第 29 卷 第 3 期 西北农林科技大学学报(自然科学版) Vol. 29 No. 3
2001 年 6 月 Jour. of Northwest Sci-Tech Univ. of Agri. and For. (Nat. Sci. Ed.) June 2001

凸函数极小值点集与梯度神经网络的极限点集

冯芙叶

(西安科技学院 基础部,西安 710054)

[摘 要] 给出了凸函数极小值点集的几何特征,它是 $\mathbf{R}^n$ 中的一个单连通子集。用神经网络求解优化问题,必须考察的问题是网络的极限点集结构;对梯度神经网络的极限点集进行详细分析,主要结果是对凸函数来说网络的极限点集就是该函数的极小值点集,而这恰是梯度网络求解凸函数总体极值时,网络能够全局稳定收敛的条件。

[关键词] 神经网络;凸函数;极小值点集;极限点集

[中图分类号] O221 [文献标识码] A [文章编号]1000-2782(2001)03-115-02

近年来用神经网络求解优化问题的研究非常活跃,这些网络模型都是由常微分方程所刻划的网络系统[1~4]。一般方法是将网络的平衡点集与目标函数的最优解集对等,在 Lyapunov 意义下说明网络的平衡点集是渐近稳定的,以期达到求解最优解的目的,文献[1～4]就是以此模式考虑了线性规划、二次规划以及凸规划的神经网络求解。由于网络是由常微分方程所刻划的,对网络的极限点集分析就显得很重要,上面所提到的文献对这个问题却未涉及。本文通过对凸函数极小值点集的详细分析,得到了一种广泛应用的梯度神经网络的极限点集,即凸函数最优解集的结论。

考虑以可微凸函数 $f(x)$ 之负梯度 $-\nabla f(x)$ 为向量场的网络系统

$$\frac{\mathrm{d}x}{\mathrm{d}t}=-\nabla f(x), \tag{1}$$

这里 f 是开凸集 $W\in\mathbf{R}^n$ 上的 $\mathbf{C}^2$ 凸函数,从而(1)式满足解的存在唯一性,f 凸是指[5]:对任意 $x^1,x^2\in D$ 及任意 $\lambda_1,\lambda_2\in[0,1]$, $\lambda_1+\lambda_2=1$ 有 $f(\lambda_1x^1+\lambda_2x^2)\leqslant\lambda_1f(x^1)+\lambda_2f(x^2)$。一个一般的 n 维网络系统是指

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x), \tag{2}$$

这里 $f:W\to\mathbf{R}^n$ 是映开集 $W\subset\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的局部 Lipschitz 映射,系统(2)式的平衡点是指 $f(x)=0$ 的点,其过 $x\in W$ 的解曲线 $x(t)$,$x(0)=x$ 的 ω 极限点集定义为

$$\Omega=\{y\mid y=\lim_{n\to\infty}x(t_n),t_n\to\infty(n\to\infty)\},$$

函数 f 的最优解集定义为

$$S^*=\{x^*\mid f(x^*)\leqslant f(x),x\in W\}。$$

关于凸函数 f 最优解集的几何结构以及它与(1)式的极限点集之间的关系是本文讨论的主要问题,其中几何特征及极限点集的非空性结果由下面的定理所描述。

定理 1 设 S^* 为 f 的最优解集,则有

① $S^*=\varphi$ 或 S^* 是 W 的单连通子集;

② 设 f 的极小值点集 S^* 非空,则由(1)式定义的神经网络任何轨线的 ω 极限点集 Ω 非空。

本文的主要结果(定理 1 与定理 2)的证明以如下几个引理为基础。

引理 1[5] x^* 是可微凸函数 $f(x)$ 的总体极小值点当且仅当 $\nabla f(x^*)=0$。

引理 2[5] 凸函数 f 的总体极小值点集 S^* 是凸集。

引理 3 对所有 $x\in W$,$f(x)$ 沿(1)式的解曲线的导数 $\dot f(x)\leqslant0$;并且 $\dot f(x)=0$,且 $x\in S^*$。这里 $\dot f(x)=\frac{\mathrm{d}f(x(t))}{\mathrm{d}t}$,其中 $x(t)$ 是(1)式的满足 $x(0)=x$ 的解曲线。

证明 由 $\dot f(x)=\frac{\mathrm{d}f(x(t))}{\mathrm{d}t}=\left(\nabla f(x),\frac{\mathrm{d}x}{\mathrm{d}t}\right)=-\|\nabla f(x)\|^2$,这里(. , .)表示内积,知 $\dot f(x)\leqslant0$;而 $\dot f(x)=0$ 当且仅当 $\|\nabla f(x)\|=0$,即 $\nabla f(x)=0$,故 $x\in S^*$,引理证毕。

[收稿日期] 2000-09-18
[作者简介] 冯芙叶(1963-),女,陕西大荔人,讲师,主要从事微分方程与最优化理论的教学和研究工作。

引理4[5] 设 f 是 $W\in\mathbf{R}^n$ 上可微凸函数,则对任意 $x,y\in W$ 有

$$(y-x)^T[\nabla f(y)-\nabla f(x)]\geqslant 0。$$

定理1的证明 若 $S^*\neq\varphi$,则由引理2知 S^* 是凸集。任取 $x^1,x^2\in S^*$,函数 $H(\theta)=\theta x^1+(1-\theta)x^2$ 是连接 x^1,x^2 的连续弧,从而 S^* 是弧连通集,而 $\mathbf{R}^n$ 中的弧连通集必为单连通集,定理的①获证。

任取 $x^*\in S^*$,由引理1知 $\nabla f(x^*)=0$,考虑函数

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

沿(1)式的导数

$$\dot{V}=(\nabla V(x),\frac{\mathrm{d}x}{\mathrm{d}t})=(x-x^*,-\nabla f(x))=$$
$$-(x-x^*)^T\nabla f(x)=$$
$$-(x-x^*)^T[\nabla f(x)-\nabla f(x^*)]。$$

由引理4知 $\dot{V}\leqslant 0$,即 V 沿(1)式的解曲线是递减的,故任何从紧集

$$G_H=\{x:\|x-x^*\|^2\leqslant H\}$$

内出发的解 $x(t)$ 满足

$$V(x(t))\leqslant V(x(0))=\frac{1}{2}\|x(0)-x^*\|^2=$$
$$\frac{1}{2}\|x-x^*\|^2\leqslant\frac{1}{2}H,$$

因此,由微分方程解的延拓定理知 $x(t)$ 在 $[0,\infty)$ 上有定义;而且上式说明 $x(t)$ 在紧集 $G_H=\{x:\|x-x^*\|^2\leqslant H\}$ 内,再由 Weierstrass 聚点原理知 Ω 非空,定理的②获证。

关于 f 的极小值点集与(2)的极限点集之间的关系有下面定理2的结果。

定理2 设 f 的极小值点集 S^* 非空,则有 $S^*=\Omega$。

证明 设 $x^*\in S^*$,则由引理1知 $\nabla f(x^*)=0$。从而对所有 $t\geqslant 0$,$x(t)=x^*$ 是(2)的平凡解,故 $x^*\in\Omega$;反之,设 x^* 是某解曲线 $x^0(t)$ 之 ω 极限点,从而存在 $t_n\to\infty,n\to\infty$,使 $x^*=\lim\limits_{n\to\infty}x^0(t_n)$。由 S^* 非空知 $f(x)$ 有下界,而由引理3可知 $f(x^0(t_n))$ 单调不增有下界,因此极限 $f_0=\lim\limits_{n\to\infty}f(x^0(t_n))$ 存在,由连续性得 $f(x^*)=f(\lim\limits_{n\to\infty}x^0(t_n))=f_0$,即 f 在 Ω 上取常数,故对 $x\in\Omega$ 有 $\dot{f}(x)=0$,再由引理3知 $x^*\in S^*$,定理2证毕。

[参考文献]

[1] Maa C Y, Shanblatt M A. Linear and quadratic programming neural network analysis[J]. IEEE Trans Neural Networks, 1992, 3(4): 580-594.

[2] Wang J. A deterministic annealing neural network for convex programming[J]. Neural Networks, 1994, 7(4): 629-641.

[3] More J J, Toroaldo G. On the solution of large quadratic programming problems with bound constraints[J]. SIAM J Optimization, 1991, 1(1): 93-113.

[4] Xia Y S, Wang J. A general methodology for designing globally convergent optimization neural networks[J]. IEEE Trans Neural Networks, 1998, 9(6): 1331-1343.

[5] Avriel M. Nonlinear Programming[M]. New Jersey: Prentice-Hall, Inc, 1976.

[6] 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.

The minimizer set of the convex function and the limit point set of neural network governed by gradient of the function

FENG Fu-ye

(Department of Basic Courses, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: The geometric symptoms of the minimizer set of the convex function are described in this paper. Analysis of the limit set for the gradient neural network is also conducted. The result being obtained here shows eminently that the two sets are exactly the same set.

Key words: neural network; convex function; minimizer set; limit set