证明：假设$\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}$是问题(2)的最优解，则有$\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}\in E$ (即$Ax^*=b$，$Cy^*=d$)。一方面固定y^*，由假设 1 可知f在$X\left(y^*\right)$上是关于x的凸函数，于是f在$x^*\in X\left(y^*\right)$处取最小值，有

$$-\nabla_x f\left(x^*,y^*\right)\in N_{X\left(y^*\right)}\left(x^*\right),$$

其中$N_{X\left(y^*\right)}\left(x^*\right)=\left\{v\in R^m:v^{\mathrm{T}}\left(x^*-x\right)\ge 0,\forall x\in X\left(y^*\right)\right\}$。

由于$X\left(y^*\right)$是中的仿射子空间，则

$$N_{X\left(y^*\right)}\left(x^*\right)=\left\{A^{\mathrm{T}}\lambda\mid\lambda\in R^p\right\},$$

故存在一个$\lambda^*\in R^p$，使得当$A^{\mathrm{T}}\lambda^*\in N_{X\left(y^*\right)}\left(x^*\right)$时，有

$$\nabla_x f\left(x^*,y^*\right)+A^{\mathrm{T}}\lambda^*=0.\tag{4}$$

同理固定x^*，由假设 1 可知f在$Y\left(x^*\right)$是关于y的凹函数，于是f在$y^*\in Y\left(x^*\right)$处取最大值，类似上述证明过程，存在一个$\mu^*\in R^q$，有

$$-\nabla_y f\left(x^*,y^*\right)+C^{\mathrm{T}}\mu^*=0.\tag{5}$$

根据$Ax^*=b$，$Cy^*=d$和(4)，(5)可知，满足(3)中所有方程，即必要性得证。

另一方面，假设存在$\lambda^*\in R^p$，$\mu^*\in R^q$，使得(3)成立，则有

$$\begin{cases}\nabla_x f\left(x^*,y^*\right)=-A^{\mathrm{T}}\lambda^*,\\ \nabla_y f\left(x^*,y^*\right)=C^{\mathrm{T}}\mu^*.\end{cases}\tag{6}$$

由假设 1 可知f在可行域E上是连续可微的，且在$X\left(y^*\right)$上是关于x的凸函数，在$Y\left(x^*\right)$是关于y的凹函数，根据凸函数和凹函数的性质，可得

$$\begin{cases}f\left(x,y^*\right)-f\left(x^*,y^*\right)\ge\nabla_x f\left(x^*,y^*\right)^{\mathrm{T}}\left(x-x^*\right),\quad\forall x\in X\left(y^*\right),\\ f\left(x^*,y\right)-f\left(x^*,y^*\right)\le\nabla_x f\left(x^*,y^*\right)^{\mathrm{T}}\left(y-y^*\right),\quad\forall y\in Y\left(x^*\right).\end{cases}\tag{7}$$

结合(6)，(7)，可得

$$\begin{cases} f\left(x,y^*\right)-f\left(x^*,y^*\right)\ge\left(-\lambda^*\right)^{\mathrm{T}}A\left(x-x^*\right)=0, & \forall x\in X\left(y^*\right), \\ f\left(x^*,y\right)-f\left(x^*,y^*\right)\le\left(\mu^*\right)^{\mathrm{T}}C\left(x-x^*\right)=0, & \forall y\in Y\left(x^*\right), \end{cases}$$

即

$$f\left(x^*,y\right)\le f\left(x^*,y^*\right)\le f\left(x,y^*\right),\quad \forall\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}\in E.$$

故 f 在 $\left(x^*,y^*\right)\in E$ 处取得最优解，即充分性得证。

基于定理 3.1，本文在文献[25]中提出的神经网络模型的基础上，将其神经网络框架拓展至等式约束下的鞍点问题(2)，构建了一种用于求解鞍点问题(2)的新型递归神经网络模型，并在定理 4.2 中证明了该神经网络的平衡点和问题(2)的等价性。本文构造的神经网络动力学方程如下：

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}=-\left(I_m-P_x\right)\nabla_x f\left(x,y\right)-A^{\mathrm{T}}\left(Ax-b\right), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=\left(I_n-P_y\right)\nabla_y f\left(x,y\right)-C^{\mathrm{T}}\left(Cy-d\right), \end{cases} \tag{8}$$

其中，$P_y=C^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C$，$\left(I_m-P_x\right)$ 表示将向量投影到矩阵 A 的零空间上，$\left(I_n-P_y\right)$ 表示将向量投影到矩阵 C 的零空间上。

由于 x 的可行域 $\{x\mid Ax=b\}$ 和 y 的可行域 $\{y\mid Cy=d\}$ 是仿射集，故当 x 和 y 在可行域内移动时，只有零空间中的方向移动才不会违反约束 $Ax=b$ 和 $Cy=d$。$-\nabla_x f\left(x,y\right)$，$\nabla_y f\left(x,y\right)$ 分别沿着函数值 $f\left(\cdot,y\right)$ 下降最快的方向和函数值 $f\left(x,\cdot\right)$ 上升最快的方向进行移动，$\left(I_m-P_x\right)$ 和 $\left(I_n-P_y\right)$ 分别对梯度进行投影，使得网络沿着负(或正)梯度方向在可行域的切空间上移动以优化目标函数。而 $\left\|Ax-b\right\|^2$ 和 $\left\|Cy-d\right\|^2$ 分别衡量了当前点 x 和 y 违反约束的程度，当 x 和 y 不在可行域内时，$-A^{\mathrm{T}}\left(Ax-b\right)$ 和 $-C^{\mathrm{T}}\left(Cy-d\right)$ 为负梯度方向，分别推动状态 x 和 y 朝着满足约束 $Ax=b$ 和 $Cy=d$ 的可行域 E 快速收敛。

为方便阅读，本文将采用下列记号来简化神经网络模型(8)，记：

$$z=\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}},\qquad I=\begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix},\qquad P=\begin{pmatrix} P_x & 0 \\ 0 & P_y \end{pmatrix},$$

$$F\left(z\right)=\begin{pmatrix} \nabla_x f\left(x,y\right) \\ -\nabla_y f\left(x,y\right) \end{pmatrix},\quad G=\begin{pmatrix} A^{\mathrm{T}}A & 0 \\ 0 & C^{\mathrm{T}}C \end{pmatrix},\quad H=\begin{pmatrix} A^{\mathrm{T}}b \\ C^{\mathrm{T}}d \end{pmatrix}.$$

神经网络(8)中的两个方程简化为如下：

$$\frac{\mathrm{d}z}{\mathrm{d}t}=\begin{pmatrix} \dfrac{\mathrm{d}x}{\mathrm{d}t} \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} \end{pmatrix}=-\left(I-P\right)F\left(z\right)-Gz+H. \tag{9}$$

定义 3.1：z^* 被称为神经网络(9)的平衡点，如果 z^* 满足以下方程：

$$0=-\left(I-P\right)F\left(z\right)-Gz+H.$$

4. 收敛性分析

在本节中，本文将分析神经网络(9)在 Lyapunov 意义下的稳定性和全局收敛性。首先证明若初始点不在可行域 E 中，则神经网络(9)的状态 $z\left(t\right)$ 将指数收敛到 E，其次证明神经网络(9)的平衡点与问题(2)的最优解是等价的，最后证明神经网络(9)在 Lyapunov 意义下是稳定的，并参考文献[26]，给出了两个充分条

件，使得神经网络(9)全局收敛到问题(2)的一个最优解。

定理 4.1：若假设 1 成立，当初始点 $z(0)=z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}\notin E$ 时，神经网络(9)的状态 $z(t)=\left(x(t)^{\mathrm{T}},y(t)^{\mathrm{T}}\right)^{\mathrm{T}}$ 将指数收敛到可行域 E 中。此外，如果初始点 $z_0\in E$，则对于 $\forall t\geq 0$，有 $z(t)\in E$ 成立。

证明：构造以下函数：

$$B(z)=B(x,y)=\frac{1}{2}\|Ax-b\|^2+\frac{1}{2}\|Cy-d\|^2, \tag{10}$$

易知 $B(z)$ 是一个可微的凸函数，且

$$\nabla_x B(z)=A^{\mathrm{T}}(Ax-b),\nabla_y B(z)=C^{\mathrm{T}}(Cy-d).$$

根据复合函数的链式法则可得

$$\begin{aligned}\frac{\mathrm{d}B(z(t))}{\mathrm{d}t}&=\nabla_x B(z(t))^{\mathrm{T}}\frac{\mathrm{d}x(t)}{\mathrm{d}t}+\nabla_y B(z(t))^{\mathrm{T}}\frac{\mathrm{d}y(t)}{\mathrm{d}t}\\&=(Ax-b)^{\mathrm{T}}A\left[-(I_m-P_x)\nabla_x f(x,y)-A^{\mathrm{T}}(Ax-b)\right]\\&\quad+(Cy-d)^{\mathrm{T}}C\left[(I_n-P_y)\nabla_y f(x,y)-C^{\mathrm{T}}(Cy-d)\right].\end{aligned} \tag{11}$$

由假设 1，A 是行满秩矩阵，易得矩阵 AA^{T} 为对称正定矩阵，则

$$\begin{cases}A(I_m-P_x)=A\left(I_m-A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}A\right)=A-A=0,\\C(I_n-P_y)=C\left(I_n-C^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C\right)=C-C=0.\end{cases}$$

令 λ_A 为矩阵 AA^{T} 的最小特征值，λ_C 为矩阵 CC^{T} 的最小特征值，且 $\lambda=\min\{\lambda_A,\lambda_C\}$，显然 $\lambda>0$。于是根据(11)，可得

$$\begin{aligned}\frac{\mathrm{d}B(z(t))}{\mathrm{d}t}&=-(Ax-b)^{\mathrm{T}}AA^{\mathrm{T}}(Ax-b)-(Cy-d)^{\mathrm{T}}CC^{\mathrm{T}}(Cy-d)\\&\leq-\lambda_A\|Ax-b\|^2-\lambda_C\|Cy-d\|^2\\&\leq-2\min\{\lambda_A,\lambda_C\}\left(\frac{1}{2}\|Ax-b\|^2+\frac{1}{2}\|Cy-d\|^2\right)\\&=-2\lambda B(z(t)),\end{aligned} \tag{12}$$

即

$$B(z(t))\leq \mathrm{e}^{-2\lambda t}B(z_0).$$

所以，当初始点 $z(0)=z_0=\notin E$ 时，神经网络(9)的状态向量 $z(t)=\left(x(t)^{\mathrm{T}},y(t)^{\mathrm{T}}\right)^{\mathrm{T}}$ 将指数收敛到可行域 E 中。

另一方面，如果 $z_0\in E$，则恒有 $B(z_0)=0$ 成立，故

$$0\leq B(z(t))\leq \mathrm{e}^{-2\lambda t}B(z_0)\leq 0,$$

即对于 $\forall t\geq 0$，

$$B(z(t))=0.$$

所以当初始点 $z_0\in E$ 时，对于 $\forall t\geq 0$，有 $z(t)\in E$ 成立。

定理 4.2 若假设 1 成立，则神经网络(9)的平衡点与问题(2)的最优解等价。

证明：假设 $z^*=\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}$ 是神经网络(9)的平衡点，根据定义 3.1 可知

$$0=-(I-P)F\left(z^*\right)-Gz^*+H, \tag{13}$$

即

$$\begin{cases}0=-\left(I_m-P_x\right)\nabla_x f\left(x^*,y^*\right)-A^{\mathrm{T}}\left(Ax^*-b\right),\\ 0=\left(I_n-P_y\right)\nabla_y f\left(x^*,y^*\right)-C^{\mathrm{T}}\left(Cy^*-d\right).\end{cases} \tag{14}$$

根据定理 4.1，若初始点 $z(0)=z_0\notin E$，则神经网络(9)的状态 $z(t)$ 将指数收敛到可行域 E 中，若初始点 $z_0\in E$，则对于 $\forall t\geq 0$，有 $z(t)\in E$ 成立。因此，可知 $z^*\in E$，由(14)可得

$$\begin{cases}0=\left(I_m-P_x\right)\nabla_x f\left(x^*,y^*\right),\\ 0=\left(I_n-P_y\right)\nabla_y f\left(x^*,y^*\right),\\ Ax^*=b,\\ Cy^*=d,\end{cases} \tag{15}$$

根据(15)，可知

$$\begin{cases}0=\left(I_m-P_x\right)\nabla_x f\left(x^*,y^*\right)=\nabla_x f\left(x^*,y^*\right)-A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}A\nabla_x f\left(x^*,y^*\right),\\ 0=\left(I_n-P_y\right)\nabla_y f\left(x^*,y^*\right)=\nabla_y f\left(x^*,y^*\right)-C^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C\nabla_y f\left(x^*,y^*\right).\end{cases}$$

令 $\lambda^*=\left(AA^{\mathrm{T}}\right)^{-1}A\nabla_x f\left(x^*,y^*\right)$，$\mu^*=\left(CC^{\mathrm{T}}\right)^{-1}C\nabla_y f\left(x^*,y^*\right)$，根据定理 3.1 可知，$z^*$ 是问题(2)的最优解。

另一方面，假设 $\bar{z}=\left(\bar{x}^{\mathrm{T}},\bar{y}^{\mathrm{T}}\right)^{\mathrm{T}}$ 是问题(2)的最优解，则 $\bar{z}\in E$（即 $A\bar{x}=b$，$C\bar{y}=d$）。由定理 3.1 可知

$$\begin{cases}-\left(I_m-P_x\right)\nabla_x f\left(\bar{x},\bar{y}\right)-A^{\mathrm{T}}\left(A\bar{x}-b\right)=\left(I_m-P_x\right)A^{\mathrm{T}}\bar{\lambda}-A^{\mathrm{T}}\left(A\bar{x}-b\right)=0,\\ \left(I_n-P_y\right)\nabla_y f\left(\bar{x},\bar{y}\right)-C^{\mathrm{T}}\left(C\bar{y}-d\right)=\left(I_n-P_y\right)C^{\mathrm{T}}\bar{\mu}-C^{\mathrm{T}}\left(C\bar{y}-d\right)=0,\end{cases} \tag{16}$$

即

$$-(I-P)F(\bar{z})-G\bar{z}+H=0.$$

根据定义 3.1 可知，$\bar{z}$ 是神经网络(9)的平衡点。

定理 4.3：若假设 1 成立，神经网络(9)的状态 $z(t)$ 在 Lyapunov 意义下是稳定的。此外，若满足下列条件之一[26]：

(A) 设 $z^*\in E^*=\left\{\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}\in E\,|\,\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}\text{ 是 }f(x,y)\text{ 的鞍点}\right\}$，$\forall z\in E$，存在一个函数 $k\left(z,z^*\right)\geq 0$，使得

$$\left(z-z^*\right)^{\mathrm{T}}F(z)=0\Rightarrow F(z)=k\left(z,z^*\right)F\left(z^*\right);$$

(B) 设 $z^*\in E^*$，$\forall z\in E$，使得

$$\left(z-z^*\right)^{\mathrm{T}}F(z)=0\Rightarrow z\in E^*.$$

对于任意初始值 $z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}\in E$，神经网络(9)的状态 $z(t)$ 全局收敛到它的一个平衡点。

证明：首先，证明神经网络(9)的状态 $z(t)$ 在 Lyapunov 意义下是稳定的。假设 $z^*=\left(\left(x^*\right)^{\mathrm{T}},\left(y^*\right)^{\mathrm{T}}\right)^{\mathrm{T}}$ 是鞍点问题(2)的最优解，则有 $z^*\in E$，由定理 4.2 可知，z^* 是神经网络(9)的平衡点，于是有下列式子成立：

$$0=-(I-P)F\left(z^*\right)-Gz^*+H=(I-P)F\left(z^*\right). \tag{17}$$

构造下列 Lyapunov 函数：

$$V(z)=\frac{1}{2}\left\|z-z^*\right\|^2, \tag{18}$$

显然，对于 $\forall z\in E$，$V(z)\geq 0$。此外，对于 $\forall\left(x^{\mathrm{T}},y^{\mathrm{T}}\right)^{\mathrm{T}}\in E$，有

$$\begin{cases}\left(x-x^*\right)^{\mathrm{T}}P_x=\left(x-x^*\right)^{\mathrm{T}}A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}A=\left(Ax-Ax^*\right)^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}A=0,\\ \left(y-y^*\right)^{\mathrm{T}}P_y=\left(y-y^*\right)^{\mathrm{T}}C^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C=\left(Cy-Cy^*\right)^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C=0,\end{cases}$$

即对于 $\forall z\in E$，

$$\begin{aligned}\left(z-z^*\right)^{\mathrm{T}}P&=\begin{pmatrix}x-x^*\\ y-y^*\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}P_x & 0\\ 0 & P_y\end{pmatrix}\\ &=\left(\left(x-x^*\right)^{\mathrm{T}}P_x \quad \left(y-y^*\right)^{\mathrm{T}}P_y\right)\\ &=0.\end{aligned} \tag{19}$$

根据(17)，(19)，可知

$$\begin{aligned}\left(z-z^*\right)^{\mathrm{T}}F\left(z^*\right)&=\left[\left(z-z^*\right)^{\mathrm{T}}P+\left(z-z^*\right)^{\mathrm{T}}(I-P)\right]F\left(z^*\right)\\ &=\left(z-z^*\right)^{\mathrm{T}}(I-P)F\left(z^*\right)\\ &=0,\end{aligned} \tag{20}$$

由于 $F(z)$ 是单调算子[30]，满足下列方程：

$$\left(z-z^*\right)^{\mathrm{T}}F(z)-\left(z-z^*\right)^{\mathrm{T}}F\left(z^*\right)\geq 0, \tag{21}$$

则根据(20)，(21)，可知，对于 $\forall z\in E$，有

$$\left(z-z^*\right)^{\mathrm{T}}F(z)\geq 0.$$

于是，对 Lyapunov 函数 $V(z)$ 求导，可得

$$\begin{aligned}\frac{\mathrm{d}V}{\mathrm{d}t}&=\left(x-x^*\right)^{\mathrm{T}}\frac{\mathrm{d}x}{\mathrm{d}t}+\left(y-y^*\right)^{\mathrm{T}}\frac{\mathrm{d}y}{\mathrm{d}t}\\ &=-\left(x-x^*\right)^{\mathrm{T}}\left(I_m-P_x\right)\nabla_x f(x,y)+\left(y-y^*\right)^{\mathrm{T}}\left(I_n-P_y\right)\nabla_y f(x,y)\\ &=-\left(z-z^*\right)^{\mathrm{T}}(I-P)F(z)\\ &=-\left(z-z^*\right)^{\mathrm{T}}F(z)\\ &\leq 0,\end{aligned} \tag{22}$$

即神经网络在 Lyapunov 意义下是稳定的。

其次证明，对于任意初始值 $z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}\in E$，神经网络(9)的状态 $z(t)$ 全局收敛到它的一个平衡点。令 $H(z)=\left(z-z^*\right)^{\mathrm{T}}F(z)$，由条件假设可知：

(a) 若 $H(z)=0$，$\forall z\in E$，根据条件(A)可得 $F(z)=k\left(z,z^*\right)F\left(z^*\right)$，则

$$\frac{\mathrm{d}z}{\mathrm{d}t}=-(I-P)F(z)=-k\left(z,z^*\right)(I-P)F\left(z^*\right)=0;$$

(b) 若 $H(z)=0$，$\forall z\in E$，根据条件(B)可得 $z\in E^*$，即 z 是问题(2)的一个最优解，根据定理 4.2 可

知，z 是神经网络(9)的一个平衡点。

因此，由给出的两个条件之一即可得，当 $H(z)=0$ 时，z 为神经网络(9)的平衡点。此外，根据(22)，对于任意初始点 $z(t_0)\in E$，有 $V(z(t))$ 随着 $t\to+\infty$ 非递增，且当 $z\to+\infty$ 时，根据(18)，易得 $V(z)\to+\infty$，即 $V(z)$ 是正定无界的。上述表明：

$$0\le V(z(t))\le V(z(0))<+\infty. \tag{23}$$

由(23)可知，$z(t)$ 是有界的，根据(9)可知，$\|\dot{z}(t)\|$ 也是有界的($\exists M\ge 0$，$\|\dot{z}(t)\|\le M$)，此外，存在一个极限点 $\bar{z}$ 和一个收敛序列

$$\{t_k \mid 0<t_1<t_2<\cdots<t_k<t_{k+1}<\cdots\}, \tag{24}$$

当 $k\to+\infty$ 时，$t_k\to+\infty$，使得 $z(t_k)\to\bar{z}$，显然，$\bar{z}\in E$。

接下来将证明 $H(\bar{z})=0$，则 $\bar{z}$ 是神经网络(9)的一个平衡点。假设 $H(\bar{z})>0$，由 $H(z)$ 的连续性可知，存在一个 $\varepsilon>0$ 和 $\delta>0$，当 $\|z-\bar{z}\|<\delta$ 时，有

$$H(z)>\varepsilon. \tag{25}$$

由(24)，存在一个正整数 N，使得 $\|z(t_k)-\bar{z}\|<\frac{\delta}{2}$ 对于所有的 $k>N$ 成立，所以当 $t\in\left[t_k-\frac{\delta}{4M},t_k+\frac{\delta}{4M}\right]$ 且 $k>N$ 时，有

$$\|z(t)-\bar{z}\|\le\|z(t)-z(t_k)\|+\|z(t_k)-\bar{z}\|\le M|t-t_k|+\frac{\delta}{2}\le\delta. \tag{26}$$

根据(25)可知，当 $t\in\left[t_k-\frac{\delta}{4M},t_k+\frac{\delta}{4M}\right]$ 且 $k>N$ 时，有 $H(z(t))>\varepsilon$。一方面，由于 $\bigcup_{k>N}\left[t_k-\frac{\delta}{4M},t_k+\frac{\delta}{4M}\right]$ 的测度是无限的，可得

$$\int_0^{+\infty}H(z(t))\mathrm{d}t>\int_0^{+\infty}\varepsilon\mathrm{d}t=+\infty. \tag{27}$$

另一方面，由(22)，(23)可知，$V(z(t))$ 单调递减且有界，即存在一个常数 V_0，使得 $\lim\limits_{t\to+\infty}V(z(t))=V_0$。根据(22)，有

$$\begin{aligned}\int_0^{+\infty}H(z(t))\mathrm{d}t&=\lim_{t\to+\infty}\int_0^t H(z(s))\mathrm{d}s\\&=-\lim_{t\to+\infty}\int_0^t\dot{V}(z(s))\mathrm{d}s\\&=-\lim_{t\to+\infty}\left[V(z(t))-V(z(0))\right]\\&=V(z(0))-V_0\\&<+\infty,\end{aligned} \tag{28}$$

与(27)矛盾，所以 $H(\bar{z})=0$，即 $\bar{z}$ 是神经网络(9)的一个平衡点。

最后，证明 $\lim\limits_{t\to\infty}z(t)=\bar{z}$，类似 $V(z)$ 的定义，定义另一个 Lyapunov 函数：

$$W(z)=\frac{1}{2}\|z-\bar{z}\|^2. \tag{29}$$

参考前面的证明，可以证明得到 $\frac{\mathrm{d}W}{\mathrm{d}t}\le 0$，$W(\bar{z})=0$ 且 $W(z)$ 是径向无界的。根据函数 $W(z)$ 的连续性，对于任意 $\varepsilon>0$，存在 $\delta>0$，使得 $\|z-\bar{z}\|\le\delta$，$W(z)<\frac{\varepsilon}{2}$。同时，由(24)可知，$z(t_k)\to\bar{z}$，并且存在一个正整数 t_N，使得 $\|z(t_N)-\bar{z}\|<\delta$，有

$$W\left(z\left(t_N\right)\right)=\left|W\left(z\left(t_N\right)\right)-W\left(\bar{z}\right)\right|<\frac{\varepsilon}{2}, \tag{30}$$

根据 $W\left(z\left(t\right)\right)$ 的单调非递增性和(29)，(30)，可知，对于 $\forall t>t_N$，

$$\left\|z\left(t\right)-\bar{z}\right\|^2=2W\left(z\left(t\right)\right)\le 2W\left(z\left(t_N\right)\right)<\varepsilon,$$

即

$$\lim_{t\to\infty} z\left(t\right)=\bar{z}.$$

因此对于任意初始点 $z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}\in E$，神经网络(9)的状态向量 $z\left(t\right)$ 将全局收敛到它的一个平衡点。

5. 数值模拟

本节对前面的理论分析提供两个数值算例来进行验证，展示本文所提出的神经网络的有效性和可行性。与文献[25]相比，本文不仅在单目标优化问题的研究基础上拓展成了 min-max 的鞍点问题，而且还继承了文献[25]中神经网络的指数收敛性和有效时间内收敛的特性。本节的所有数值实验均在 MATLAB 编程环境下进行。

例 5.1：考虑如下等式约束的鞍点优化问题：

$$\begin{aligned}&\min_x\max_y && f\left(x,y\right)=x_1^2+2x_2^2+3x_3^2+x_1y_1-x_2y_1+x_3y_2-\left(y_1^2+2y_2^2\right),\\ &\text{s.t.} && Ax=b,\\ &&& Cy=d,\end{aligned} \tag{31}$$

其中 $x\in\mathbb{R}^3$，$y\in\mathbb{R}^2$，A，C 均为行满秩矩阵，令：

$$A=\begin{pmatrix}1 & 1 & 1\end{pmatrix},\quad b=1,\quad C=\begin{pmatrix}1 & -1\end{pmatrix},\quad d=0.$$

显然，问题(31)的目标函数 f 关于 x 是凸函数(由于 Hessian 矩阵 $\nabla_{xx}^2 f=\mathrm{diag}\left(2,4,6\right)$ 正定)，关于 y 是凹函数(Hessian 矩阵 $\nabla_{yy}^2 f=\mathrm{diag}\left(-2,-4\right)$ 负定)，满足本文假设条件。通过求解 KKT 条件方程组，可得问题(31)的理论鞍点为 $x^*=\left[0.569,0.241,0.190\right]^{\mathrm{T}}$，$y^*=\left[-0.086,-0.086\right]^{\mathrm{T}}$，目标函数的最优解为 $f\left(x^*,y^*\right)=0.481$。

接下来本文将采用文中所描述的神经网络来进行求解，则与问题(31)所对应的神经网络方程(8)可改写为如下方程：

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=-\left(I_3-P_x\right)\nabla_x f\left(x,y\right)-A^{\mathrm{T}}Ax+A^{\mathrm{T}}b,\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=\left(I_2-P_y\right)\nabla_y f\left(x,y\right)-C^{\mathrm{T}}Cy+C^{\mathrm{T}}d,\end{cases} \tag{32}$$

其中

$$P_x=A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}A=\frac{1}{3}\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix},\quad P_y=C^{\mathrm{T}}\left(CC^{\mathrm{T}}\right)^{-1}C=\frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix},$$

$$\nabla_x f\left(x,y\right)=\begin{pmatrix}2x_1+y_1\\ 4x_2-y_1\\ 6x_3+y_2\end{pmatrix},\qquad \nabla_y f\left(x,y\right)=\begin{pmatrix}x_1-x_2-2y_1\\ x_3-4y_2\end{pmatrix}.$$

根据神经网络模型(32)，在 MATLAB 环境下进行仿真实验，随机生成 10 组初始点 $z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}$，其中 $x_0\in\mathbb{R}^3$，$y_0\in\mathbb{R}^2$，仿真时间 $t\in\left[0,2\right]$。数值结果如图 1 和图 2 所示。

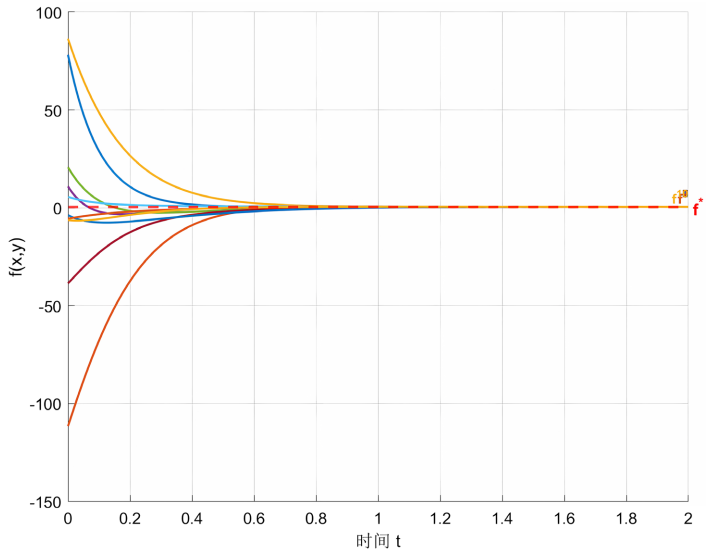

Figure 1. The convergence behavior plot of the objective function in Example 5.1
图 1. 例 5.1 中的目标函数收敛过程图

图 1 展示了例 5.1 中的目标函数值 $f\left(x(t),y(t)\right)$ 从不同初始值均收敛到最优值 f^* 的过程。

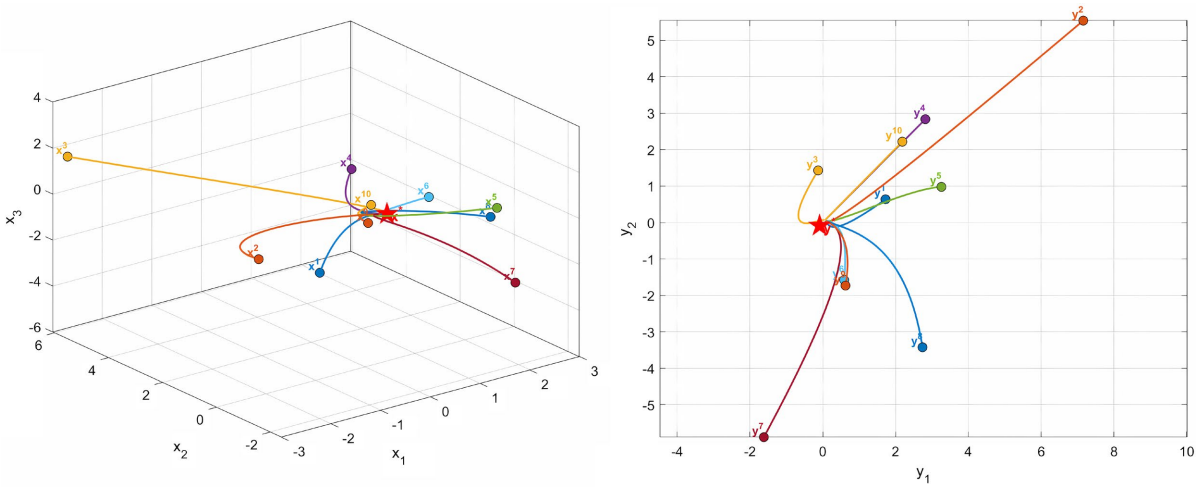

Figure 2. Convergence trajectories plot of each component of the state variables in Example 5.1
图 2. 例 5.1 中状态变量的每个分量收敛轨迹图

图 2 中的左右两图分别展示了 x，y 分量的收敛轨迹，10 条不同颜色的轨迹线从随机初始点出发，均收敛到点 x^* 和点 y^*。可以看出，尽管初始点分布在较广的范围内，所有轨迹均能快速收敛到理论最优值，验证了神经网络的全局收敛特性。

例 5.2：为了进一步验证神经网络(9)在处理复杂问题时的性能，本文考虑如下较高维度的鞍点问题：

$$\begin{aligned}\min_{x}\max_{y}\quad & f(x,y)=2x_1^2+3x_2^2+x_3^2+x_1x_2-2x_2x_3+2x_1y_1-x_2y_2+3x_3y_3\\ & -\left(y_1^2+2y_2^2+y_3^2\right)+y_1y_2-y_2y_3,\\ \text{s.t}\quad & Ax=b,\\ & Cy=d,\end{aligned}\tag{33}$$

其中 $x\in\mathbb{R}^3$， $y\in\mathbb{R}^3$，A，C 均为行满秩矩阵，令：

$$A=\begin{pmatrix}1&2&-1\\3&-1&2\end{pmatrix},\quad b=\begin{pmatrix}4\\1\end{pmatrix},\quad C=\begin{pmatrix}2&-1&1\\1&3&-2\end{pmatrix},\quad d=\begin{pmatrix}3\\-2\end{pmatrix}.$$

与例 5.1 中的问题(31)相比，问题(33)中目标函数的自变量相对维数更高一些，包含交叉项以及更复杂的非线性结构，更具有研究性，且关于 x 是凸函数，关于 y 是凹函数，满足本文的假设条件。通过求解 KKT 系统，得到该问题的理论鞍点为 $x^*=[1.293,0.845,-1.017]^{\mathrm{T}}$， $y^*=[0.976,-0.878,0.171]^{\mathrm{T}}$，目标函数 f 的最优解为 $f(x^*,y^*)=8.8457$ 。

将文中所提的神经网络(9)应用到问题(33)上，令 $z=(x^{\mathrm{T}},y^{\mathrm{T}})^{\mathrm{T}}$ 可得：

$$\frac{\mathrm{d}z}{\mathrm{d}t}=-(I-P)F(z)-Gz+H, \tag{34}$$

其中

$$I-P=\begin{pmatrix}\frac{1}{83}\begin{pmatrix}9&-15&-21\\-15&25&35\\-21&35&49\end{pmatrix} & 0\\ 0 & \frac{1}{75}\begin{pmatrix}1&-5&-7\\-5&25&35\\-7&35&49\end{pmatrix}\end{pmatrix},$$

$$F(z)=\begin{pmatrix}4x_1+x_2+2y_1\\6x_2+x_1-2x_3-y_2\\2x_3-2x_2+3y_3\\-2x_1+2y_1-y_2\\x_2+4y_2-y_1+y_3\\-3x_3+2y_3+y_2\end{pmatrix},G=\begin{pmatrix}10&-1&5&0&0&0\\-1&5&-4&0&0&0\\5&-4&5&0&0&0\\0&0&0&5&1&0\\0&0&0&1&10&-7\\0&0&0&0&-7&5\end{pmatrix},H=\begin{pmatrix}7\\7\\-2\\4\\-9\\7\end{pmatrix}.$$

同样，本文用改写后的神经网络(34)在 MATLAB 环境下采取随机生成 10 组初始点来进行数值模拟，数值仿真结果如图 3 和图 4 所示。与例 5.1 相比，该问题具有更高的维度和更复杂的动力学特性，但本文提出的神经网络仍能有效收敛。

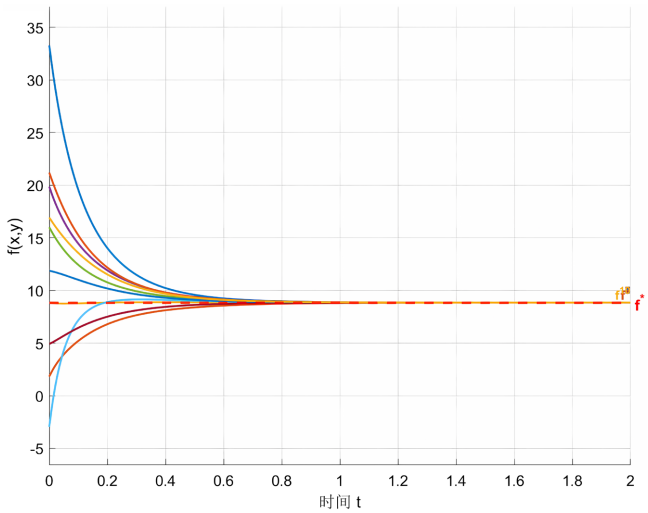

Figure 3. The convergence behavior plot of the objective function in Example 5.2
图 3. 例 5.2 中的目标函数收敛过程图

图 3 展示了例 5.2 中的目标函数值 $f\left(x(t),y(t)\right)$ 从不同初始值 $z_0=\left(x_0^{\mathrm{T}},y_0^{\mathrm{T}}\right)^{\mathrm{T}}$ 均收敛到最优值 f^* 的过程。

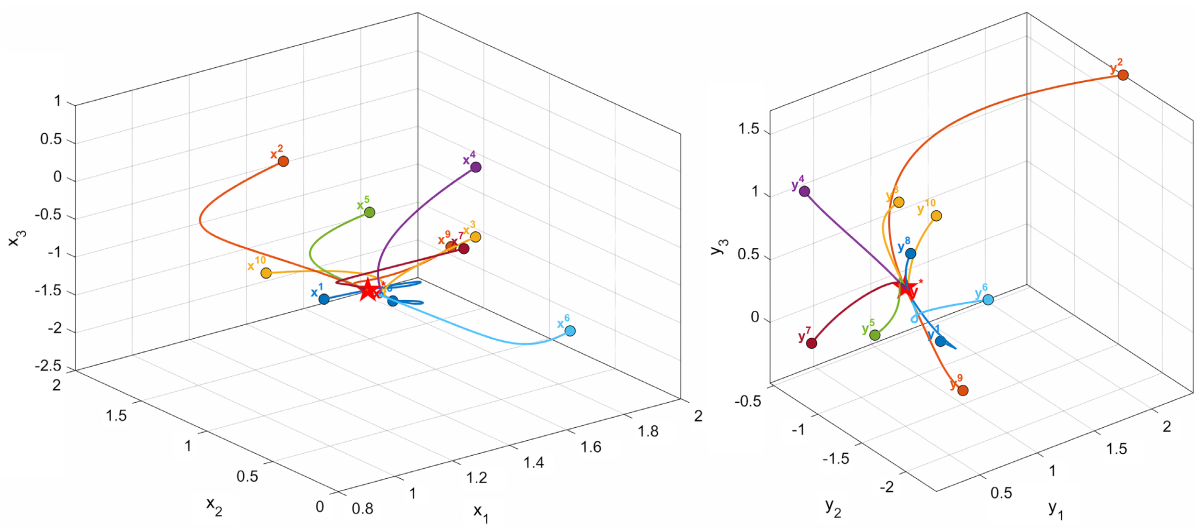

Figure 4. Convergence trajectories plot of each component of the state variables in Example 5.2
图 4. 例 5.2 中状态变量的每个分量收敛轨迹图

图 4 的左右两图分别展示了 x 和 y 分量的收敛轨迹，所有轨迹线从鞍点附近的扰动初始点出发，快速收敛到理论鞍点，验证了神经网络的快速收敛特性。

综上，本文对比了上面两个例子，表 1 总结了两个算例的收敛性能指标，包括平均迭代步数、约束精度和计算时间等。

Table 1. Comparison of neural network convergence performance
表 1. 神经网络收敛性能比较

算例	变量为度	平均收敛步数	约束精度	计算时间(s)	收敛速率
5.1	3 × 2	124	1.2×10^{-6}	0.45	指数收敛
5.22	3 × 2	89	8.7×10^{-7}	0.38	指数收敛

数值实验表明，本文提出的递归神经网络为等式约束鞍点问题提供了一种有效的求解方法，具有收敛快、精度高、易于实现等优点。

6. 总结

基于非线性规划在工程与经济等领域的重要应用背景，本文聚焦于等式约束下的鞍点问题这一核心课题。相较于 Qin 等人[25]针对单目标伪凸优化提出的递归神经网络，本文将其框架拓展至更具挑战性的 min-max 鞍点问题场景，实现了从单一优化到鞍点问题的理论跨越。本文构建了一种新的神经网络模型，在适当的假设下，证明了神经网络平衡点与鞍点问题最优解的等价性。进一步地，证明了神经网络在 Lyapunov 意义下是稳定的，在一些附加条件下，神经网络全局收敛到鞍点问题的最优解，为之后鞍点问题的研究提供了一种新的理论支持。

本文的理论分析基于若干关键假设，这些假设在保证理论严密性的同时，也带来了一定的应用局限

性，例如本文对目标函数的凸－凹结构和光滑性有较强的依赖。在此基础上，如何将此 RNN 框架扩展到非凸－非凹问题(可能需要引入更复杂的动态或正则化项)，以及如何处理非光滑的目标函数成为了未来的研究方向之一。

参考文献

[1] Ben-Tal, A., Ghaoui, L.E. and Nemirovski A. (2009) Robust Optimization. Princeton University Press.

[2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014) Generative Adversarial Nets. *Neural Information Processing Systems*, **27**, 2672-2680.

[3] Boyd, S. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine Learning*, **3**, 1-122. https://doi.org/10.1561/2200000016

[4] Zafar, M.B., Valera, I., Gomez Rodriguez, M. and Gummadi, K.P. (2017) Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment. *Proceedings of the 26th International Conference on World Wide Web*, Perth, 3-7 April 2017, 1171-1180. https://doi.org/10.1145/3038912.3052660

[5] Madry, A., Makelov, A., Schmidt, L., Tsipras, D. and Vladu, A. (2018) Towards Deep Learning Models Resistant to Adversarial Attacks. *International Conference on Learning Representations*, Vancouver.

[6] Arrow, K.J. and Debreu, G. (1954) Existence of an Equilibrium for a Competitive Economy. *Econometrica*, **22**, 265-290. https://doi.org/10.2307/1907353

[7] Arrow, K.J., Hurwicz, L. and Uzawa, H. (1958) Studies in Linear and Non-Linear Programming. Stanford University Press.

[8] Daskalakis, C., Ilyas, A., Syrgkanis, V. and Zeng, H.Y. (2018) Training GANs with Optimism. *International Conference on Learning Representations*, Vancouver.

[9] Cheng, Z., Ma, J., Wang, W., Zhu, Z., de Silva, C.W. and Lee, T.H. (2024) Alternating Direction Method of Multipliers-Based Parallel Optimization for Multi-Agent Collision-Free Model Predictive Control. *IEEE Transactions on Artificial Intelligence*, **5**, 4176-4191. https://doi.org/10.1109/tai.2024.3364127

[10] Gao, X.-B., Liao, L.-Z. and Xue, W. (2004) A Neural Network for a Class of Convex Quadratic Minimax Problems with Constraints. *IEEE Transactions on Neural Networks*, **15**, 622-628. https://doi.org/10.1109/tnn.2004.824405

[11] Gao, X. and Liao, L. (2006) A Novel Neural Network for a Class of Convex Quadratic Minimax Problems. *Neural Computation*, **18**, 1818-1846. https://doi.org/10.1162/neco.2006.18.8.1818

[12] Ghasabi-Oskoei, H., Malek, A. and Ahmadi, A. (2007) Novel Artificial Neural Network with Simulation Aspects for Solving Linear and Quadratic Programming Problems. Computers *& Mathematics with Applications*, **53**, 1439-1454. https://doi.org/10.1016/j.camwa.2006.10.022

[13] Gao, X.B. and Li, C.P. (2017) A New Neural Network for Convex Quadratic Minimax Problems with Box and Equality Constraints. *Computers & Chemical Engineering*, **104**, 1-10. https://doi.org/10.1016/j.compchemeng.2017.03.022

[14] Wang, Z.D., Liu, Y.R., Li, M.Z. and Liu, X.H. (2006) Stability Analysis for Stochastic Cohen-Grossberg Neural Networks with Mixed Time Delays. *IEEE Transactions on Neural Networks*, **17**, 814-820. https://doi.org/10.1109/tnn.2006.872355

[15] 喻昕, 陈昭蓉. 一类非光滑非凸优化问题的神经网络方法[J]. 计算机应用研究, 2019, 36(9): 2575-2578.

[16] 喻昕, 林植良. 解决一类非光滑伪凸优化问题的新型神经网络[J]. 计算机科学, 2022, 49(5): 227-234.

[17] Hopfield, J.J. (1982) Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences*, **79**, 2554-2558. https://doi.org/10.1073/pnas.79.8.2554

[18] 王雨嫣, 廖柏林, 彭晨, 等. 递归神经网络研究综述[J]. 吉首大学学报(自然科学版), 2021, 42(1): 41-48.

[19] Xia, Y. and Wang, J. (2005) A Recurrent Neural Network for Solving Nonlinear Convex Programs Subject to Linear Constraints. *IEEE Transactions on Neural Networks*, **16**, 379-386. https://doi.org/10.1109/tnn.2004.841779

[20] Liu, Q.S. and Wang, J. (2011) A One-Layer Recurrent Neural Network for Constrained Nonsmooth Optimization. *IEEE Transactions on Systems*, *Man*, *and Cybernetics*, *Part B (Cybernetics)*, **41**, 1323-1333. https://doi.org/10.1109/tsmcb.2011.2140395

[21] Tank, D. and Hopfield, J. (1986) Simple 'Neural' Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit. *IEEE Transactions on Circuits and Systems*, **33**, 533-541. https://doi.org/10.1109/tcs.1986.1085953

[22] Kennedy, M.P. and Chua, L.O. (1988) Neural Networks for Nonlinear Programming. *IEEE Transactions on Circuits and Systems*, **35**, 554-562. https://doi.org/10.1109/31.1783

[23] Zhang, S. and Constantinides, A.G. (1992) Lagrange Programming Neural Networks. *IEEE Transactions on Circuits and Systems II*: *Analog and Digital Signal Processing*, **39**, 441-452. https://doi.org/10.1109/82.160169

[24] Xia, Y. and Wang, J. (2004) A General Projection Neural Network for Solving Monotone Variational Inequalities and Related Optimization Problems. *IEEE Transactions on Neural Networks*, **15**, 318-328. https://doi.org/10.1109/tnn.2004.824252

[25] Qin, S., Bian, W. and Xue, X. (2013) A New One-Layer Recurrent Neural Network for Nonsmooth Pseudoconvex Optimization. *Neurocomputing*, **120**, 655-662. https://doi.org/10.1016/j.neucom.2013.01.025

[26] Gao, X.B. (2003) Exponential Stability of Globally Projected Dynamic Systems. *IEEE Transactions on Neural Networks*, **14**, 426-431. https://doi.org/10.1109/tnn.2003.809409

[27] Murray, R.M., Li, Z. and Sastry, S.S. (2017) A Mathematical Introduction to Robotic Manipulation. 1st Edition, CRC Press. https://doi.org/10.1201/9781315136370-1

[28] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.

[29] 张从军. 集值分析与经济应用[M]. 北京: 科学出版社, 2004.

[30] 唐昊. 不等式约束鞍点问题的罚函数方法[D]: [硕士学位论文]. 重庆: 重庆交通大学, 2024.