

求解约束 Minimax 问题的神经网络模型

陶 卿

方廷健

(中国科学技术大学电子工程与信息科学系·合肥, 230027) (中国科学院合肥智能机械研究所·合肥, 230031)

摘要: 本文提出一种求解约束二次 Minimax 问题的神经网络模型, 给出了它的 Lyapunov 能量函数, 运用 LaSalle 不变性原理证明了它的大范围渐近稳定性, 作为应用考察了 L_1 范数极小化问题。

关键词: Minimax 问题; 神经网络模型; 大范围渐近稳定性; L_1 范数极小化问题

文献标识码: A

0225

对策理论
数学规划

The Neural Network Model for Solving Minimax Problems with Constraints

TAO Qing

(Department of Electronic Engineering & Information Science, University of Science & Technology of China · Hefei, 230027, P. R. China)

FANG Tingjian

(Hefei Institute of Intelligent Machines, Academia Sinica · Hefei, 230031, P. R. China)

Abstract: This paper presents a kind of neural network models for solving quadratic minimax problems with constraints, and shows its global asymptotic stability by using LaSalle invariant principle. The L_1 -norm minimizing problems are considered as applications.

Key words: minimax problems; neural network models; global asymptotic stability; L_1 -norm minimizing problems

1 引言(Introduction)

Minimax 问题是对策理论研究的主要对象, 它在军事筹划和自动控制中有重要应用。在处理实际问题时, 往往需要求得数值解, 因此, 设计一种实时的求解方法是非常必要的。从数学的观点看, Minimax 问题是数学规划问题的自然推广。近年来, 利用神经网络求解数学规划问题取得了很大的进展(见文献[1~4])。人们自然想到能否用神经网络求解 Minimax 问题, 1997 年, 文献[5]建立一种模型, 但其所解决的问题是无约束的。

本文将提出一种求解约束 Minimax 问题的神经网络模型, 证明了它的大范围渐近稳定性。当目标函数是二次泛函时, 文献[5]的结论是本文无约束时的特例, 作为应用还考察了 L_1 范数极小化问题(见文献[6,7]), 此时我们的模型和文献[7]一样。

2 Minimax 问题(Minimax problem)

考虑下述 Minimax 问题

$$\min_{x \in U_1} \max_{y \in U_2} \frac{1}{2} x^T A_1 x + \frac{1}{2} y^T A_2 y + x^T B y + a^T x + a^T y. \quad (1)$$

其中 A_1 和 A_2 分别是对称半正定和半负定 $m \times m$ 矩阵, Q_1, Q_2 和 $Q \subset \mathbb{R}^m$ 是闭凸集, $x = (x_1, x_2, \dots, x_m)^T$, 以下假定问题(1)的鞍点集非空, 记 (x^*, y^*) 是问题(1)的一个鞍点, $\|x\|$ 表示 x 的欧几里德范数。

引理 1 对 $\forall x \in \mathbb{R}^m$, 存在唯一一点 $P(x)$, 满足

$$\|x - P(x)\| = \inf_{y \in Q} \|x - y\|.$$

称 P 为 Q 上的投影算子, 且对

$$\forall \hat{x} \in \mathbb{R}^m, x_0 \in Q,$$

$$\langle \hat{x} - x_0, x - x_0 \rangle \leq 0, \forall x \in Q$$

的充要条件是

$$x_0 = P(\hat{x}).$$

投影算子 P 还满足

$$\|P(x) - P(y)\| \leq \|x - y\|, \forall x, y \in \mathbb{R}^m.$$

证 见文献[8]的第二章定理 2。

类似于文献[4]引理 3 和定理 1, 可得下列结论:

引理 2 (x^*, y^*) 是问题(1)的鞍点的充要条

件是

$$\langle A_1 x^* + B y^* + a_1, x - x^* \rangle \geq 0, \quad \forall x \in Q_1,$$

$$\langle A_2 y^* + B^T x^* + a_2, y - y^* \rangle \leq 0, \quad \forall y \in Q_2.$$

定理 1 记 P_1 和 P_2 分别表示 Q_1 和 Q_2 上的投影算子, (x^*, y^*) 是问题(1)的解的充要条件是

$$x^* = P_1(x^* - A_1 x^* - B y^* - a_1),$$

$$y^* = P_2(y^* + A_2 y^* + B^T x^* + a_2).$$

3 神经网络模型及其动态分析 (The neural network model and its dynamic analysis)

我们提出以下连续神经网络模型来求解二次 Minimax 问题(1).

$$\begin{cases} \frac{dx}{dt} = (I + A_1)(P_1(x - A_1 x - B y - a_1) - x) - \\ \quad B(P_2(y + A_2 y + B^T x + a_2) - y), \\ \frac{dy}{dt} = (I + A_2)(P_2(y + A_2 y + B^T x + a_2) - y) + \\ \quad B^T(P_1(x - A_1 x - B y - a_1) - x). \end{cases} \quad (2)$$

定理 2 令

$$V_1(x, y) = \frac{1}{2} \|x - x^*\|^2,$$

$$V_2(x, y) = \frac{1}{2} \|y - y^*\|^2,$$

$$V(x, y) = V_1(x, y) + V_2(x, y),$$

则

$$\begin{aligned} \dot{V}(x, y) &\leq -\langle A_1(x - x^*), x - x^* \rangle + \\ &\quad \langle A_2(y - y^*), y - y^* \rangle - \\ &\quad \|x - P_1(x - A_1 x - B y - a_1)\|^2 - \\ &\quad \|y - P_2(y + A_2 y + B^T x + a_2)\|^2, \end{aligned} \quad (3)$$

证

$$\begin{aligned} \dot{V}_1(x, y) &= \langle x - x^*, \frac{dx}{dt} \rangle = \\ &\langle x - x^*, (I + A_1)(P_1(x - A_1 x - B y - a_1) - x) \rangle - \\ &\langle x - x^*, B(P_2(y + A_2 y + B^T x + a_2) - y) \rangle. \end{aligned}$$

用类似于文献[4]引理 3 的方法

$$\begin{aligned} \langle x - x^*, (I + A_1)(P_1(x - A_1 x - B y - a_1) - x) \rangle &= \\ &= -\langle A_1(x - x^*), x - x^* \rangle - \\ &\quad \|x - P_1(x - A_1 x - B y - a_1)\|^2 - \\ &\quad \langle B(y - y^*), P_1(x - A_1 x - B y - a_1) - x^* \rangle - \\ &\quad \langle x - x^*, B(P_2(y + A_2 y + B^T x + a_2) - y) \rangle - \\ &\quad \langle A_1 x^* + B y^* + a_1, P_1(x - A_1 x - B y - a_1) - x^* \rangle. \end{aligned}$$

根据引理 1 和引理 2 可得

$$\begin{aligned} \dot{V}_1(x, y) &\leq \\ &= -\langle A_1(x - x^*), x - x^* \rangle - \\ &\quad \|x - P_1(x - A_1 x - B y - a_1)\|^2 - \\ &\quad \langle B(y - y^*), P_1(x - A_1 x - B y - a_1) - x^* \rangle - \\ &\quad \langle B(y - y^*), x - x^* \rangle - \\ &\quad \langle x - x^*, B(P_2(y + A_2 y + B^T x + a_2) - y) \rangle. \end{aligned} \quad (4)$$

同理

$$\begin{aligned} \dot{V}_2(x, y) &\leq \\ &\langle A_2(y - y^*), y - y^* \rangle - \\ &\quad \|y - P_2(y + A_2 y + B^T x + a_2)\|^2 + \\ &\quad \langle B^T(x - x^*), P_2(y + A_2 y + B^T x + a_2) - y \rangle + \\ &\quad \langle B^T(x - x^*), y - y^* \rangle + \\ &\quad \langle y - y^*, B^T(P_1(x - A_1 x - B y - a_1) - x) \rangle. \end{aligned} \quad (5)$$

(4)和(5)式相加即得(3)式. 证毕.

定理 3 i) 神经网络模型(2)大范围渐近收敛于问题(1)的鞍点集.

ii) 若 A_1 正定, A_2 负定, 问题(1)有唯一鞍点 (x^*, y^*) , 且存在 $c^* > 0$, 使

$$\frac{1}{2} \|x(t) - x^*\|^2 + \frac{1}{2} \|y(t) - y^*\|^2 \leq$$

$$O(e^{-c^* t}), \quad t \rightarrow +\infty.$$

证 i) $\forall (x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}^m$, 由引理 1 知投影算子是 Lipschitz 映射, 此时可设 $(x(t), y(t))$ 是模型(2)以 (x_0, y_0) 为初值的解, 其最大存在区间为 $[0, \beta(x_0))$, 令

$$G = \{(x, y); V(x, y) \leq \frac{1}{2} \|x_0 - x^*\|^2 + \frac{1}{2} \|y_0 - y^*\|^2, (x, y) \in \mathbb{R}^m \times \mathbb{R}^m\}.$$

根据定理 2 知, $V(x, y)$ 是 G 上(2)的 Lyapunov 函数(见文献[10]), 且 $x(t) \in G$. 显然 G 为有界集, 根据常微分方程解的存在性理论(见文献[9]), 必有 $\beta(x_0) = +\infty$.

根据 LaSalle 不变性原理(见文献[10])存在常数 r , 使

$$(x(t), y(t)) \rightarrow M \cap V^{-1}(r), \quad t \rightarrow +\infty \text{ 时.}$$

其中 M 是 $E = \{(x, y); \dot{V}(x, y) = 0, (x, y) \in \bar{G}\}$ 的最大不变集.

根据定理 2 和定理 1 知 M 中的每一个点都是问题(1)的一个鞍点. 由 (x_0, y_0) 的任意性, i) 得证.

ii) 若 A_1 正定, A_2 负定, 则存在 $c_1, c_2 > 0$, 使

$$\langle A_1(x - x^*), x - x^* \rangle \geq c_1 \|x - x^*\|^2,$$

$$-\langle A_2(y - y^*), y - y^* \rangle \geq c_2 \|y - y^*\|^2.$$

令 $c^* = 2(c_1 + c_2)$, 根据定理 2 知问题 (1) 有唯一鞍点 (x^*, y^*) , 且

$$\frac{d}{dt}(e^{c^* t} V(x(t), y(t))) =$$

$$e^{c^* t} (\dot{V}(x(t), y(t)) + c^* V(x(t), y(t))) \leq 0.$$

故 $e^{c^* t} V(x(t))$ 是 $0 \leq t < +\infty$ 上的减函数, 因此

$$\frac{1}{2} \|x(t) - x^*\|^2 + \frac{1}{2} \|y(t) - y^*\|^2 \leq$$

$$V(x(t), y(t)) \leq V(x(0), y(0))e^{-c^* t},$$

得证.

注 当 Q_1 和 Q_2 为超闭方体时, 模型 (2) 可由电路方便地实现, 其原理与文献 [3] 和文献 [7] 完全相同. 当问题 (1) 没有约束, 即 P_1 和 P_2 为恒等算子时, 模型 (2) 与文献 [5] 目标函数是二次泛函时模型相同.

例 1 考虑下述混合约束下的半正定二次规划问题 (见文献 [3])

$$\min \quad \frac{1}{2} x^T A x + a^T x$$

$$\text{s.t.} \quad D x = b, \quad x \in Q.$$

根据凸规划的鞍点定理, 例 1 等价于

$$\min_{x \in Q} \max_{y \in \mathbb{R}^n} \left(\frac{1}{2} x^T A x + a^T x + y^T (D x - b) \right).$$

此时模型 (2) 与文献 [3] 的模型相同.

例 2 考虑下述 L_1 范数极小化问题 (见文献 [6])

$$\min \|A x - b\|_{L_1}.$$

根据文献 [7], 例 2 等价于

$$\min_{x \in \mathbb{R}^n} \max_{\|y\|_\infty \leq 1} y^T (A x - b).$$

此时模型 (2) 与文献 [6] 的模型相同.

参考文献 (References)

- [1] Tank D W and Hopfield J J. Simple "neural" optimization networks [J]. IEEE Trans. Circuits and Systems, 1986, 35(5): 533-541
- [2] Bouzerdoum A and Pettison T R. Neural network for quadratic optimization with bound constraints [J]. IEEE Trans. Neural Network, 1993, 4(2): 293-304
- [3] Xia Y. A new neural network for solving linear and quadratic problems [J]. IEEE Trans. Neural Network, 1996, 7(6): 1544-1547
- [4] 陶卿, 王常波, 方廷健. 一种求解闭凸集上二次规划问题的神经网络模型 [J]. 模式识别与人工智能, 1998, 11(1): 83-87
- [5] 叶仲泉, 张邦礼, 曹长修. Minimax 神经网络收敛性分析 [J]. 信息与控制, 1997, 26(1): 1-6
- [6] 张青富, 保铮. 求解超定线性方程组 L_1 -范数最优解的神经网络方法 [J]. 电子学报, 1996, 24(1): 97-101
- [7] 夏又生, 叶大振. 解 L_1 范数极小化的神经网络 [J]. 电子学报, 1997, 25(11): 99-101
- [8] 张学铭, 李训经, 陈祖浩. 最优控制系统的常微分方程理论 [M]. 北京: 高等教育出版社, 1989
- [9] 郭太均. 非线性泛函分析 [M]. 济南: 山东科技出版社, 1985
- [10] LaSalle J P. The Stability of Dynamical Systems [M]. Philadelphia: SIAM, 1976

本文作者简介

陶卿 1965 年生, 博士, 中国人民解放军炮兵学院作战指挥实验中心副教授. 现在中国科技大学做博士后研究工作. 研究领域包括: 神经网络系统理论, SVM 理论和应用数学.

方廷健 1939 年生, 博士生导师, 中国科学院合肥智能机械研究所研究员. 研究方向为智能信息处理.