

**ВЫБОР АДАПТИВНОГО ШАГА ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ С  
ИСПОЛЬЗОВАНИЕМ МЕТОДА НАЙСКОРЕЙШЕГО СПУСКА**

**И.И. Гладкий**

*(БГТУ, Брест)*

Задача обучения нейронной сети с фиксированной функцией активации  $F$  состоит в нахождении весовых коэффициентов  $\omega_{ij}$  ( $i = \overline{1, n}$ ,  $j = \overline{1, m}$ ) и порогов нейронных элементов  $T_j$  ( $j = \overline{1, m}$ ), которые минимизируют некоторую

ошибку сети  $E_S$ , как отклонение выходных значений  $y_j^k(t)$  от эталонных значений  $t_j^k$  -  $j$ -ого нейрона сети для  $k$ -ого образа. Для определения адаптивного шага обучения  $\alpha(t)$  необходимо определить такой шаг обучения сети, который минимизирует среднеквадратичную ошибку  $E_S(t+1) = \sum_k \sum_j E_j^k(t+1) = \frac{1}{2} \sum_k \sum_j (y_j^k(t+1) - t_j^k)^2$ .

**Теорема 1.** Для нейронной сети величина адаптивного шага обучения  $\alpha(t)$  с использованием метода наискорейшего спуска в момент времени  $t$  определяется соотношением:

$$\alpha(t) = \frac{\sum_{k=1}^L \sum_{j=1}^m (y_j^k(t) - t_j^k) F'(S_j^k(t)) a_j^k(t)}{\sum_{k=1}^L \sum_{j=1}^m \left( (F'(S_j^k(t)))^2 + (y_j^k(t) - t_j^k) F''(S_j^k(t)) \right) (a_j^k(t))^2}, \quad (1)$$

где  $S_j^k(t) = \sum_i \omega_{ji}(t) x_i^k - T_j(t)$ ,  $a_j^k(t) = \sum_{p=1}^L (y_j^p(t) - t_j^p) F'(S_j^p(t)) \left( \sum_{i=1}^n x_i^p x_i^k + 1 \right)$ ,  
 $j = \overline{1, m}$ ,  $k = \overline{1, L}$ .

**Следствие 1.** В случае одного  $j$  - ого выходного нейронного элемента, соотношение (1), примет вид:

$$\alpha(t) = \frac{\sum_{k=1}^L F'(S_j^k) (y_j^k - t_j^k) a_j^k}{\sum_{k=1}^L \left( (F'(S_j^k))^2 + (y_j^k - t_j^k) F''(S_j^k) \right) (a_j^k)^2},$$

где  $a_j^k = \sum_{p=1}^L (y_j^p - t_j^p) F'(S_j^p) \left( \sum_{i=1}^n x_i^p x_i^k + 1 \right)$ ,  $k = \overline{1, L}$ .

**Следствие 2.** В случае одного образа, т. е. при  $L = 1$ , соотношение (1), принимает вид:

$$\alpha(t) = \frac{\sum_{j=1}^m F'(S_j) (y_j - t_j) a_j}{\sum_{j=1}^m \left( (F'(S_j))^2 + (y_j - t_j) F''(S_j) \right) (a_j)^2},$$

Материалы V Республиканской научной конференции студентов и аспирантов "Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях", Гомель 18-20 марта 2002 г.

---

где  $a_j = (y_j - t_j) F^n \left( S_j^p \left( \sum_{i=1}^n x_i^2 + 1 \right) \right)$ ,  $j = \overline{1, m}$ .