

РАЗДЕЛ I. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И НЕЙРОННЫЕ СЕТИ

ОБУЧЕНИЕ МНОГОСЛОЙНЫХ НЕЙРОННЫХ СЕТЕЙ
С ИСПОЛЬЗОВАНИЕМ МАТРИЧНЫХ ОПЕРАЦИЙ

Грязев Е.В., БГТУ, Брест

Рассмотрим многослойную гетерогенную нейронную сеть [1], состоящую из N нейронных блоков (рис.1), каждый из которых имеет структуру, представленную на рис. 2.

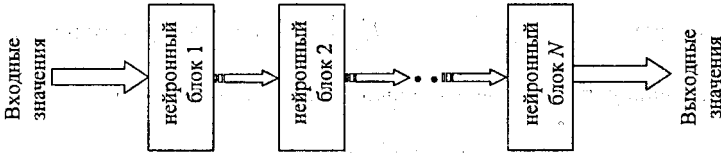


Рис.1. Блочное представление многослойной нейронной сети

Входными значениями для каждого нейронного блока являются выходы предыдущего; для первого – последовательность входных образов $\bar{x}^k = (x_1^k, \dots, x_{m_0}^k)$, ($k = \overline{1, L}$). Выходное значение i_n -го нейрона n -ого блока сети для k -ого образа определяется рекуррентным соотношением:

$$y_{i_n}^{(n),k} = F_n(S_{i_n}^{(n),k}),$$

где

$$S_{i_n}^{(n),k} = \sum_{i_{n-1}=1}^{m_{n-1}} w_{i_{n-1} i_n}^{(n)} y_{i_{n-1}}^{(n-1),k} - T_{i_n}^{(n)}, \quad i_n = \overline{1, m_n}, \quad k = \overline{1, L}$$

При этом формируется вектор $Y^{(n),k} = (y_1^{(n),k} \quad y_2^{(n),k} \quad \dots \quad y_{m_n}^{(n),k} \quad -1)^T$.

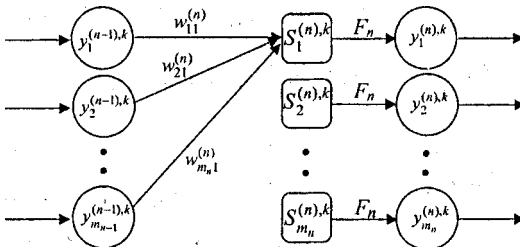


Рис.2. Архитектура n -ого блока многослойной нейронной сети

Задача обучения данной многослойной гетерогенной нейронной сети состоит в нахождении матриц весовых коэффициентов

$$W^{(n)} = \begin{pmatrix} W_{11}^{(n)} & W_{21}^{(n)} & \dots & W_{m_{n-1}1}^{(n)} \\ W_{12}^{(n)} & W_{22}^{(n)} & \dots & W_{m_{n-1}2}^{(n)} \\ \dots & \dots & \dots & \dots \\ W_{1m_n}^{(n)} & W_{2m_n}^{(n)} & \dots & W_{m_{n-1}m_n}^{(n)} \end{pmatrix}_{m_n \times m_{n-1}}$$

и столбцов порогов $\overline{T^{(n)}} = (T_1^{(n)}, T_2^{(n)}, \dots, T_{m_n}^{(n)})^T$, $n = \overline{1, N}$, которые минимизируют некоторую ошибку сети E_S , как отклонение выходных значений сети $y_{i_N}^{(N),k}$ от эталонных $t_{i_N}^k$ — i_N -ого нейрона сети для k -ого образа. В качестве ошибки рассматривается усредненное по количеству образов «квадратичное отклонение»

$$E_S = \frac{1}{2L} \sum_{k=1}^L \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k)^2.$$

Теорема 1. Модификация синаптических связей и порогов многослойной гетерогенной нейронной сети (рис.2) производится в соответствии с формулами:

$$w_{j_n - j_n}^{(n)}(t+1) = w_{j_n - j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k},$$

$$T_{j_n}^{(n)}(t+1) = T_{j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n(m_{n-1}+1)}^{(n)} \cdot Y^{(n-1),k},$$

где $C^{(n)}$ вычисляется рекуррентно:

$$C^{(n)} = C^{(n+1)} \cdot W^{(n+1)} \cdot MF_n, \quad C^{(N)} = \varepsilon^k \cdot MF_N,$$

$$\varepsilon^k = \left((y_1^{(2),k} - t_1^k) \quad (y_2^{(2),k} - t_2^k) \quad \dots \quad (y_{m_2}^{(2),k} - t_{m_2}^k) \right),$$

$$MF_n = \begin{pmatrix} F_n'(S_1^{(n),k}) & 0 & \dots & 0 \\ 0 & F_n'(S_2^{(n),k}) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F_n'(S_{m_n}^{(n),k}) \end{pmatrix} \text{ — матрица размерности } m_n \times m_n,$$

а матрица $M_{j_n j_{n-1}}^{(n)}$ размерности $m_n \times (m_{n-1} + 1)$ состоит из числа 1 на позиции $j_n j_{n-1}$ и нулей в качестве остальных элементов матрицы.

Изменение синаптических связей и порогов сети производится начиная с последнего N -ого до первого блока сети.

Доказательство: Вычислим градиент ошибки для k -ого образа:

$$\begin{aligned} \frac{\partial E_s^{(k)}}{\partial w_{j_n-1 j_n}^{(n)}} &= \frac{\partial \left(\sum_{i_N=1}^{m_N} \frac{1}{2} (y_{i_N}^{(N),k} - t_{i_N}^k)^2 \right)}{\partial w_{j_n-1 j_n}^{(n)}} = \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k) \cdot \frac{\partial y_{i_N}^{(N),k}}{\partial w_{j_n-1 j_n}^{(n)}} = \\ &= \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k) \cdot F_N' (S_{i_N}^{(N),k}) \cdot \frac{\partial S_{i_N}^{(N),k}}{\partial w_{j_n-1 j_n}^{(n)}} = \\ &= \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k) \cdot F_N' (S_{i_N}^{(N),k}) \cdot \sum_{i_{N-1}=1}^{m_{N-1}} w_{i_N-1 i_N}^{(N)} \cdot \frac{\partial y_{i_{N-1}}^{(N-1),k}}{\partial w_{j_n-1 j_n}^{(n)}} = \\ &= \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k) \cdot F_N' (S_{i_N}^{(N),k}) \cdot \sum_{i_{N-1}=1}^{m_{N-1}} w_{i_N-1 i_N}^{(N)} \cdot F_{N-1}' (S_{i_{N-1}}^{(N-1),k}) \cdot \frac{\partial S_{i_{N-1}}^{(N-1),k}}{\partial w_{j_n-1 j_n}^{(n)}} = \\ &= \dots = \\ &= \sum_{i_N=1}^{m_N} (y_{i_N}^{(N),k} - t_{i_N}^k) \cdot F_N' (S_{i_N}^{(N),k}) \cdot \sum_{i_{N-1}=1}^{m_{N-1}} w_{i_N-1 i_N}^{(N)} \cdot F_{N-1}' (S_{i_{N-1}}^{(N-1),k}) \cdot \dots \cdot \\ &\dots \cdot \sum_{i_n=1}^{m_n} w_{i_n-1 i_n}^{(n+1)} \cdot F_n' (S_{i_n}^{(n),k}) \cdot y_{j_n-1}^{(n-1),k} \cdot \delta_{j_n}^{(n)} = \\ &= \varepsilon^k \cdot MF_N \cdot W^{(N)} \cdot MF_{N-1} \cdot \dots \cdot W^{(n+1)} \cdot MF_n \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k} = \\ &= C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k}, \end{aligned}$$

где

$$C^{(n)} = C^{(n+1)} \cdot W^{(n+1)} \cdot MF_n, \quad C^{(N)} = \varepsilon^k \cdot MF_N.$$

Т.к. $\frac{\partial E_s}{\partial z} = -\frac{\partial \left(\frac{1}{L} \sum_{k=1}^L E_s^{(k)} \right)}{\partial z} = \frac{1}{L} \sum_{k=1}^L \frac{\partial E_s^{(k)}}{\partial z}$, то формулы настройки синаптиче-

ских связей $w_{j_n-1/j_n}^{(n)}(t+1) = w_{j_n-1/j_n}^{(n)}(t) - \alpha^{(n)} \frac{\partial E_s}{\partial w_{j_n-1/j_n}^{(n)}}$, $j_{n-1} = \overline{1, m_{n-1}}$, $j_n = \overline{1, m_n}$, при-

нимают вид:

$$w_{j_n-1/j_n}^{(n)}(t+1) = w_{j_n-1/j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot y^{(n-1),k}$$

Аналогичным образом получаются формулы для настройки порогов нейронной сети.

Теорема доказана.

Данная теорема позволяет привести к алгоритму изменение весов и порогов сети в процессе обучения градиентным методом сведением их модификации к матричным операциям. Предложенный алгоритм обучения был реализован в среде MatLab 6.1.

Литература. 1. Осовский С. Нейронные сети для обработки информации / Пер. с польского Н.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.

АЛГОРИТМ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СОПРЯЖЕННЫХ ГРАДИЕНТОВ

Кокош Н.В., БГУ, г. Минск

1. Введение

Рассмотрим нейронную сеть, состоящую из n нейронных элементов распределительного слоя и m - выходного слоя (рис. 1).

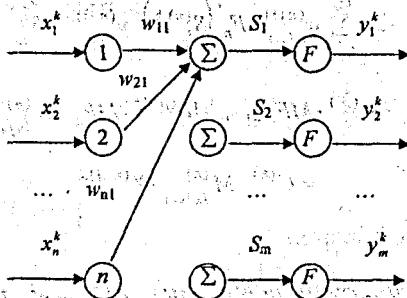


Рис. 1. Схема функционирования нейронной сети