DOI:10.16383/j.aas.2003.05.016
第29卷 第5期
2003年9月
自 动 化 学 报
ACTA AUTOMATICA SINICA
Vol. 29, No. 5
Sept., 2003

基于精确罚函数的一类广义非线性神经网络模型[1)]

孟志青[1]　胡奇英[1]　杨晓琪[2]
[1](西安电子科技大学经济管理学院　西安　710071)
[2](香港理工大学应用数学系　香港)
(E-mail: mengzhiqing@xtu.edu.cn)

摘　要　针对一般的非线性优化问题定义了一种2次非线性罚函数，证明了在一定条件下对应的罚优化问题的精确罚定理，由此引进了一种广义非线性神经网络模型，并证明了这种网络的平衡点与能量函数之间的联系，在一定条件下对应的平衡点收敛到原问题的最优解.这种神经网络模型对于求解许多优化问题具有重要的作用.

关键词　神经网络，非线性罚函数，最优解，平衡点，稳定点
中图分类号　O122.2

A General Model of Non-Linear Neural Networks Based on Exact Penalty Function

MENG Zhi-Qing[1]　HU Qi-Ying[1]　YANG Xiao-Qi[2]
[1](*School of Economics and Management, Xidian University, Xi'an*　710071)
[2](*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong*)
(E-mail: mengzhiqing@xtu.edu.cn)

Abstract　A double non-linear penalty function is defined for the non-linear optimality problems (NP) and the exact penalty theorem is exacted under some conditions. A new general model of non-linear neural networks is introduced and the relationship between the equilibrium points and the energy function is showed. Under the given condition, the equilibrium point of the neural networks converges to a solution of NP. This model plays an important part in many optimal problems.

Key words　Neural networks, non-linear penalty function, optimal solution, equilibrium point, stable point

1) 国家自然科学基金(69904008)资助
Supported by National Natural Science Foundation of P. R. China(69904008)
收稿日期　2001-04-11　　收修改稿日期　2001-11-15
Received April 11, 2001; in invised form November 15, 2001

1 引言

自 Hopfield 引进反馈式神经网络用来求解优化问题取得成功以后,用神经网络来研究最优化问题成为当前国际上的研究前沿性课题. 神经网络解决优化问题的过程主要是侧重于能量函数和动力系统的构造,以保证系统的平衡点与优化问题的最优解对应. Hopfield 网络成功地解决了一些约束优化问题,如旅行商问题、最短路问题和分类问题等. 许多研究所引进的神经网络模型都类似于 Hopfield 模型. 近年来,许多研究集中在研究非线性神经网络模型用来求解含约束的优化问题,如文[1~4]. 神经网络在解决含约束优化问题,主要是利用罚函数的方法来构造能量函数与相应的微分动力系统,最大的问题是网络的平衡点不一定是优化问题的最优解. 在文[1]中 Xia 和 Wang 建立了一类广义神经网络,主要证明了在凸性条件下的网络平衡点收敛到优化问题最优解. 非线性罚函数[5~7]是近年来解决非线性优化的一个新的工具,其理论研究基本趋于成熟. 本文借助一个 2 次非线性罚函数,建立了一种新的神经网络模型,主要证明平衡点可以收敛到优化问题的可行解,这个可行点满足 Fritz John 条件,在凸线性条件下它是最优解.

2 非线性罚函数法的精确性

近年来,非线性罚函数的研究取得了很大进展,非线性罚函数具有良好的特性,如光滑性和精确性,为解决约束优化问题提供了新的途径. 下面讨论非线性优化问题(NP).

(NP) $\min f(\boldsymbol{x})$

$$\text{s.t. } \boldsymbol{x}\in X,\ g_i(\boldsymbol{x})\leqslant 0,\ i=1,2,\cdots,m,\ h_j(\boldsymbol{x})=0,\ j=1,2,\cdots,q.$$

设 $g_i^+(\boldsymbol{x})=\max\{0, g_i(\boldsymbol{x})\}$, $i\in I$. 设可行解集

$$X_0=\{\boldsymbol{x}\in X \mid g_i(\boldsymbol{x})\leqslant 0,\ i=1,2,\cdots,m,\ h_j(\boldsymbol{x})=0,\ j=1,2,\cdots,q\}.$$

考虑下面(NP)对应的非线性光滑罚函数

$$F(\boldsymbol{x},\rho)=\Big(f(\boldsymbol{x})^2+\rho\sum_{i\in I}g_i^+(\boldsymbol{x})^2+\rho\sum_{j\in J}h_j(\boldsymbol{x})^2\Big)^{1/2},$$

其中 $\rho\geqslant 0$. 如果 f, $g_i(\forall i\in I)$ 和 $h_j(\forall j\in J)$ 是连续可微的,显然 $F^2(\boldsymbol{x},\rho)$ 也是连续可微的. 下面均设 $f(\boldsymbol{x})$ 是非负函数,考虑对应的罚函数优化问题

(P(ρ)) $\min F(\boldsymbol{x},\rho)$, s.t. $\boldsymbol{x}\in X$.

定理 1. 假设 X_0 是紧集,函数 f 满足 $\lim\limits_{\boldsymbol{x}\to\infty,\ \boldsymbol{x}\in X} f(\boldsymbol{x})=+\infty$,则

$$\max_{\rho\in R_+^1}\min_{\boldsymbol{x}\in X}F(\boldsymbol{x},\rho)=\min_{\boldsymbol{x}\in X_0}f(\boldsymbol{x}).$$

设 $b(\boldsymbol{x})=\max\{g_i(\boldsymbol{x}), h_j(\boldsymbol{x}), -h_j(\boldsymbol{x}) \mid i=1,2,\cdots,m;\ j=1,2,\cdots,q\}$,那么 $b(\boldsymbol{x})\leqslant 0$ 等价于 $\boldsymbol{x}$ 是(NP)的可行解.

定义 1. 如果 $\boldsymbol{x}^*$ 是(NP)的最优解,我们称 $\boldsymbol{y}^*$ 和 $\boldsymbol{z}^*$ 是强 Lagrange 乘子向量,如果以下公式成立

$$\nabla f(\boldsymbol{x}^*)=-\sum_{i\in I}y_i^*\nabla g_i(\boldsymbol{x}^*)-\sum_{j\in J}z_j^*\nabla h_j(\boldsymbol{x}^*) \tag{1}$$

$$y_i^*g_i(\boldsymbol{x}^*)=0,\ y_i^*\geqslant 0,\ g_i(\boldsymbol{x}^*)\leqslant 0,\ i=1,2,\cdots,m,\ z_j^*\geqslant 0,\ h_j(\boldsymbol{x}^*)=0,\ j=1,2,\cdots,q \tag{2}$$

定理 2. 设 $\boldsymbol{x}^*$ 是(NP)的最优解,$(\boldsymbol{y}^*, \boldsymbol{z}^*)\in R^m\times R^q$ 是(NP)的强 Lagrange 乘子向量.

又设 X 是凸集，f，$g_i(\forall i\in I)$和 $h_j(\forall j\in J)$是可微凸函数，$\boldsymbol{x}_\rho^*$ 是(P(ρ))的最优解. 则

1) $f(\boldsymbol{x}^*)=F(\boldsymbol{x}_\rho^*,\rho)$，若 $b(\boldsymbol{x}_\rho^*)\leqslant 0$;

2) $f(\boldsymbol{x}^*)=F(\boldsymbol{x}_\rho^*,\rho)$，若 $b(\boldsymbol{x}_\rho^*)>0$，$\rho\geqslant\max\left\{\lambda^2(m+q)^2,\ \dfrac{2(m+q)\lambda f(\boldsymbol{x}^*)}{b(\boldsymbol{x}_\rho^*)}\right\}$和 $\lambda=\max\{y_i^*,\ z_j^*\ |\ i=1,\cdots,\ m;\ j=1,2,\cdots,q\}$.

定理 1 和定理 2 的证明见附录. 定理 1 和定理 2 说明罚函数 $F(x,\rho)$在一定条件下是一个精确罚函数，即存在一个罚参数 ρ 使得罚问题(P(ρ))的最优解也是原问题(NP)的最优解. 定理 2 的条件是对凸情况下的精确罚定理，文[2] 证明了对于非凸情况下的精确罚函数定理，但验证条件成立比较困难. 另外一个重要的事实是非线性罚函数与非线性神经网络有着密切的联系. 下面我们来进一步研究这个问题.

3 非线性神经网络模型

设向量函数 $b: R^n\times R^1\to R^n$，考虑动力学微分系统

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=Bb(\boldsymbol{x},t) \tag{3}$$

其中 $\boldsymbol{x}(t)=(x_1(t),x_2(t),\cdots,x_n(t))^{\mathrm{T}}$，$B=\mathrm{diag}(\beta_1,\ \beta_2,\ \cdots,\ \beta_n)$和 $\beta_i>0(i=1,2,\ \cdots,\ n)$为收敛系数，$\boldsymbol{x}$ 是状态变量.

定义 **2**. 如果点 $\bar{\boldsymbol{x}}\in X$ 满足 $b(\bar{\boldsymbol{x}},t)=0(\forall t)$，则称 $\bar{\boldsymbol{x}}$ 是动力系统(3)的平衡点.

定义 **3**. 设点 $\bar{\boldsymbol{x}}\in X$，如果对任给的 ε 和 $t_0\geqslant 0$，存在 $\delta(\varepsilon,t_0)>0$，当 $t\geqslant t_0$，$\|\boldsymbol{x}_0-\bar{\boldsymbol{x}}\|<\delta(\varepsilon,t_0)$时，有 $\|\boldsymbol{x}(t;x_0,t_0)-\bar{\boldsymbol{x}}\|<\varepsilon$，其中 $\boldsymbol{x}(t;x_0,t_0)$表示以 $\boldsymbol{x}_0$为起点 t_0为起始时间满足系统(3)的一条参数曲线，则称 $\bar{\boldsymbol{x}}$ 是动力系统(3)稳定点.

下面讨论求解问题(NP)的广义神经网络模型，(NP)的能量函数定义为

$$E(\boldsymbol{x})=F(\boldsymbol{x},\rho)^2=f(\boldsymbol{x})^2+\rho\sum_{i\in I}g_i^+(\boldsymbol{x})^2+\rho\sum_{j\in J}h_j(\boldsymbol{x})^2 \tag{4}$$

对应的微分动力系统定义为

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=-B\ \text{镩}\ WP\ MathALApE(\boldsymbol{x}) \tag{5}$$

这种神经网络模型非常类似于文[1] 的模型. 其实，Hopfield 的神经网络模型具有 2 个显著的特征：状态微分等于能量函数的负梯度和神经元连接权系数等于能量函数的 2 次偏导数(Hesse 矩阵). 因此，Hopfield 的神经网络模型的动力微分系统与能量函数也有与式(5)一样的关系式. 设 n 个神经元通过状态 $x_i(t)$两两相连，如 $E(\boldsymbol{x})$具有 2 次微分，那么网络的连接权系数定义为

$$w_{ij}=\frac{\partial^2E(\boldsymbol{x})}{\partial x_i\partial x_j},\quad i,j=1,2,\cdots.$$

定理 **3**. 假设 $\bar{x}$ 是系统(5)的在罚参数 ρ 下的平衡点，对于 $\boldsymbol{x}\neq 0$ 有 $E(\boldsymbol{x})\neq 0$，则 $\bar{\boldsymbol{x}}$ 是动力系统(5)的稳定点. 进一步，如果 $\bar{\boldsymbol{x}}$ 对应的权系数矩阵$(w_{ij})_{n\times n}$是半正定的，则 $\bar{\boldsymbol{x}}$ 是罚优化问题(P(ρ))的局部最优解.

证明. 由假设，对于 $\boldsymbol{x}\neq 0$，有 $E(\boldsymbol{x})>0$，因为

$$\frac{\mathrm{d}E(\boldsymbol{x})}{\mathrm{d}t}=\sum_{k=1}^n\frac{\partial E}{\partial x_k}\frac{\mathrm{d}x_k(t)}{\mathrm{d}t}=\sum_{k=1}^n\frac{\partial E}{\partial x_k}B\left(-\frac{\partial E}{\partial x_k}\right)\leqslant 0.$$

根据 Lyapunov 稳定性定理知 $\bar{x}$ 是动力系统(5)的稳定点，第 2 个结论显然成立. 证毕.

定理 4. 设 $\bar{x}$是罚优化问题(P(ρ))的一个最优解,则 $\bar{x}$是系统(5)的在罚参数 ρ 下的平衡点.

定理 4 显然成立. 定理 3 和定理 4 给出了罚优化问题的最优解与动力系统的平衡点的之间的联系.我们容易证明下面定理和推论成立.

定理 5. 假设函数 f, g_i($\forall i\in I$)和 h_j($\forall j\in J$)均是连续可导函数, $\lim\limits_{x\to\infty, x\in X} f(x)=+\infty$. 设$\{\rho_k\}$是趋于无穷大的正数序列. 假设序列$\{x^k\}$是对应的动力系统(4)在罚参数 ρ_k下的平衡点,对应的值$\{E(x^k)\}$有界, 则序列$\{x^k\}$的任意聚点 $\bar{x}$是有界的并且是(NP)的可行解,存在数 λ, α_1, α_2, …, α_q和$\mu_i\geqslant 0$, $i=1,2,\cdots,m$ 满足

$$\lambda \text{镝 } WP\ MathALApf(\bar{x})+\sum_{i\in I}\mu_i \text{镝 } WP\ MathALApg_i(\bar{x})+\sum_{j\in J}\alpha_j \text{镝 } WP\ MathALAph_j(\bar{x})=0 \tag{6}$$

推论 1. 又设 X 是凸集, f, g_i($\forall i\in I$)和 h_j($\forall j\in J$)是可微凸函数, 假设定理5 的条件全部成立,且 $\lambda>0$, μ_i, $_j\alpha\geqslant 0$, $i\in I$, $j\in J$, 则序列$\{x^k\}$的任意聚点 $\bar{x}$是(NP)的最优解.

定理 5 给出了网络平衡点与原问题之间的联系,式(6)是 Fritz John 条件[8]. 特别是在凸性条件下,推论 1 表明平衡点可以导出原问题的最优解. 因此,可以给出下面的一个算法.

Step1. 给定开始点 x^k, ρ_0, $\varepsilon>0$, $N>1$, 置 $k=0$.

Step2.

1) 使用同一个开始点 x^k, 设 $y^0=x^k$, 置 $j=0$;

2) 如果一镝 $WP\ MathALApE(y^j)\neq 0$, 令 $D=-$镝 $WP\ MathALApE(y^j)$, 存在 t 使得 $E(y^j+Dt)<E(y^j)$, 令 $y^{j+1}=y^j+Dt$, 置 $j+1=j$, 重复 Step2,否则转入 3);

3) 设 $x^k=y^j$, 即 y^j 是动力系统的平衡点,转入 Step3.

Step3. 如果 $|E(x^{k+1})-E(x^k)|<\varepsilon$, 设 $\rho_{k+1}=N^{\rho_k}$, 置 $k:=k+1$, 重复 Step2.

注. 我们使用该算法进行了一些数值实验,表明算法在较大的罚参数下可以较快地得到满意精度的可行解或最优解.

References

1 Youshen Xia, Jun Wang. A general methodology for designing globally convergent optimization neural networks. *IEEE Transactions on Neural Networks*, 1998, **9**(6): 1331~1444

2 Staoshi Matsud. "Optimal" Hopfield network for combinatorial optimization with linear cost function. *IEEE Transactions on Neural Networks*, 1998, **9**(6): 1319~14329

3 Smith K, Palaniswami M, Krishnamoorthy M. Neural techniques for combinatorial optimization with applications. *IEEE Transactions on Neural Networks*, 1998, **9**(6): 1307~1318

4 Teixeira M C M, Zak S H. Analog neural nonderivative optimizers. *IEEE Transactions on Neural Networks*, 1998, **9**(4): 629~638

5 Rubinov A M, Glover B M, Yang X Q. Extended lagrange and penalty functions in continuous optimization. *Optimization*, 1999, **46**(3): 327~351

6 Yang X Q. An exterior point method for computing point that satisfies second-order necessary condition for a C^{1, 1} optimization problem. *Journal of Mathematics Analysis and Application*, 1994, **187**(1): 118~133

7 Rubinov A M, Glover B M, Yang X Q. Decreasing functions with applications to penalization. *SIAM Journal of Optimization*, 1999, **10**(1): 289~313

8 Fiacco A V, McCormick Garth P. Nonlinear Programming. Philadelphia, USA: The SIAM Press, 1990

附　录

定理 1 的证明. 首先将(NP)的约束统一转化为不等式约束:对任何 $x\in X$，令

$$f_i(\boldsymbol{x}) = g_i(\boldsymbol{x}), i = 1,2,\cdots,m; f_{m+j}(\boldsymbol{x}) = h_j(\boldsymbol{x}), j = 1,2,\cdots,q; f_{m+q+j}(\boldsymbol{x}) = -h_j(\boldsymbol{x}), j = 1,2,\cdots,q$$

设 $X_1=\{\boldsymbol{x}\in R^n \mid \boldsymbol{x}\in X,\ f_i(\boldsymbol{x})\leqslant 0,\ i=1,2,\cdots,m+2q\}$, $\bar{I}=\{1,2,\cdots,m+2q\}$. 显然 $X_0=X_1$. 因此(NP)问题可等价写成(NP) $\min\{f(\boldsymbol{x}) \mid \boldsymbol{x}\in X,\ f_i(\boldsymbol{x})\leqslant 0, i=1,2,\cdots, m+2q\}$. 考虑下面对应的罚函数

$$F(\boldsymbol{x},\bar{\rho}) = (f^2(\boldsymbol{x}) + \sum_{i\in\bar{I}}\rho_i p(f_i^+(\boldsymbol{x})))^{1/2}$$

其中 $\bar{\rho}=(\rho_1,\rho_2,\cdots,\rho_{m+2q})$, $\rho_i\geqslant 0$, $i\in\bar{I}$. 由文[5]的 Theorem3.1，有 $\max\limits_{\bar{\rho}\in R_+^{m+2q}}\min\limits_{\boldsymbol{x}\in X}F(\boldsymbol{x},\bar{\rho})=\min\limits_{\boldsymbol{x}\in X_0}f(\boldsymbol{x})$. 对所有的 $i=1,2,\cdots, m+2q$,不妨设 $\rho_i=\rho$,则定理结论成立. 证毕.

定理 2 的证明. 结论 1)是显然的. 仅需证明结论 2). 因 f, $g_i(\forall i\in I)$和 $h_j(\forall j\in J)$是可微凸函数，则有

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*),\qquad \forall \boldsymbol{x}\in X$$

$$g_i(\boldsymbol{x}) \geqslant g_i(\boldsymbol{x}^*) + \nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*),\qquad \forall \boldsymbol{x}\in X,\ i=0,1,2,\cdots,m$$

$$h_j(\boldsymbol{x}) \geqslant h_j(\boldsymbol{x}^*) + \nabla h_j(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*),\qquad \forall \boldsymbol{x}\in X,\ j=0,1,2,\cdots,q$$

又因 $\boldsymbol{x}^*$ 是(NP)的最优解，$\boldsymbol{y}^*$ 和 $\boldsymbol{z}^*$ 是强 Lagrange 乘子向量. 利用式(1),(2)和上面不等式，获得

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*) = f(\boldsymbol{x}^*) - \sum_{i\in I}y_i^*\nabla g_i(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*) - \sum_{j\in J}z_j^*\nabla h_j(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*) \geqslant$$

$$f(\boldsymbol{x}^*) - \sum_{i\in I}y_i^* g_i(\boldsymbol{x}) - \sum_{j\in J}z_j^* h_j(\boldsymbol{x})$$

设 $I^+(\boldsymbol{x}_\rho^*)=\{i\in I \mid g_i(\boldsymbol{x}_\rho^*)>0\}$和 $J^+(\boldsymbol{x}_\rho^*)=\{j\in J \mid h_j(\boldsymbol{x}_\rho^*)>0\}$，有

$$f(\boldsymbol{x}_\rho^*) \geqslant f(\boldsymbol{x}^*) - \sum_{i\in I^+(\boldsymbol{x}_\rho^*)}y_i^* g_i(\boldsymbol{x}_\rho^*) - \sum_{j\in J^+(\boldsymbol{x}_\rho^*)}z_j^* h_j(\boldsymbol{x}_\rho^*)$$

取 $\lambda=\max\{y_i^*,\ z_j^* \mid i=1,\cdots,m,\ j=1,\cdots,q\}$，则有 $\lambda\geqslant 0$ 和

$$-y_i^* g_i(\boldsymbol{x}_\rho^*) \geqslant -\lambda g_i(\boldsymbol{x}_\rho^*),\ \forall i\in I^+(\boldsymbol{x}_\rho^*),\qquad -z_j^* h_j(\boldsymbol{x}_\rho^*) \geqslant -\lambda h_j(\boldsymbol{x}_\rho^*),\ \forall j\in J^+(\boldsymbol{x}_\rho^*)$$

从而有

$$f(\boldsymbol{x}_\rho^*) \geqslant f(\boldsymbol{x}^*) - \lambda\sum_{i\in I^+(\boldsymbol{x}_\rho^*)}g_i(\boldsymbol{x}_\rho^*) - \lambda\sum_{j\in J^+(\boldsymbol{x}_\rho^*)}h_j(\boldsymbol{x}_\rho^*)$$

由 $b(x)$的定义,有 $f(\boldsymbol{x}_\rho^*)\geqslant f(\boldsymbol{x}^*)-\lambda(m+q)b(\boldsymbol{x}_\rho^*)$. 如果 $f(\boldsymbol{x}^*)<\lambda(m+q)b(\boldsymbol{x}_\rho^*)$，显然有

$$F^2(\boldsymbol{x}_\rho^*,\rho) \geqslant (f(\boldsymbol{x}^*))^2 - (\lambda(m+q)b(\boldsymbol{x}_\rho^*))^2 + \rho\sum_{i\in I^+(\boldsymbol{x}_\rho^*)}g_i^+(\boldsymbol{x}_\rho^*)^2 + \rho\sum_{j\in J}h_j(\boldsymbol{x}_\rho^*)^2 \geqslant$$

$$f^2(\boldsymbol{x}^*) - (\lambda(m+q)b(\boldsymbol{x}_\rho^*))^2 + \rho b(\boldsymbol{x}_\rho^*)^2 \tag{A1}$$

如果 $f(\boldsymbol{x}^*)\geqslant\lambda(m+q)b(\boldsymbol{x}_\rho^*)$，可得到

$$F^2(\boldsymbol{x}_\rho^*,\rho) \geqslant (f(\boldsymbol{x}^*) - \lambda(m+q)b(\boldsymbol{x}_\rho^*))^2 + \rho\sum_{i\in I^+(\boldsymbol{x}_\rho^*)}g_i^+(\boldsymbol{x}_\rho^*)^2 + \rho\sum_{j\in J}h_j(\boldsymbol{x}_\rho^*)^2 \geqslant$$

$$f^2(\boldsymbol{x}^*) - 2\lambda(m+q)b(\boldsymbol{x}_\rho^*)f(\boldsymbol{x}^*) + (\lambda(m+q)b(x_\rho^*))^2 + \rho b(x_\rho^*)^2 \geqslant$$

$$f^2(\boldsymbol{x}^*) - 2\lambda(m+q)b(\boldsymbol{x}_\rho^*)f(\boldsymbol{x}^*) + \rho b(\boldsymbol{x}_\rho^*)^2 \tag{A2}$$

因此,取 $\rho\geqslant\max\left\{\lambda^2(m+q)^2, \dfrac{2(m+q)\lambda f(\boldsymbol{x}^*)}{b(\boldsymbol{x}_\rho^*)}\right\}$，由式(A1)和(A2),有 $F(\boldsymbol{x}_\rho^*,\rho)\geqslant f(\boldsymbol{x}^*)$,它的反向等式由已知成立,故结论 2)成立. 证毕.

定理 5 的证明. 根据已知,存在一个大于 0 的数 L,使得 $L>E(\boldsymbol{x}^k)$, $k=0,1,2,\cdots$. 假设$\{\boldsymbol{x}^k\}$无界. 那么，$\boldsymbol{x}^k\to\infty$当 $k\to+\infty$，由上面得到 $L>(f(\boldsymbol{x}^k))^2\to+\infty, k=0,1,2,\cdots$，导致与已知矛盾. 下面证明$\{\boldsymbol{x}^k\}$的任意极限点属于 X_0. 不失普遍性,不妨假设 $\boldsymbol{x}^k\to\bar{x}$. 如果 $\bar{x}\notin X_0$，那么存在某个 i 或 j 使得 $g_i^+(\bar{x}))^2>0$ 或

$h_j^2(\bar{x})>0$. 因为 $f, g_i(i\in I)$ 和 $h_j(x)(j\in J)$ 是连续的，$E(x)$ 也是连续的. 我们有 $L>E(x^k)=(f(x^k))^2+\rho_k\sum_{i\in I}g_i^+(x^k)^2+\rho_k\sum_{j\in J}h_j(x^k)^2$. 令 $k\to+\infty$，则有 $\rho_k\to+\infty$ 导出 $L>E(x^k)\to+\infty$，得到一个矛盾.

式(6)的证明. 由已知 $-\nabla E(x^k)=0$，$k=0,1,2,\cdots$. 得到

$$2f(x^k)\nabla f(x^k)+\rho_k\sum_{i\in I}2g_i^+(x^k)\nabla g_i(x^k)+\rho_k\sum_{j\in J}2h_j(x^k)\nabla h_j(x^k)=0 \tag{A3}$$

令 $x^k\to\bar{x}$，$\beta_k=2\mid f(x^k)\mid+\rho_k\sum_{i\in I}2g_i^+(x^k)+\rho_k\sum_{j\in J}2\mid h_j(x^k)\mid+1>0$. 因为 $f, g_i(i\in I)$ 和 $h_j(x)(j\in J)$ 是连续可导的，则序列 $\{2f(x^k)\beta_k^{-1}\}$，$\{\rho_k 2g_i^+(x^k)\beta_k^{-1}\}$ 和 $\{\rho_k 2h_j(x^k)\beta_k^{-1}\}$ 是有界序列，因此它们存在收敛的子序列，不妨设它们收敛，故令

$$2(f(x^k))\beta_k^{-1}\to\lambda;\ 2\rho_k g_i^+(x^k)\beta_k^{-1}\to\mu_i\geqslant 0,\ i\in I;\ 2\rho_k h_j(x^k)\beta_k^{-1}\to\alpha_j,\ j\in J$$

式(A3)两边同时乘以 β_k^{-1}，再令 $k\to+\infty$，即得到式(6). 证毕.

例. 我们使用 Visual C++6.0 在 PentiumIII(CPU，128M 内存)机上对下面模型进行了数值实验

$$\begin{aligned}\max\quad & f(x)=x_1+\cdots+x_n,\\ \text{s.t.}\quad & a_{i1}x_1+\cdots+a_{in}x_n\leqslant 10,\ i=1,\cdots,m,\\ & x_1,\cdots,x_n=0\ \text{或}\ 1,\end{aligned}$$

其中 $a_{ij}(i=1,\cdots,m,j=1,\cdots,n)$ 是 0 到 2 之间的随机数. 通过在 200×200 以内的许多实验表明，多数最优目标值是 7 和 8，在较短时间内可得到最优解. 表 1 给出了 6 个模型的计算结果.

表 1 6 个模型的计算结果

Table 1 The calculated results of six models

模型大小	步数k	目标值	小步率 1（步/秒）	小步率 2（步/秒）	模型大小	步数k	目标值	小步率 1（步/秒）	小步率 2（步/秒）
30×30	2	9.000000	92	92	100×100	4	7.000000	39	39
60×60	17	8.000000	67	67	200×200	12	8.000000	7.2	7.5
90×90	9	8.000001	41	41	300×300	2	7.000001	4.0	4.3

说明: 1) 表中目标值是对应步数 k 得到的目标值, 由于模型规模和系数值不同所以迭代次数不同; 2) 表中小步率 1 是 *Step*2 算法中的每一步迭代的速度, 即每秒计算一步目标值的次数.

孟志青 西安电子科技大学博士研究生, 现任湘潭大学信息工程学院计算机系教授. 主要研究方向为最优决策与最优控制.

(**MENG Zhi-Qing** He is a professor at Computer Department of Xiangtan University. Ph. D. candidate at Xidian University. His research interests include optimal decision making and optimal control.)

胡奇英 教授, 博士生导师, 现为西安电子科技大学管理学院院长. 主要研究方向为最优决策与最优控制.

(**HU Qi-Ying** He is a professor at school of Economy & Management of Xidian University. His research interests include optimal decision making and optimal control.)

杨晓琪 博士, 现为香港理工大学应用数学系副教授. 主要研究方向为非线性优化计算与最优控制.

(**YANG Xiao-Qi** Now, he is an associate professor at Department of Applied Mathematics of Hong Kong Polytechnic University. His research interests include nonlinear optimization computing and optimal control.)