

# 改进的FP学习和综合方法\*

吴英全 王建波 刘明珠

(哈尔滨工业大学 哈尔滨 150001)

**摘要** 对FP算法进行了改进,首先对“共同特征位”进行了优化处理,其次对 $x^0$ 样本的选取进行优化,最后,对阈值设计进行优化处理.经过优化后的样本吸引域明显地较为均匀,模拟实验数据也证明了这一点.

**关键词** 神经网络,多层前馈网络,FP算法,共同特征位,吸引域.

**中图法分类号** TP18

对多层前馈神经网络,目前最流行的学习算法是BP算法,但此算法存在几个主要缺点:①学习收敛速度慢;②产生大量假吸引中心;③所得网络容错能力差. Arai在文献[1]中证明了用三层网络可以完成任何样本的对应,并证明作为通用联想记忆器( $p$ 个样本)的三层网络,其隐含的元件至少要有 $(p-1)$ 个,同时,给出这种网络的求解方法.但是所给出的方法计算量比较大,获得的网络性能也不佳,故到目前为止,对NN网络尚无有效的综合方法.

FP算法以记忆样本作权值,得到了良好的效果.样本的吸引半径达到最大,隐层的神经元个数达到了最少,而且网络易于实现.

本文从3个方面对FP算法进行优化.首先,提出共同特征位的概念,对样本共同特征位进行优化,然后进一步对网络的阈值进行优化,最后对网络的特殊输入样本的选取进行优化.

## 1 算法设计

### 第1层设计

给定训练样本集

$$K = \{r^0 = (x^0, y^0), r^1 = (x^1, y^1), \dots, r^{p-1} = (x^{p-1}, y^{p-1})\} \quad x^i \in \{-1, +1\}^n, y^i \in \{-1, +1\}^m \quad i = 0, 1, \dots, p-1.$$

第1层网络的输入 $x$ 与输出 $z$ 的对应关系为

$$z = \text{sgn}(W * x - \theta), \tag{1}$$
$$\text{sgn}(a) = \begin{cases} +1 & a \geq 0 \\ -1 & a < 0 \end{cases}$$

其中 $W$ 为权系数矩阵, $\theta$ 为阈值向量.为此,我们先引用FP算法的第1层结构设计.

**命题1.**  $[2]$   $x^0, x^1, \dots, x^{p-1}$ 是 $p$ 个互不相同的 $n$ 维向量,作 $p$ 个 $(p-1)$ 维向量如下.

$$\begin{aligned} z^0 &= (-1, -1, -1, \dots, -1) \\ z^1 &= (+1, -1, -1, \dots, -1) \\ z^2 &= (-1, +1, -1, \dots, -1) \\ &\dots\dots\dots \\ z^{p-1} &= (-1, -1, -1, \dots, +1) \end{aligned}$$

有

$$z^i = F(x^i) = \text{sgn}(W * x^i - \theta_i) \quad i = 1, 2, \dots, p-1 \tag{2}$$

$$z^0 = F(x^0) = \text{sgn}(W * x^0 - \theta_0) \tag{3}$$

其中

\* 作者吴英全,1974年生,硕士,主要研究领域为二进神经网络算法设计、研究.王建波,1974年生,博士生,主要研究领域为控制系统故障诊断,神经网络在故障检测和诊断中的应用.刘明珠,1941年生,博士,教授,博导,主要研究领域为常微分和延迟微分方程稳定性及数值稳定性.

本文通讯联系人:王建波,哈尔滨 150001,哈尔滨工业大学 458 信箱

本文 1997-08-22 收到原稿,1997-10-21 收到修改稿

$$\theta_i = \begin{cases} n-d_i+1 & n \text{ 为偶数} \\ n-d_i & n \text{ 为奇数} \end{cases} \quad (4)$$

$$w_i = x^i \quad (5)$$

上述权阵设置有如下性质.

命题 2. 如果输入样本  $x$  满足

$$d(x, x^i) < \left[ \frac{d_i+1}{2} \right] \quad i \neq 0, \quad (6)$$

这里  $[ \ ]$  表示取整函数, 则按命题 1 提供的权阵  $W$  和阈值  $\theta$ , 有  $F(x) = x^i$ .

由于第 2 层的设计是把  $x^i$  一一对应地映射到  $y^i$ , 因此, 我们的目标在于, 在海明距离意义下尽可能扩大域  $F^{-1}(x^i)$ , 下面我们对命题 1 作如下改进.

命题 3. 取

$$\theta_i = n-d_i \quad w_i = x^i \quad i = 1, 2, \dots, p-1, \quad (7)$$

同样有

$$\begin{aligned} z^i &= F(x^i) = \text{sgn}(W \times x^i - \theta_i) \quad i = 1, 2, \dots, p-1, \\ z^0 &= F(x^0) = \text{sgn}(W \times x^0 - \theta_0). \end{aligned}$$

命题 3 的证明与命题 1 的证明完全一致.

命题 4. 与  $x^i$  海明距离为  $\left[ \frac{d_i}{2} \right]$  且与  $x^i$  海明距离最小或海明距离小于  $\left[ \frac{d_i}{2} \right]$  的输入样本  $x$  完全收敛到  $x^i (i=0, 1, 2, \dots, m)$ .

证明:  $i \neq j$  时,

$$\begin{aligned} w_j * x - \theta_j &= x^j * x - (n-d_j) = n-2d(x^j, x) - (n-d_j) = d_j - 2d(x^j, x) \\ &\leq d_j - 2\max(d(x^j, x^i) - d(x^i, x), \left[ \frac{d_i}{2} \right] + 1) = \min(-2 * \left[ \frac{d_i}{2} \right] - 2 + d_j, d_j - 2d(x^j, x^i) + 2 * \left[ \frac{d_i}{2} \right]) \\ &\leq \min(-2 * \left[ \frac{d_i}{2} \right] + d_j - 2, 2 * \left[ \frac{d_i}{2} \right] - d_j) < 0 \quad j = 1, 2, \dots, m \end{aligned}$$

当  $i \neq 0$  时,

$$w_i * x - \theta_i = (x^i, x) - (n-d_i) = n-2d(x^i, x) - (n-d_i) = d_i - 2d(x^i, x) \geq 0 \quad \square$$

我们用“ $\wedge$ ”运算代表样本间的共同位,  $v = x^0 \wedge x^1 \wedge x^2 \wedge \dots \wedge x^{p-1}$  表示  $p$  个样本间的共同位构成的集合, 我们称之为“共同特征位”.

命题 5.  $x^0, x^1, \dots, x^{p-1}$  是  $p$  个互不相同的  $n$  维向量, 设

$$v = x^0 \wedge x^1 \wedge x^2 \wedge \dots \wedge x^{p-1} \quad (8)$$

则取

$$\begin{aligned} w_j &= \begin{cases} 0 & j \in v \quad i = 1, 2, \dots, p-1 \\ x_j^i & j \notin v \quad i = 1, 2, \dots, p-1 \end{cases} \quad (9) \\ \theta_i &= n - |v| - d_i \end{aligned}$$

有

$$\begin{aligned} z^i &= F(x^i) = \text{sgn}(W * x^i - \theta) \quad i = 1, 2, \dots, p-1 \\ F(x^0) &= z^0 \end{aligned}$$

实际上, 由于在共同特征位设置权值为零, 即相当于把标准输入样本的共同特征位去掉, 降维后的  $n - |v|$  维输入标准样本  $\tilde{x}^1, \tilde{x}^2, \dots, \tilde{x}^m$  之间的海明距离与原来一致.

命题 3 的意义在于屏蔽掉输入标准样本的共同特征位, 即相当于降维处理. 这是因为输入样本不论哪个标准样本的变形, 在共同特征位的取值都是一致的, 从而无需考虑输入样本在共同特征位的取值.

从第 1 层结构可以看出:  $x^0$  是一个特殊输入样本. 对  $x^0$  的选取我们采用如下方法:

· 求取  $x^0, x^1, x^2, \dots, x^{p-1}$  中  $d_i$  的最小值. 它所对应的样本  $x^i$  我们认为选作  $x^0$  是最合适的, 这是因为这个值越小, 命题 4 定义的吸引域也越小. 于是我们把这个  $x^i$  与原来的  $x^0$  对调.

### 第 2 层设计

命题 6. [2]任给  $p$  个  $m$  维输出样本  $\{y^0, y^1, y^2, \dots, y^{p-1}\}$ , 则存在  $m \times (p-1)$  维权矩阵  $U$  和阈值向量  $\xi$ , 使得

$$y^j = G(z^j) = \text{sgn}(U \times z^j - \xi) \quad j=0, 1, 2, \dots, p-1. \tag{10}$$

综上所述我们有: 训练样本集  $K = \{r^0, r^1, r^2, \dots, r^{p-1}\}$  都是吸引中心.

## 2 模拟结果

我们取 5 个标准的字母  $A, B, C, D, E$  为样本, 要求对它们实现分类, 分成 5 类(如图 1 所示).

我们对标准的样本随机地倒置黑白像素, 分别用本文算法和 FP 算法进行学习, 然后用带噪声的测试样本进行输入, 统计其识别的正确率. 我们定义“正确率”为: 在与  $x^j$  海明距离最近的所有输入样本中, 输出结果吸引到  $y^j$  的百分比(见表 1).

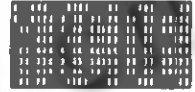


图 1

表 1

学习算法	0 位反转	1 位反转	2 位反转	3 位反转	4 位反转	5 位反转	6 位反转
改进算法	100%	100%	100%	89.5%	74.8%	59.0%	48.0%
FP 算法	100%	100%	100%	60.1%	40.6%	12.7%	9.43%

## 3 结 论

本文的讨论基于 FP 前馈神经网络改进模型. 理论分析和模拟结果都表明, 文中建立的模型有如下优点: (1) 每个样本都是吸引中心; (2) 网络具有最少的元件个数; (3) 网络的学习复杂度和记忆复杂度达到最优; (4) 易于实现知识的扩充.

### 参考文献

- 1 Arai M. Mapping abilities of three-layer neural networks. In: Proceedings of IEEE IJCNN-1989. 1989. 419~423
- 2 张铃, 吴福朝等. 多层前馈神经网络的学习和综合算法. 软件学报, 1995, 6(7): 440~448  
(Zhang ling, Wu Fu-chao et al. A learning and synthesis algorithm of multilayered feedforward neural networks. Journal of Software, 1995, 6(7): 440~448)

## Improvement of the FP Learning and Synthesize Algorithm

WU Ying-quan WANG Jian-bo LIU Ming-zhu

(Harbin Institute of Technology Harbin 150001)

**Abstract** FP algorithm is improved in this paper. First, the “common characteristic bit” is optimized, then the selection of  $x^0$  sample is optimized. Finally, the design of threshold is optimized. After these optimizations, the attractive domain of sample appears more even. The solution is proved by the test data.

**Key words** Neural network, multilayered feedforward network, FP algorithm, common characteristic bit, attractive domain.