

基于吸引域的总体极小化问题的神经网络求解

张全举¹, 曲小钢¹, 陈开周²

(1. 西安建筑科技大学 理学院, 陕西 西安 710055; 2. 西安电子科技大学 理学院, 陕西 西安 710071)

摘要: 神经网络求解优化问题具有非常强大的实时计算功用, 因此近年来受到了密切的关注. 这里考察了求解无约束总体极小化问题的神经网络方法, 提出了一种新的网络求解模型. 从基于吸引域分析方法为出发点证明了所给网络平稳点集合的全局吸引性. 分析了网络的电路实现, 并估计了各个平稳点的吸引域. 这些理论分析与估计是构造所提神经网络模型的依据, 同时也是网络可靠运行的基础. 此外, 数值模拟试验也充分揭示了这个网络模型在实际运行中都能够很好地求解总体极小化问题, 是一个十分有效的神经网络系统. 这里的结果表明: 这里提出的网络模型无论从理论上还是实际运行中都能够可靠且稳定地求解总体极值问题, 基于吸引域构造神经网络的方法是一种很有潜力的神经网络求解优化问题的研究方向.

关键词: 总体极值; 神经网络; 全局吸引性; 吸引域

中图分类号: TP273 **文献标识码:** A

Neural networks based on the attractive region for solving global minimization

ZHANG Quan-ju¹, QU Xiao-gang¹, CHEN Kai-zhou²

(1. School of Science, Xi'an University of Architecture & Technology, Shanxi Xi'an 710055, China;

2. School of Science, Xidian University, Shanxi Xi'an 710071, China)

Abstract: Neural networks for solving optimisation problems own very robust real time computation ability and much attention had paid for it in recent years. Neural network method for solving global unconstrained minimized problems was investigated and a new neural network model was then proposed. Starting from method on analysis of attractive region, the global attraction of equilibrium point set was demonstrated for the proposed model. The circuit realization of the given model was analysed and the attraction region for individual equilibrium was estimated also. All these theoretical analyses and estimation results were the foundations for constructing the proposed neural network model. At the same time, they also provide the basic supports for the network's reliable running. In addition, numerical experimental results revealed enough evidence to show that the proposed neural network model could work well in practice on finding solutions for the global minimized problems, which means the model was a very efficient neural network system. The results showed that the proposed neural network model could conduct reliably and stably, whether in theory or in practice, to solve the global minimized problems and hence methods based on the attractive region of neural network construction is a promising research direction for solving optimisation problems.

Key words: global optimization; neural network; global attractability; attractive region

用神经网络求解优化问题是近几年来较活跃的研究领域, 其实时计算的强大功能日益受到科学界的关注. 当目标函数是凸函数或能够化为凸函数时, 网络能 Lyapunov 全局渐近稳定地收敛于最优解, 如对线性规划^[1~3]、凸规划^[4~7]问题. 本文将考虑无约束总体极小化问题的神经网络求解.

1 网络及其吸引性 (Neural networks and its attractability)

对 $\mathbb{R}^n$ 上的总体极小化问题: $\min f(x), x \in \mathbb{R}^n$, 若 $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ 具有足够的光滑性, 则该问题有解的一个充分条件是:

引理 1 若连续函数 $f(x)$ 的某水平集 $S(f, \alpha) = \{x \mid f(x) \leq \alpha\}$ 非空有界, 则 $f(x)$ 的总体极小值点集非空有界.

证 由 $S(f, \alpha)$ 非空有界及 f 的连续性知其 $S(f, \alpha)$ 上必有最小值点 x^* , 该点也是 f 在 $\mathbb{R}^n$ 上的

总体极小值点.再由 $S(f, \alpha)$ 的有界性可知 f 的总体极小值点集有界.

$$\text{考虑以 } f(x) \text{ 的负梯度 } -\nabla f(x) = -\left\{ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right\} \text{ 为向量场的微分方程}$$

$$\frac{dx}{dt} = -\nabla f(x) \quad (1)$$

所刻划的神经网络,其吸引性定义为

定义 1 记网络(1)的平稳点集为 $M = \{x \mid \nabla f(x) = 0\}$.若对网络(1)的任何从 x 出发的解曲线 $x(t) (x(0) = x \in \mathbb{R}^n)$,均有某一 $p \in M$ 使 $\lim_{t \rightarrow \infty} x(t) = p$,即集合 M 吸引任何轨线,则称网络(1)是全局吸引的.

假定 $f(x)$ 的所有水平集 $S(f, c)$ 有界且其平稳点集 M 由有限个点 $p_1, p_2, \dots, p_m$ 构成,则有

定理 1 网络(1)在上述条件下是全局吸引的.

证 任取 $x \in \mathbb{R}^n$, 由 $\dot{f}|_{(1)} = \frac{df(x(t))}{dt} = -\|\nabla f(x(t))\|^2 \leq 0$ 知 $f(x(t)) \leq f(x(0)) = f(x)$, 故 $x(t) \in S(f, f(x))$, 即从紧集 $S(f, f(x))$ 内出发的解恒在其内.由微分方程解的延拓定理知 $x(t)$ 的存在区间为 $(-\infty, +\infty)$, 再由 $x(t)$ 的有界性知其 ω 极限集 $\Omega(x)$ 是一连通不变紧集.任取 $y \in \Omega(x)$, 由 $\Omega(x)$ 的不变性知对所有 $t \geq 0$ 有 $y(t) \in \Omega(x)$, 故 $f(y) = \inf_{t \geq 0} f(x(t)) = f_0$ (由 f 沿 $x(t)$ 递减及其连续性知 f_0 存在), 从而, 过 y 的轨线满足 $\dot{f}|_{(1)} = 0$ 及 $\nabla f(y) = 0$, 因此, y 是式(1)的平稳点.

由 M 中点的孤立性以及 $\Omega(x)$ 的连通性可知, $\Omega(x)$ 仅由一点组成, 即对所有 $t_n \rightarrow \infty, x(t_n)$ 的极限点唯一, 故 $\lim_{t \rightarrow \infty} x(t)$ 存在且等于某 $p_i \in M, 1 \leq i \leq m$. 证毕.

注 1 此吸引性结论对于与式(1)对偶的网络

$$\frac{dx}{dt} = \nabla f(x) \quad (2)$$

的负向吸引性 $\lim_{t \rightarrow -\infty} x(t) \in M$ 同样成立, 这是因为式(1)的 ω 极限点正好是式(2)的 α 极限点, 反之亦然.

由此及定理 1 可得如下的推论.

推论 1 对网络 2 的负向吸引性有 $x(t) \rightarrow p \in M$ 或 $x(t) \rightarrow \infty$ 成立.

2 初值控制(Control of initial values)

上节的吸引性结论保证了网络的可靠运行, 即任一初值输入都有一个确定的平稳点输出. 下面分析各平稳点 $p \in M$ 的吸引域 $B(p) = \{x \in \mathbb{R}^n \mid \lim_{t \rightarrow \infty} x(t)$

$= p\}$, 以便确定初值的输入范围, 以搜索平稳点.

I) 若 p 是极大值点, 则 $B(p) = \{p\}$. 因为 p 是极大值点, 故存在某邻域 $U(p)$, 对任意 $x \in U(p) - \{p\}$ 有 $f(x) < f(p)$. 若某轨线 $f(t)$ 在 $t^* > 0$ 时进入 $U(p)$, 则有 $f(x(t)) \leq f(x(t^*)) < f(p)$, 故不可能有 $\lim_{t \rightarrow \infty} x(t) = p$.

为进一步分析, 需引入正不变集的概念. 若对集 P 中每个点 x 及所有 $t \geq 0$, 网络(1)或(2)的流 $\Phi_t(x)$ 有定义且在 P 中, 则称集 P 关于系统(1)或(2)是正向不变的.

II) 若 p 是极小值点, 记 $D_c(p) = \{x \mid \|x - p\| \leq c < r\}$, 其中 r 是距 p 最近的平稳点到 p 的距离. 假定 $D_c(p)$ 是正向不变的, 则有 $D_c(p) \subset B(p)$.

考虑 $D_c(p)$ 上的定正函数 $V(x) = f(x) - f(p)$, 易知对 $x \in D_c(p) - \{p\}$ 有 $\dot{V}(x)|_{(1)} = \dot{f}(x)|_{(1)} = -\|\nabla f(x(t))\|^2 < 0$, 即 $V(x)$ 是 $D_c(p)$ 上的严格 Lyapunov 函数, 故由经典的 Lyapunov 稳定性理论知 p 在 $D_c(p)$ 内是 Lyapunov 渐近稳定的. 又因 $D_c(p)$ 是正向不变的, 从而对任一 $x \in D_c(p)$ 都有 $\lim_{t \rightarrow \infty} x(t) = p$ 成立.

注 2 若对 $c > 0, V(x) = c$ 是有界闭曲面, 在其内部 $\dot{V}(x) < c$, 则 $\{x \mid V(x) \leq c\} \subset B(p)$, 它一般会较 $D_c(p)$ 大一些.

III) p 是鞍点的情况比较复杂. 首先 p 点不可能是吸引的: 在 p 的无论多么小的邻域 U 内, 总有 $x \in U$, 使 $f(x) < f(p)$, 从而过 p 点的轨线当 $t \geq 0$ 时有 $f(x(t)) \leq f(x) < f(p)$, 对此 x , 不可能有 $\lim_{t \rightarrow \infty} x(t) = p$.

不排除对某些特别的点此式成立, 这种对部分初值的吸引性称为条件吸引性. 为方便起见, 设 $p = 0$. 考虑网络(1)在鞍点处的线性化网络

$$\frac{dx}{dt} = -H(0)x + \varphi(x). \quad (3)$$

其中: $H(0)$ 是 f 在 0 点的 Hessian 矩阵, $\varphi(x) = -\nabla f(x) + H(0)x$.

由

$$\begin{aligned} \varphi(x) &= -\nabla f(x) + H(0)x = \\ &= -\int_0^1 H(tx)x dt + \int_0^1 H(0)x dt \end{aligned}$$

可得

$$\begin{aligned} \|\varphi(x)\| &= \\ \left\| \int_0^1 (H(0) - H(tx))x dt \right\| &\leq \end{aligned}$$

$$\int_0^1 \|H(tx) - H(0)\| \|x\| dt \leq$$

$$M \|x\|, x \in \{x \mid \|x\| \leq K\}.$$

该估计式和下述引理2说明网络(3)的吸引性决定于常系数线性网络 $\frac{dx}{dt} = -H(0)x$, 而此线性网络的吸引性又决定于 $H(0)$ 的特征值.

引理2^[8] 若对 $\forall \varepsilon > 0, \exists \delta > 0$, 当 $\|x\| \leq \delta$ 时, $\|g(t, x)\| \leq \varepsilon \|x\|$, 则 $\frac{dx}{dt} = A(t)x + g(t, x)$ 的指数吸引性由 $\frac{dx}{dt} = A(t)x$ 确定.

令 $A(t) = -H(0)$, $g(x, t) = \varphi(x)$, 可知网络(3)是这里的特例.

现设 $H(0)$ 有 k 个正特征值、 $n - k$ 个负特征值, 则由上述讨论及常系数线性方程组解的结构知, 在 $p = 0$ 的某邻域中存在 k 维流形 S_k^+ 和 $n - k$ 维流形 S_{n-k}^- , 使得若 $x(0) \in S_k^+$, 则 $\lim_{t \rightarrow \infty} x(t) \rightarrow 0$; 若 $x(0) \in S_{n-k}^-$, 则 $\lim_{t \rightarrow -\infty} x(t) \rightarrow 0$, 即 $S_k^+ \subset B(0)$.

总结以上 I) ~ III) 可得

定理2 平稳点 p 的吸引域 $B(p)$ 有如下估计:

1) 若 p 是极大值点, 则 $B(p) = \{p\}$;

2) 若 p 是极小值点且 $D_c(p)$ 是正不变集, 则 $D_c(p) \subset B(p)$;

3) 若 p 是鞍点, 则存在 p 点的 k 维流形 $S_k^+(p)$, 使得 $S_k^+(p) \subset B(p)$ (当 $t \rightarrow +\infty$) 及 $n - k$ 维流形 $S_{n-k}^-(p)$, 使得 $S_{n-k}^-(p) \subset B(p)$ (当 $t \rightarrow -\infty$).

结合定理2与注1不难推出如下关于网络(2)的相应结论.

推论2 网络(2)的平稳点 p 的吸引域 $B(p)$ 有如下估计:

1) 若 p 是极大值点且 $D_c(p)$ 是负不变集, 则 $D_c(p) \subset B(p)$;

2) 若 p 是极小值点, 则 $B(p) = \{p\}$;

3) 若 p 是鞍点, 则存在 p 点的 $n - k$ 维流形 $S_{n-k}^+(p)$, 使得 $S_{n-k}^+(p) \subset B(p)$ (当 $t \rightarrow +\infty$) 及 k 维流形 $S_k^-(p)$, 使得 $S_k^-(p) \subset B(p)$ (当 $t \rightarrow -\infty$).

注3 由定理2可知, 当 $x \in \mathbb{R}^n$ 不是 f 的平稳点时, 以 x 为初值运行网络(1)将输出一个极小值点 p 或鞍点 p . 为了确定 p 附近的平稳点, 根据定理2的式(3), 可在 p 附近取所有形如 $x = (p_1, \dots, p_i, p_{i+1} + \delta, \dots, p_{i+k} + \delta, \dots, p_n)$ 的点 (共有 $2^n - 2$ 个) 为初始输入, 交替运行网络(1)和(2), 可将 p 点附近的所有平稳点输出. 当然, 可能有部分鞍点漏网, 但这并不影响问题的解决. 对新得的平稳点输入上述初始点, 再次交替运行网络(1)和(2)即可得到新的平稳点. 由于已假定

问题的平稳点的个数是有限的, 依此法将求出所有平稳点 p_i , 比较诸 $f(p_i)$ 的值即可求得最小值点 p^* .

这种网络模式的块状构造见图1, 其中, 各单元由线积分器和反馈梯度网络组成.

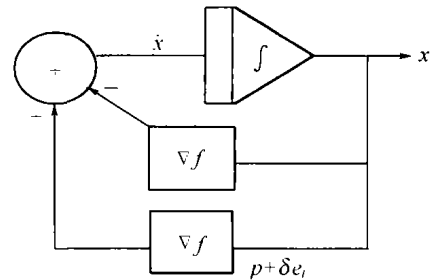


图1 网络流程图

Fig. 1 Diagram of main networks

3 数值模拟(Numerical simulations)

3.1 计算步骤(Steps)

步骤1 任选初值, 运行网络(1), 输出一个平稳点.

步骤2 在该平稳点附近另取初值运行网络(1). 若输出仍为该点, 则可断定其为极小值点, 记录函数值后转步骤3; 若输出不为该点, 则可断定其为鞍点, 此时, 重复步骤2对所得输出点进行判定, 记录极小值后转步骤3.

步骤3 以上两步所获得的平稳点附近的点作为初值运行网络(2), 记录输出点后转步骤2.

步骤4 当不再有新的平稳点输出时, 即可比较诸极小值, 输出总体极小值.

注意, 运行网络(2)时, 可事先选定一充分大的数, 若函数值大于该数, 则另取初值.

3.2 数值算例(Examples)

首先考虑 Goldstein-Price 函数

$$f(x, y) = [1 + (x + y + 1)^2(19 - 14x + 3x^2 - 14y + (6xy + 3y^2))][30 + (2x - 3y)^2(18 - 32x + 12x^2 + 48y - 36xy + 27y^2)].$$

它有4个极小值点(图2): $B(1.2, 0.8)$, $D(1.8, 0.2)$, $E(-0.6, -0.4)$, $F(0, -1)$; 1个极大值点 $C(0.8, 0.2)$; 此外, 该函数还有4个鞍点. 计算时, 以 $A(0.5, 0.9)$ 为初值, 搜索轨线见图2, 图中箭头表示搜索方向.

作为第2个例子, 考虑6峰驼背函数

$$g(x, y) = 4y^2 - 2.1y^4 + \frac{1}{3}y^6 + xy - 4x^2 + 4x^4.$$

它有6个极小值点(图3): $E(-1.61, -0.57)$, $F(1.70, -0.80)$, $G(1.61, 0.57)$, $H(-1.70, 0.79)$, $K(-0.09, 0.71)$, $L(0.09, -0.71)$, 2个极大值点

$I(-1.23, -0.17), J(1.23, 0.17)$. 计算时, 以 $(\pm 4, \pm 4)$ 4个点为初值, 搜索轨线见图3. 限于版面, 图中仅画出了搜索轨线进入图示范围内的那一部分.

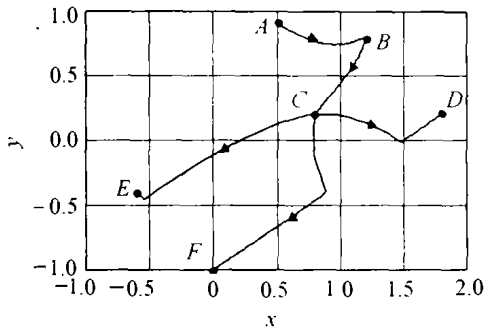


图2 Goldstein-Price函数的计算机模拟图示
Fig. 2 Diagram of Goldstein-Price function

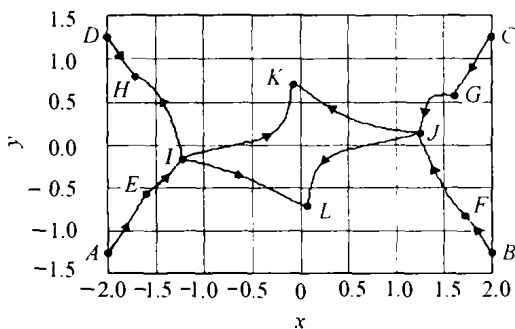


图3 6峰驼背函数的计算机模拟图示
Fig. 3 Diagram of 6-humped function

4 结论(Conclusions)

提出了一个求解优化问题的新的神经网络方法, 分析了网络平稳点集合的总体吸引性, 得到了一个基本定理; 估计了各平稳点的吸引域, 并利用吸引域构造了相应的神经网络模型. 计算机数值模拟证实了本文提出的网络能够稳定且有效地解决优化问题.

参考文献(References):

[1] XIA Youshen. A new neural network for solving linear programming

problems and its applications [J]. *IEEE Trans on Neural Networks*, 1996, 7(3): 643-651.

[2] XIA Youshen. Neural networks for solving extended linear programming problems [J]. *IEEE Trans on Neural Networks*, 1997, 8(6): 519-525.

[3] 张青富, 焦李成, 保铮. 一种新的求解线性规划的神经网络[J]. *电子学报*, 1992, 20(10): 44-49.

(ZHANG Qingfu, JIAO Li-cheng, BAO Zheng. A new neural network for solving linear programming [J]. *J of Electronics*, 1992, 20(10): 44-49.)

[4] WANG Jun. A deterministic annealing neural network for convex programming [J]. *IEEE Trans on Neural Networks*, 1994, 7(4): 629-641.

[5] MAA C Y, SHANBLAT M A. Linear and quadratic programming neural network analysis [J]. *IEEE Trans on Neural Networks*, 1992, 3(4): 580-594.

[6] CHUA L C, LIN G N. Nonlinear programming without computation [J]. *IEEE Trans on Circuits and Systems*, 1984, 31(2): 182-188.

[7] WANG Jun, XIA Youshen. Analysis and design of primal-dual assignment networks [J]. *IEEE Trans on Neural Networks*, 1998, 9(1): 183-193.

[8] 廖晓昕. 稳定性的理论、方法和应用[M]. 武汉: 华中理工大学出版社, 1998.

(LIAO Xiaoxin. *Theory, Methods and Applications of Stability* [M]. Wuhan: Hua Zhong University Press, 1998.)

作者简介:

张全举 (1964—), 男, 2002年毕业于西安电子科技大学应用数学系, 获博士学位, 西安建筑科技大学副教授. 主要研究方向为最优化理论与算法以及控制工程等. E-mail: zhangquanju@sina.com;

曲小钢 (1959—), 男, 1990年毕业于西安建筑科技大学理学院, 获硕士学位, 西安建筑科技大学副教授, 中国计算物理学会理事. 主要研究方向为计算物理, 优化与控制等. E-mail: qxg7828@sina.com;

陈开周 (1930—), 男, 博士生导师, 西安电子科技大学教授. 主要研究方向为最优化理论与神经网络等.