



# 单层神经网络的快速学习算法研究

冯昭志      黄载禄      杨叔子

(华中理工大学电子与信息工程系 武汉 430074)

**摘 要** 该文提出一种适用于单层神经网络(SNN)训练的新颖的广义误差函数,给出了 SNN 新的快速学习算法(FLA). 进一步提出了一种广义系统辨识模型,对 FLA 的收敛性进行了理论分析. 实验表明:文中给出的新 FLA 比 Karayiannis 的 LFA 具有更快的收敛速度.

**关键词** 单层神经网络,快速 BP 算法,统计分析.

最近, Karayiannis<sup>[1]</sup>提出了单层神经网络(SNN)的一系列快速学习算法(FLA),较好地改善了 SNN 的收敛性能. 然而, Karayiannis 没有讨论神经元状态为连续的情况,也未讨论 FLA 的收敛性. 虽然 Bershad<sup>[2]</sup>等人在 SNN 的 BP 算法分析方面取得了突破,可他们提出的系统辨识模型不适用于对 FLA 的理论分析. 针对这些问题,本文将试图对 SNN 的 FLA 及其收敛性分析作一新的探讨.

## 1 SNN 的快速学习算法

单层神经网络(SNN)的结构如图 1 所示. 在图中,  $N_{in}$ 、 $N_{out}$  分别是输入层和输出层的神经元数;  $w_{ij}$  是输入层第  $j$  个神经元到输出层第  $i$  个神经元之间的连接权. 设  $\mathbf{x}_k^* = [x_{0,k} \ x_{1,k} \ \cdots \ x_{N_{in},k}]$ ,  $\mathbf{y}_k = [y_{1,k} \ y_{2,k} \ \cdots \ y_{N_{out},k}]$  分别为 SNN 的输入向量和输出向量,  $k = 1, 2, \dots, N_p$ ,  $N_p$  为输入模式数;  $x_{0,k} = 1.0$ , 则 SNN 的回忆过程的系统动力学方程为

$$y_{i,k} = f(\bar{y}_{i,k}) = f(\mathbf{x}_k^* \mathbf{w}_i) = f\left(\sum_{j=0}^{N_{in}} w_{ij} x_{j,k}\right) \quad (1)$$

式中  $\mathbf{w}_i^* = [w_{i0} \ w_{i1} \ \cdots \ w_{iN_{in}}]$ ;  $f(x) = \text{sigmoid}(x)$ ; \* 为矩阵转置运算.

给定网络输入模式  $\mathbf{x}_k^*$  和训练目标模式  $\mathbf{t}_k [t_{1,k} \ t_{2,k} \ \cdots \ t_{N_{out},k}]$ , SNN 的学习过程就是迭代地修改网络权值,从而使广义误差函数(能量函数)的值为最小. 为了加速 SNN 的收敛进程,我们提出了新的广义误差函数

$$G_k(\lambda) = \lambda \sum_{i=1}^{N_{out}} \varphi_2(e_{i,k}) + (1 - \lambda) \sum_{i=1}^{N_{out}} \varphi_1(e_{i,k}), \quad (2)$$

其中  $\varphi_2(x) = \frac{1}{2}x^2$ ,  $\varphi_1(x) = \frac{1}{\lambda - 1} [t_{i,k} \ln(t_{i,k} - x) + (t_{i,k} - \lambda - \lambda t_{i,k})(t_{i,k} - x)]$ ,  $\lambda = \lambda(E) = \exp(-\mu/E^2)$ ,  $E = \frac{1}{2} \sum_{i=1}^{N_{out}} (t_{i,k} - y_{i,k})^2$ ,  $\mu$  是一给定常数,  $\mu \in (0, 1)$ .

用梯度下降法可推导出 SNN 新的快速学习算法的递归计算公式. 权值矩阵  $w_{ij}^{(k)}$  和反向误差校正信号  $\delta_{i,k}(\lambda)$  的数学表达式为

$$w_{ij}^{(k)} \leftarrow w_{ij}^{(k)} - \eta \frac{\partial G_k(\lambda)}{\partial w_{ij}^{(k)}} = w_{ij}^{(k)} + \eta \delta_{i,k}(\lambda) x_{j,k},$$

$$1 \leq k \leq N_p, 1 \leq i \leq N_{out}, \quad 1 \leq j \leq N_{in}, \quad (3)$$

$$\delta_{i,k}(\lambda) = \frac{\partial G_k(\lambda)}{\partial y_{i,k}} = (1 - y_{i,k}^2)(t_{i,k} - \lambda y_{i,k}). \quad (4)$$

显然在公式(4)中,  $(t_{i,k} - \lambda y_{i,k})$  为输出层第  $i$  个神经元的输出误差;  $(1 - y_{i,k}^2)$  则为传统的导函数因子项. 因此, 可以认为新 FLA 是用  $(1 - y_{i,k}^2)$  替代了 Karayiannis 的 FLA 中的导函数因子项  $(1 - y_{i,k})y_{i,k}$  (当  $f(x) = \text{sigmoid}(x)$ ). 进一步考虑到不等式  $(1 - y_{i,k})y_{i,k} \leq 1 - y_{i,k}^2 \leq \exp(-y_{i,k}^2) \leq \cos(y_{i,k})$  成立, 我们可得到 SNN 的一个导函数集合  $g(y_{i,k}) = \{(1 - y_{i,k})y_{i,k}, 1 - y_{i,k}^2, \exp(-y_{i,k}^2), \cos(y_{i,k})\}$ . SNN 新的快速学习算法的详细描述如下:

#### Algorithm Fast BP algorithm

Initialize Randomly  $W$

Select  $E_0, \lambda, \mu, \eta, \alpha; \lambda = 1; \text{num} = 0$

I:  $\text{num} = \text{num} + 1; k = 0; E = 0$

II:  $k = k + 1$

$$y_{i,k} = f\left(\sum_{j=0}^{N_{in}} w_{ij} x_{j,k}\right) \quad 1 \leq i \leq N_{out};$$

$$\delta_{i,k}(\lambda) = g(y_{i,k})(t_{i,k} - \lambda y_{i,k})$$

$$w_{ij} \leftarrow w_{ij} + \eta \delta_{i,k}(\lambda) x_{j,k}$$

$$y_{i,k} = f\left(\sum_{j=0}^{N_{in}} w_{ij} x_{j,k}\right)$$

$$E \leftarrow E + \frac{1}{2} \sum_{i=1}^{N_{out}} (t_{i,k} - y_{i,k})^2$$

if  $k < m$ ; then: go to II

$\lambda = \exp(-\mu/E^2)$ ; if  $E > E_0$ ; then: go to I End

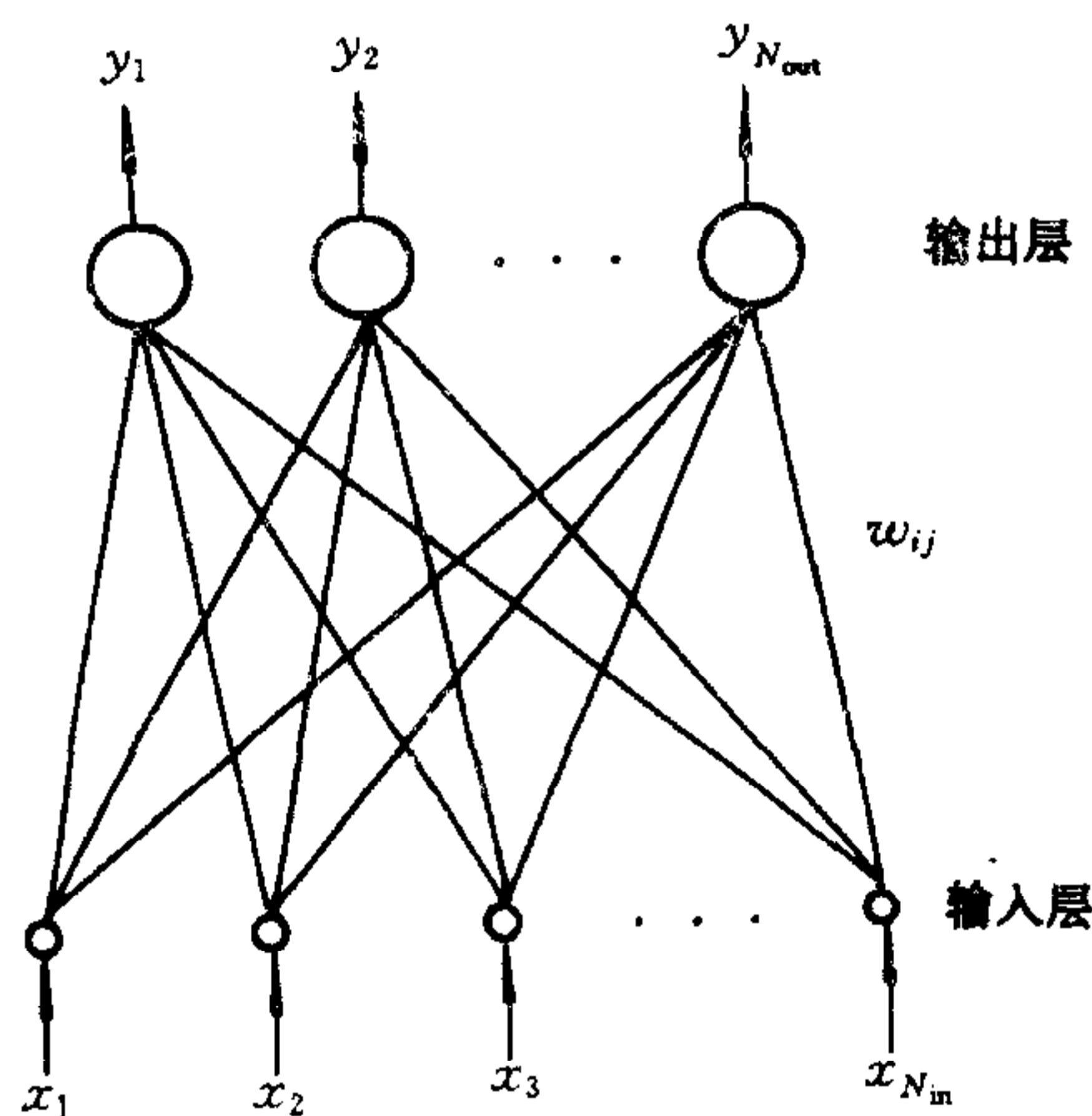


图1 单层神经网络结构

## 2 算法收敛性分析

为了对新 FLA 进行分析, 我们推广了 Bershad 等人的系统辨识模型, 提出了广义系统辨识模型(见图 2). 在图 2 中,  $X(n)$  是高斯型输入向量;  $F, W(n)$  分别是系统辨识模型权值和 SNN 的连接权值;  $e_1(n), e(n)$  则分别是广义系统辨识模型的内部误差和系统误

差;函数  $f(x)$ 、 $\lambda(x)$  为非线性函数. 为了讨论方便,取  $f(x) = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \int_0^x e^{-z^2/2\sigma^2} dz$  (有关符号的定义均与文献[2]的定义相同,以后不再说明).

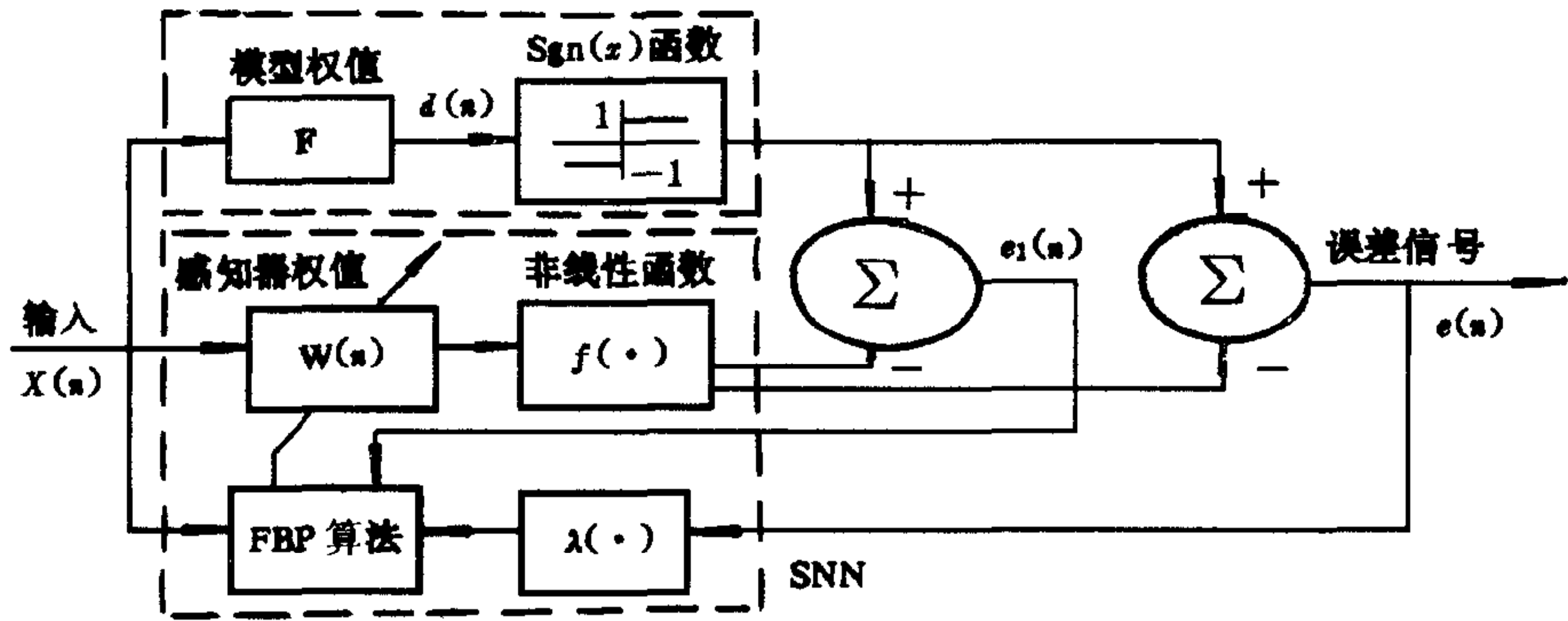


图2 具有期望响应模型的单层神经网络  
(广义系统辨识模型)

SNN 的新 FLA 的迭代计算公式的向量方程为

$$\begin{aligned} W(n+1) &= W(n) + \mu e_1(n) f[W(n)^T(n)X(n)] \\ &\quad [\lambda \phi_2'(e_1(n)) + (1-\lambda)\phi_1'(e_1(n))]X(n) \\ e_1(n) &= \text{sgn}[F^T X(n)] - f[W^T(n)X(n)] \\ e(n) &= \text{sgn}[F^T X(n)] - f[W^T(n+1)X(n)] \end{aligned} \quad (5)$$

类似于文献[2],可推导出权值矩阵  $W(n+1)$  的条件数学期望

$$\begin{aligned} E[W(n+1)|W(n)] &= W(n) + \mu \frac{2}{\pi} \left[ \frac{\sigma_x}{[\sigma^2 + \sigma_x^2 W^T(n)W(n)(1-\lambda^2)]^{1/2}} \right. \\ &\quad \cdot \frac{1}{\sqrt{F^T F}} \left\{ F - \frac{\sigma_x^2 F^T W(n)W(n)}{\sigma_x^2 W^T(n)W(n) + \sigma^2} \right\} \\ &\quad \left. - \frac{\sigma \sigma_x^2 W(n)}{[\sigma_x^2 W^T(n)W(n) + \sigma^2][2\sigma_x^2 W^T(n)W(n) + \sigma^2]^{1/2}} \right] \end{aligned}$$

设  $W(n)$  的平均权值为  $\bar{W}(n) = E[W(n+1)|W(n)]$ , 则

$$\begin{aligned} \bar{W}(n+1) &= \bar{W}(n) + \mu \frac{2}{\pi} \left[ \frac{\sigma_x}{[\sigma^2 + \sigma_x^2 \bar{W}^T(n)\bar{W}(n)(1-\lambda^2)]^{1/2}} \right. \\ &\quad \cdot \frac{1}{\sqrt{F^T F}} \left\{ F - \frac{\sigma \sigma_x^2 F^T \bar{W}(n)\bar{W}(n)}{\sigma_x^2 \bar{W}^T(n)\bar{W}(n) + \sigma^2} \right\} \\ &\quad \left. - \frac{\sigma \sigma_x^2 \bar{W}(n)}{[\sigma_x^2 \bar{W}^T(n)\bar{W}(n) + \sigma^2][2\sigma_x^2 \bar{W}^T(n)\bar{W}(n) + \sigma^2]^{1/2}} \right] \end{aligned} \quad (6)$$

其中  $\lambda = \lambda(E) = \exp(-\mu/E^2)$ . 该式较好地描述了 SNN 的连接权的非线性特征. 当  $\bar{W}(0) = \mathbf{0}$  (零向量) 时, 有

$$\bar{W}(1) = \mu \frac{2}{\pi} \left[ \frac{\sigma_x}{\sigma} \frac{F}{\sqrt{F^T F}} \right]$$

即在一次迭代计算之后,  $\bar{W}(n)$  是沿着逼近  $F$  的方向变化, 从而网络的训练是成功的.

在公式(6)中,当  $n \rightarrow N$ ,  $\lambda$  是以指数曲线上升,并且逼近 1. 然而在传统 EBP 算法中<sup>[2]</sup>,  $\lambda = \frac{F^T W(n)}{\sqrt{F^T F W^T(n) W(n)}}$  是线性地逼近 1. 因此,我们认为这就是新 FLA 比传统 EBP 算法收敛速度快的机理.

### 3 实验

利用本文提出的快速学习算法,我们成功地完成了对某油井井壁介质结构的模式分类实验. 实验中采用的神经网络参数为  $N_{in}=120, N_{out}=4, \eta=0.3, \mu=0.1$ . 网络的初始化权值为  $(-0.5, 0.5)$  之间的随机数,训练误差精度  $E_0=0.01$ . 取井壁介质的超声回波信号的 120 个采样数据为神经网络的输入向量,将第一次对油井井壁介质的采样数据模式作为网络的训练模式,而将第二次获取的介质结构数据作为待识别的模式. 对该油井的井壁介质结构分类效果能达到 100% 的正确率. 对应于  $g(y)$  ( $g(y) = \{(1-y)y, 1-y^2, e^{-y^2}, \cos(y)\}$ ) 的网络自适应学习过程如图 3 所示. 显然,本文提出的新的快速学习算法具有比 Karayiannis 快速 BP 算法更快的收敛速度,且所需的 CPU 计算时间也较少. 该实验说明:文中提出的新学习算法是有效的和可行的.

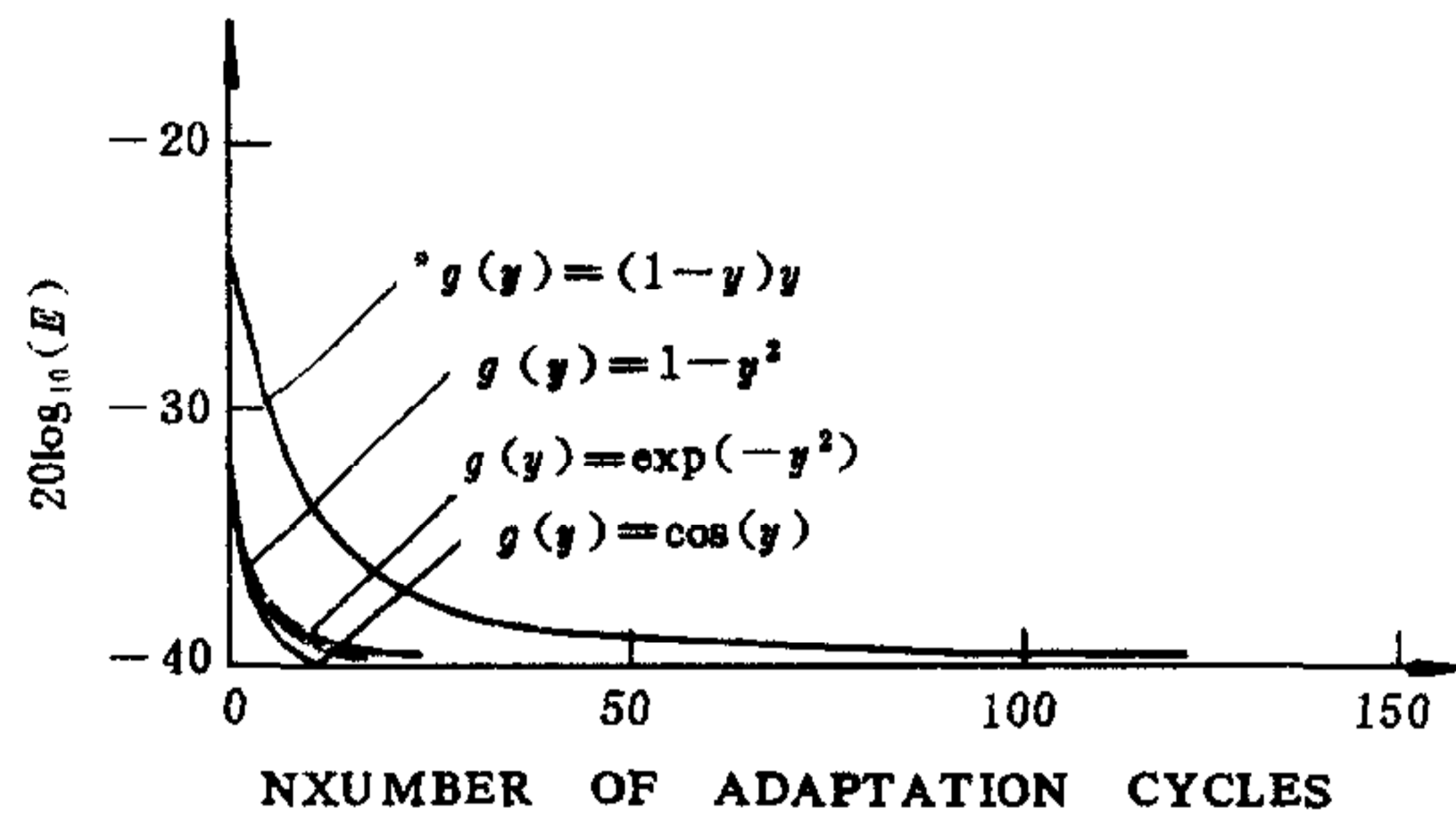


图 3 石油测井应用中的网络自适应学习过程

\* Karayiannis FBP

### 参 考 文 献

- [1] Karayiannis N B, Venetsanopoulos A N. Fast learning algorithm for neural networks. *IEEE Trans. Cas- I*, 1992, **39**(7): 453—474.
- [2] Bershad N J, Shynk J J, Feintuch P L. Statistical analysis of the single-layer backpropagation algorithm: Part I - mean weight behavior. *IEEE Trans. on Signal Processing*, 1993, **41** (2): 573—582.

## STUDY ON THE FAST LEARNING ALGORITHM OF SINGLE-LAYER NEURAL NETWORKS

FENG ZHAOZHI    HUANG ZAILU    YANG SHUZI

*(Dept. of Electronic & Information Eng., Huazhong University of science and Technology, Wuhan 430074)*

**Abstract** This paper proposes a new generalized criterion for the training of single-layer neural networks, which leads to a novel fast learning algorithm for single-layer neural network. In order to analyse the convergent properties of the fast algorithm we developed, a new generalized system identification model is also presented. Experiment results show that the fast algorithm proposed in this paper performs the training of neural network faster than the corresponding learning algorithm given by Karayiannis.

**Key words** Single-layer NN, fast BP algorithm, statistical analysis.