第25卷 第5期
1999年9月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 25, No. 5
Sep., 1999

研究简报

一种新的线性规划问题的神经网络解法

田大钢 费 奇
(华中理工大学系统工程研究所 武汉 430074)

关键词 线性规划,神经网络,熵障碍对偶法.

NEW ALGORITHM FOR LINEAR PROGRAMMING WITH NEURAL NETWORKS

TIAN Dagang FEI Qi
(*Institute of System Engineering, Huazhong University of Science and Technology, Wuhan* 430074)

Key words Linear programming, neural network, entropic perturbation.

1 引言

单纯形法是解线性规划问题的最常用方法,可它不是一种多项式算法[1]. 椭圆算法[2]的提出,使人们认识到线性规划问题存在多项式解法. 但椭圆算法本身在实际中的应用却并不成功. 内点法[3-5]是新的一类多项式算法,尽管它在求解大规模线性规划问题方面显示了相当的潜力,其算法的精度和软件的开发都有待完善和发展.

神经网络方法展示了一种新的计算思想. 由于固有的并行性和学习、联想能力,其应用和发展前景未可估量. 对线性规划问题而言,Hopfield 和 Tank 提出的 TH 算法[6]是这种方法的代表,然而,TH 算法所提供的能量函数以及沿着同样方法构造的能量函数只保证网络收敛到局部极小(不是全局最小)点,有时甚至不收敛,或者解不可行[7]. 这些缺陷的不良影响当然不仅仅限制在 TH 方法本身.

本文详细讨论了熵障碍对偶方法[8]的目标函数. 指出利用此函数作为能量函数建立的神经网络具有良好的收敛性质,是一种求解线性规划问题的令人放心的算法.

2 线性规划问题的熵障碍对偶方法及其收敛性

考虑标准形式的线性规划问题 P:min $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{X}$, $\boldsymbol{X}$ 满足约束条件 $A\boldsymbol{X}=\boldsymbol{b}, \boldsymbol{X}\geqslant 0$,以及 P 的对偶 Q:max $b^{\mathrm{T}}\boldsymbol{W}$, $\boldsymbol{W}$ 满足条件 $A^{\mathrm{T}}\boldsymbol{W}\leqslant\boldsymbol{C}$,这里 A 是 $m\times n$ 约束矩阵. 不直接进行 P,Q 的

收稿日期 1998-02-17 收修改稿日期 1998-10-25

求解，而是讨论 P 的熵扰动形式

$$P_\mu:\min \boldsymbol{C}^{\mathrm{T}}\boldsymbol{X}+\mu\sum_{j=1}^{n}x_j\log x_j \qquad \text{s.t. } A\boldsymbol{X}=\boldsymbol{b},\boldsymbol{X}\geqslant 0.$$

重要的是 P_μ 有无约束对偶[8]：

$$D_\mu:\max \boldsymbol{b}^{\mathrm{T}}\boldsymbol{W}-\mu\sum_{j=1}^{n}\exp\left\{\left[\left(\sum_{i=1}^{m}a_{ij}w_i-c_j\right)/\mu\right]-1\right\},$$

这里 $\mu>0$. D_μ 的完全无约束性以及目标函数的良好性质使得可以用各种无约束方法对其进行求解，这是熵扰动与传统障碍法的主要区别.

在某些“标准”条件下（例如内点条件），P_μ 有唯一最优解 $\boldsymbol{X}^*(\mu)$，满足 $\lim\limits_{\mu\to 0}\boldsymbol{X}^*(\mu)=\boldsymbol{x}^*$，$\boldsymbol{x}^*$ 是 P 的最优解. 同样的 D_μ 有唯一最优解 $\boldsymbol{W}^*(\mu)$，满足 $\lim\limits_{\mu\to 0}\boldsymbol{W}^*(\mu)=\boldsymbol{w}^*$，$\boldsymbol{w}^*$ 是 D 的最优解. 这样，将 LP 问题转化为求无约束问题 D_μ 的最优解.

讨论 D_μ 的求解. 对任意 $\varepsilon>0,\mu>0$，设

$$f(\boldsymbol{w},\mu,\varepsilon)=\mu\sum_{j=1}^{n}\exp\left\{\left[\left(\sum_{j=1}^{n}a_{ij}w_i-c_j\right)/\mu\right]-1\right\}+\varepsilon\|\boldsymbol{w}\|^2-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{w},$$

显然，当 ε 趋于零时，只要 $\boldsymbol{w}(\mu,\varepsilon)$ 有界，$f(\boldsymbol{w},\mu,\varepsilon)$ 的极小点 $\boldsymbol{w}(\mu,\varepsilon)$ 趋于 D_μ 的解 $\boldsymbol{w}(\mu)$.

考虑梯度系统

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=-\nabla f(\boldsymbol{w},\mu,\varepsilon),\\ \boldsymbol{w}(0,\mu,\varepsilon)=\boldsymbol{w}_0(\mu,\varepsilon).\end{cases} \tag{1}$$

由梯度系统的性质[9]知，式(1)若有平衡点，则必定是渐近稳定的. 而本文有如下的定理.

定理1. 对任何 $\mu>0,\varepsilon>0$，式(1)存在唯一的平衡点 $\boldsymbol{w}(\mu,\varepsilon)$，它是全局指数型渐进稳定的. 而当 A 行满秩时，对于 $\varepsilon=0$，式(1)的平衡点 $\boldsymbol{w}(\mu)$（若存在）是唯一的且指数型渐进稳定的.

证. 取 $f(\boldsymbol{w},\mu,\varepsilon)$ 作为式(1)的 Liapunov 函数，应用文献[10]中定理8.7和文献[9]中定理3. 证毕.

从自动计算的角度考虑，下面讨论 ε 随时间 t 变化的情况. 考虑

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=-\nabla f(\boldsymbol{w},\mu,0)-\varepsilon(t)\boldsymbol{w},\\ \boldsymbol{w}(0,\mu)=\boldsymbol{w}_0(\mu).\end{cases} \tag{2}$$

定理2. 设 $\boldsymbol{w}(\mu)$ 是 $\varepsilon=0$ 时，(1)的平衡点. 若 $\varepsilon(t)$ 满足 $\lim\limits_{t\to\infty}\varepsilon(t)=0$，$\int_1^{\infty}\varepsilon(\tau)\mathrm{d}\tau=\infty$，则式(2)的任意解都趋于 $\boldsymbol{w}(\mu)$.

证. 应用文献[10]中定理10.1. 证毕.

3 熵障碍对偶法的神经网络求解

将式(1)写成标量形式

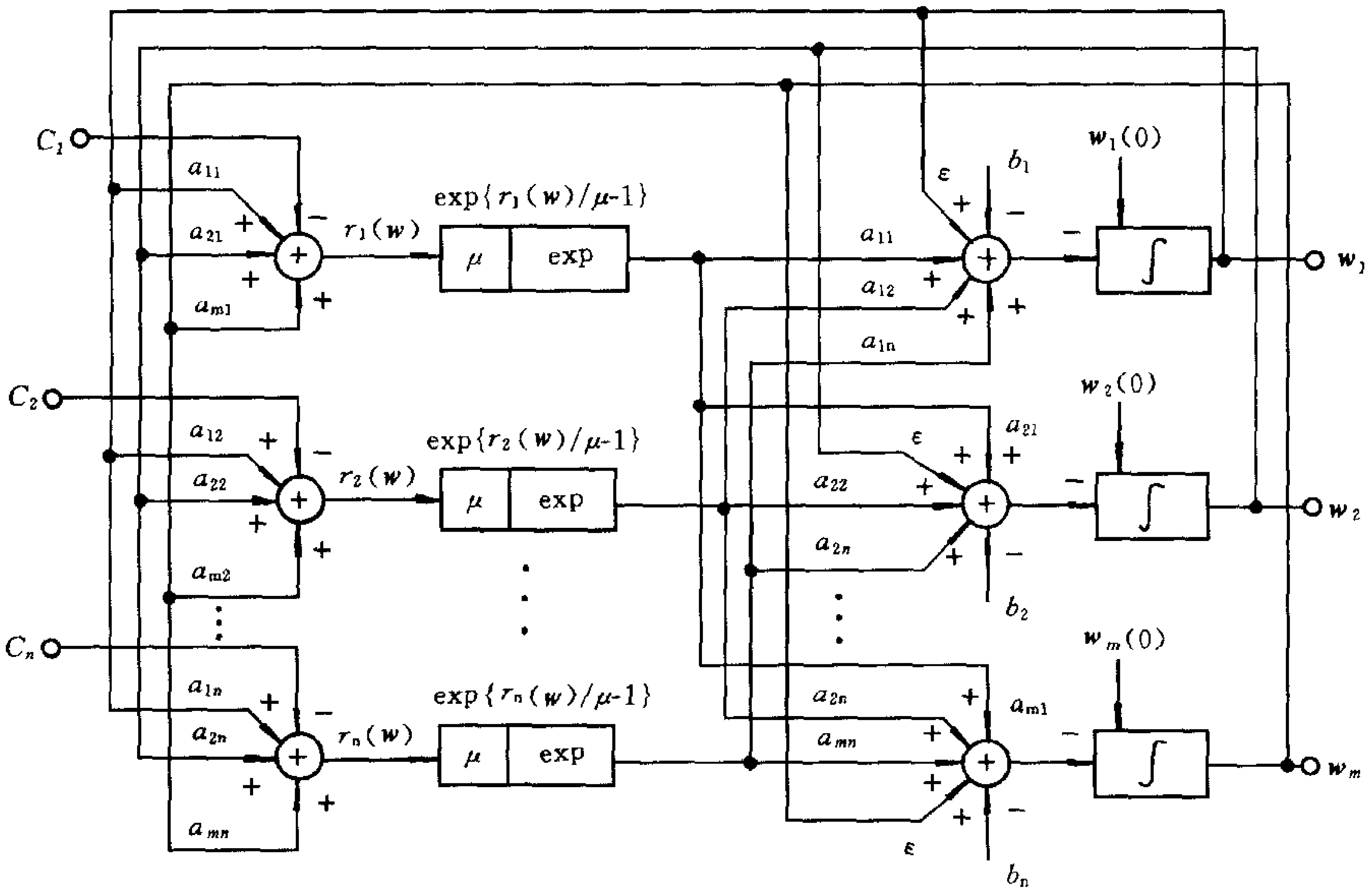


图1　系统(1)的神经网络实现

$$\frac{\mathrm{d}w_i}{\mathrm{d}t} = b_i - 2\varepsilon w_i - \sum_{j=1}^{n} a_{ij}\exp\left\{\left[\left(\sum_{k=1}^{m} a_{kj} w_k - c_j\right)/\mu\right] - 1\right\}, \quad i = 1,2,\cdots,m \tag{3}$$

图1为式(3)的神经网络求解框图，包括加法器、积分器和函数产生器．其中 $r_i(w) = \sum_{k=1}^{m} a_{ki} w_k - c_i, i=1,2,\cdots,n$. 与用一般障碍法实现的神经网络比较，式(3)的性态更好.

4　仿真和结论

例1. $\min\ z = 3x_1 + 2x_2 + x_3 + 4x_4$　,

s.t　$2x_1 + 4x_2 + 5x_3 + x_4 - x_5 = 230$,

$3x_1 - x_2 + 7x_3 - 2x_4 - x_6 = 46$,

$5x_2 + 2x_2 + x_3 + 6x_4 - x_7 = 345$,

$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7 \geqslant 0$.

例2. *病态问题*

$\min\ z = x_1 + x_2$,

s.t $2x_1 - 2x_2 + x_3 + x_4 + 0.5x_5 = 2$,

$x_1 + x_2 + x_3 \quad + 0.5x_5 = 1, 3x_1 - x_2 + 2x_3 + x_4 + x_5 = 3, x_1, x_2, x_3, x_4, x_5 \geqslant 0$,

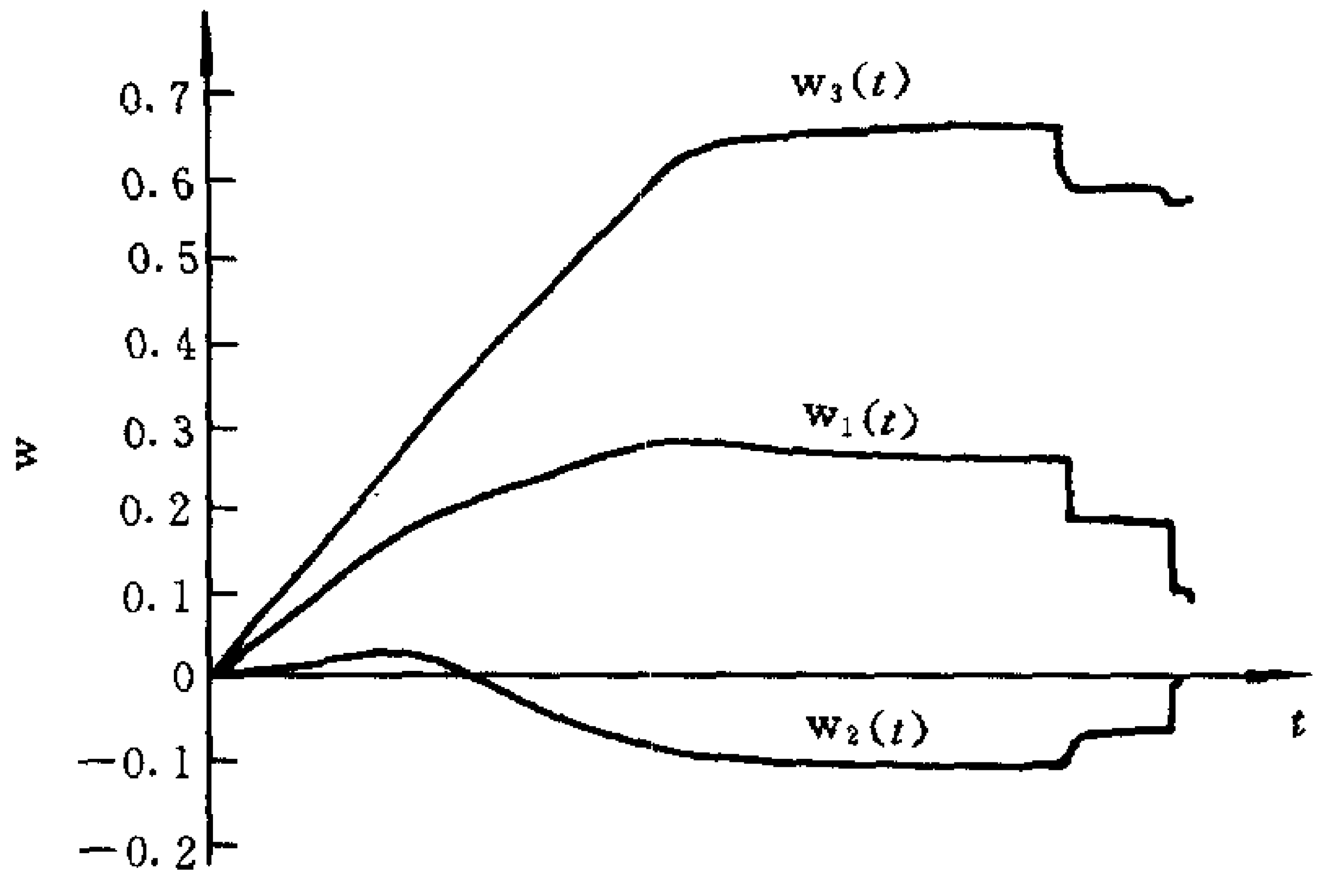
计算机仿真结果如图2. μ,ε 从0.000 1变到1.0e—10. 例1的最优值为251.0002，其中：

$$w_1 = 0.086\ 957\ 51, w_2 = -\ 0.000\ 000\ 67,\ w_3 = 0.565\ 217\ 50.$$

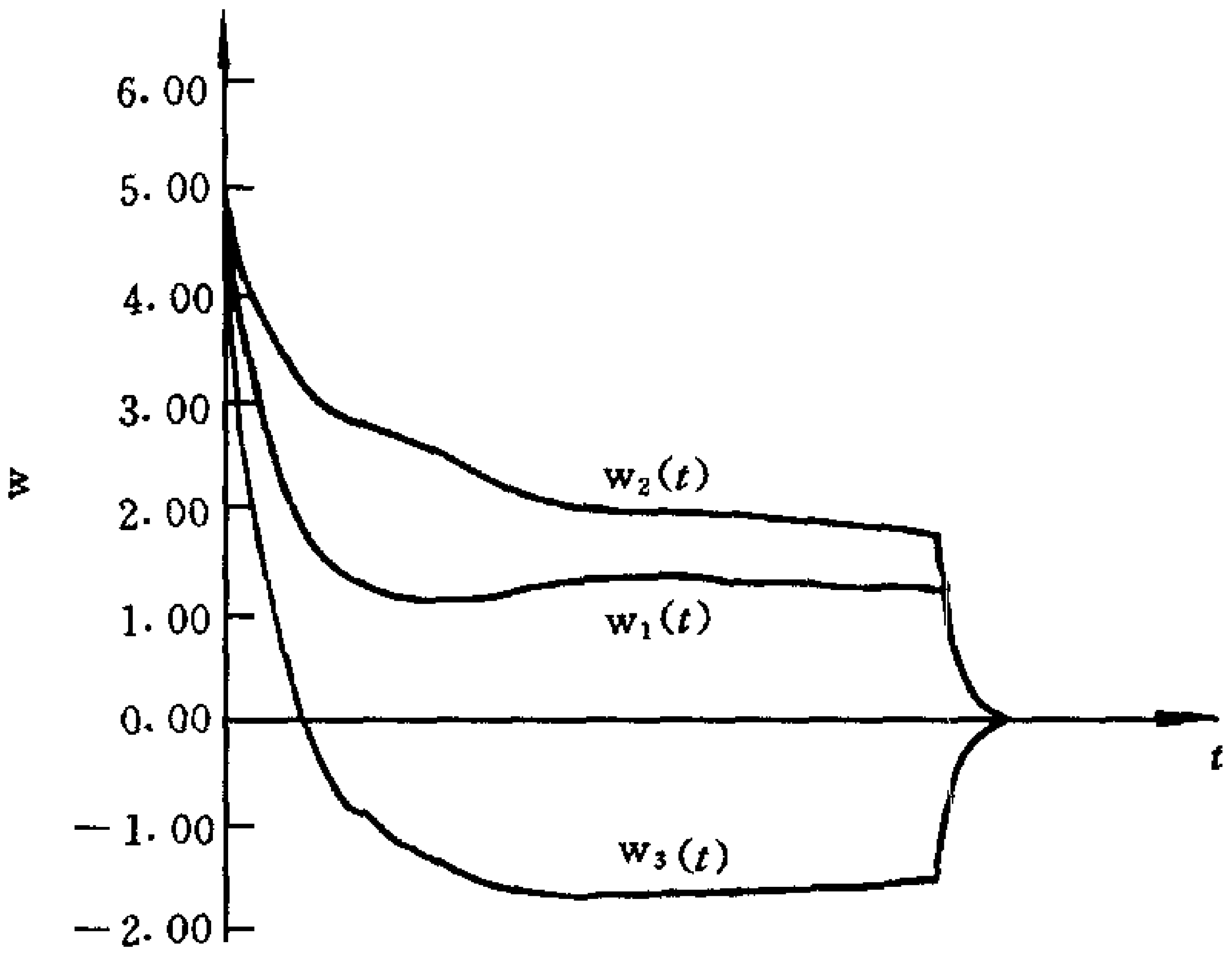
例2的最优值和解向量均为零. 计算采用龙格-库塔法，且进行了步长减半矫正. 由于 exp 函数的存在，计算精度要求较高(否则容易产生溢出)，本文将矫正精度定在1.0e—15. 基步长定为0.000 1，最后的时间长度分别为0.008 33，0.002 37.

对任何 $\varepsilon>0, \mu>0$，式(1)的全局收敛性使得最优解的计算不必从内点进行；式(1)的良好性质，也使得我们可以采取其它优化方法，例如牛顿法，求解线性规划问题，可以得到

二阶收敛. 还可以通过 $\varepsilon>0$时，式(1)的计算结果来判断问题 D_μ 解的存在性；在 Hopfield 型网络的计算中，也不必考虑初态的选取和可行性验证，且原问题和对偶问题的解可以同时得到[8]. 定理2和例2表明病态并不影响计算. 进一步研究表明，扰动不必取 exp 函数，这将极大地提高计算效率，关于这一点将另文发表.



(a)例1



(b)例2

图2 例1,例2的仿真曲线

参 考 文 献

1 Klee V, Minty G J. How good is the simplex algorithm? In Inequalities III(Edited by Shisha O), New York: Academic Press, 1972, 159—175

2 Bland R G, Goldfarb D, Todd M J. The ellipsoid method: a survey. *Oper. Res*, 1981, **29**:1039—1091

3 Karmarkar N K. A new polynomial-time algorithm for linear programming. *Combinatorica*, 1984, **4**(4):375—395

4 Hooker J N. Karmarkar's linear programming algorithm. *Interfaces*, 1986, **14**(1):75—90

5 Ghellinck G de, Vial J P. A polynomial Newton method for linear programming. *Algorithmica*, 1986, **1**(3): 425—453

6 Tank D W, Hopfield J J. Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Trans. Circuits Syst.*, 1986, **33**(4):533—541

7 Jun W, Vira C. Recurrent neural networks for linear programming: analysis and design principles. *Computers Ops. Res.*, 1992, **19**(3—4): 297—311

8 Fang S-C, Tsao H-S T. Linear programming with entropic perturbation. ZOR., 1993, **37**(2):171—186

9 张锦炎. 常微分方程几何理论与分支问题. 北京:北京大学出版社, 1981, 179—183

10 Yoshizawa T 著, 郑祖庥等译. 稳定性理论与周期解和概周期解的存在性. 南宁:广西人民出版社, 1985, 36—136

田大钢 40岁，博士生. 研究方向：神经网络理论及其应用、决策和决策支持系统、反应扩散方程、分形理论在决策中的应用等.

费 奇 59岁，博士生导师. 研究方向：决策支持系统理论、建模支持、知识处理、不确定性分析、分布式环境下的群决策、分形理论在决策中的应用等.