

# Лекция 6

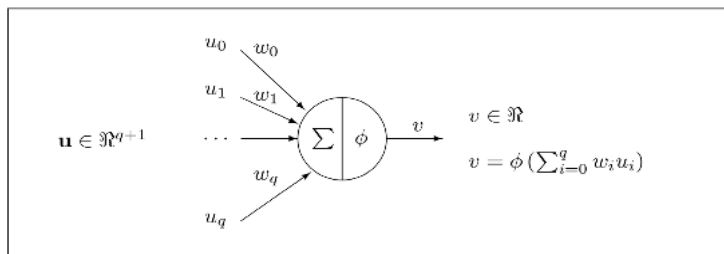
## Нейросетевые методы, перцептрон Розенблатта, многослойный перцептрон

*Лектор – Сенько Олег Валентинович*

Курс «Математические основы теории прогнозирования»  
4-й курс, III поток

- 1 Нейросетевые методы
- 2 перцептрон Розенблатта
- 2 многослойный перцептрон
- 3 метод обратного распространения ошибки

В основе нейросетевых методов лежит попытка компьютерного моделирования процессов обучения, используемых в живых организмах. Когнитивные способности живых существ связаны с функционированием сетей связанных между собой биологических нейронов – клеток нервной системы. Для моделирования биологических нейросетей используются сети, узлами которых являются искусственные нейроны (т.е. математические модели нейронов), Можно выделить три типа искусственных нейронов: нейроны-рецепторы, внутренние нейроны и реагирующие нейроны. Каждый внутренний или реагирующий нейрон имеет множество входных связей, по которым поступают сигналы от рецепторов или других внутренних нейронов. Пример модели внутреннего или реагирующего нейрона представлен на рисунке 1.



**Рис.1. Модель внутреннего или реагирующего нейрона.**

Представленный на рисунке 1 нейрон имеет  $r$  внешних связей, по которым на него поступают входные сигналы  $u_1, \dots, u_r$ . Поступившие сигналы суммируются с весами  $w_1, \dots, w_r$ . На выходе нейрона вырабатывается сигнал  $\Phi(z)$ , где  $z = w_0 + \sum_{i=1}^r w_i u_i$ ,  $w_0$  - параметр сдвига. Может быть использована также форма записи  $z = \sum_{i=0}^r w_i u_i$ , где фиктивный «сигнал»  $u_0$  тождественно равен 1.

Функцию  $\Phi(z)$  обычно называют активационной функцией. Могут использоваться различные виды активационных функций, включая

- пороговую функцию, задаваемую с помощью пороговой величины  $b$ :  $\Phi(z) = 1$  при  $z \geq b$ ,  $\Phi(z) = 0$  при  $z < b$ ;
- - сигмоидная функция  $\Phi(z) = \frac{1}{1+e^{-az}}$ , где  $a$ -вещественная константа;
- гиперболический тангенс;
- тождественное преобразование  $\Phi(z) = z$ .

Методы, основанные на использовании искусственных нейронов могут быть использованы для решения самых разнообразных задач распознавания. При этом сигналы, поступающие на вход перцептрона, интерпретируются как входные признаки  $X_1, \dots, X_n$ .

Первой нейросетевой моделью стал перцептрон Розенблатта, предложенный в 1957 году. В данной модели используется единственный реагирующий нейрон. Модель, реализующая линейную разделяющую функцию в пространстве входных сигналов, может быть использована для решения задач распознавания с двумя классами, помеченными метками 1 или -1. В качестве активационной функции используется пороговая функция:  $\Phi(z) = 1$  при  $z \geq 0$ ,  $\Phi(z) = -1$  при  $z < 0$ . Особенностью модели Розенблатта является очень простая, но вместе с тем эффективная, процедура обучения, вычисляющая значения весовых коэффициентов  $(w_0, \dots, w_n)$ . Настройка параметров производится по обучающим выборкам, совершенно аналогичных тем, которые используются для обучения статистических алгоритмов. На первом этапе производится преобразование векторов сигналов (признаковых описаний) для объектов обучающей выборки. В набор исходных признаков добавляется тождественно равная 1 нулевая компонента. Затем вектора описаний из класса  $K_2$  умножаются на -1. Вектора описаний из класса  $K_1$  не изменяются.

**Процедура обучения перцептрона.** Нулевое приближение вектора весовых коэффициентов  $(w_0^{(0)}, \dots, w_n^{(0)})$  выбирается случайным образом. Преобразованные описания объектов обучающей выборки  $\tilde{S}_t$  последовательно подаются на вход перцептрона. В случае если описание  $\mathbf{x}^{(k)}$ , поданное на  $k$ -ом шаге классифицируется неправильно, то происходит коррекция по правилу  $w^{(k+1)} = w^{(k)} + \mathbf{x}$ . В случае правильной классификации  $w^{(k+1)} = w^{(k)}$ . Отметим, что правильной классификации всегда соответствует выполнение равенства  $(w^{(k)}, \mathbf{x}^{(k)}) \geq 0$ , а неправильной  $(w^{(k)}, \mathbf{x}^{(k)}) < 0$ . Процедура повторяется до тех пор, пока не будет выполнено одно из следующих условий:

- - достигается полное разделение объектов из классов  $K_1$  и  $K_2$ ;
- - повторение подряд заранее заданного числа итераций не приводит к улучшению разделения;
- - оказывается исчерпанным заранее заданный лимит итераций.

Для описанной процедуры обучения справедлива следующая теорема.  
**Теорема.** В случае, если описания объектов обучающей выборки линейно разделимы в пространстве признаков описаний, то процедура обучения перцептрона построит линейную гиперплоскость разделяющую объекты двух классов за конечное число шагов.

Отсутствие линейной разделимости двух классов приводит к бесконечному закликиванию процедуры обучения перцептрона. Существенно более высокой аппроксимирующей способностью обладают нейросетевые методы распознавания, задаваемые комбинациями связанных между собой нейронов. Таким методом является **Многослойный перцептрон**.



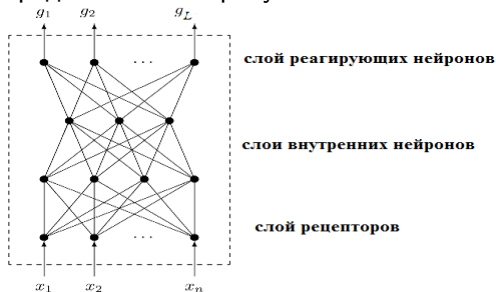
В методе Многослойный перцептрон сеть формируется из нескольких слоёв нейронов. В их число входит слой входных рецепторов, подающих сигналы на нейроны из внутренних слоёв. Слои внутренних нейронов осуществляют преобразование сигналов. Слой реагирующих нейронов производит окончательную классификацию объектов на основании сигналов, поступающих от нейронов, принадлежащих внутренним слоям.

Обычно соблюдаются следующие правила формирования структуры сети.

- Допускаются связи между только между нейронами, находящимися в соседних слоях.
- Связи между нейронами внутри одного слоя отсутствуют.
- Активационные функции для всех внутренних нейронов идентичны.

Для решения задач распознавания с  $L$  классами  $K_1, \dots, K_L$  используется конфигурация с  $L$  реагирующими нейронами.

Схема многослойного перцептрона с двумя внутренними слоями представлена на рисунке 2.



Отметим, что сигналы  $g_1, \dots, g_L$ , вычисляемые на выходе реагирующих нейронов, интерпретируются как оценки за классы  $K_1, \dots, K_L$ . Весовые коэффициенты  $w$  сопоставлены каждой из связей между нейронами из различных слоёв.

Рассмотрим процедуру распознавания объектов с использованием многослойного перцептрона. Предположим, что конфигурация нейронной сети включает наряду со слоем рецепторов и слоем реагирующих нейронов также  $H$  внутренних слоёв искусственных нейронов. Заданы также количества нейронов в каждом слое. Пусть  $n$  – число входных нейронов-рецепторов,  $r(h)$  – число нейронов в внутреннем слое  $h$ . На первом этапе вектор рецепторы формируют по информации, поступающей из внешней среды, вектор входных переменных (сигналов)  $u_1^0, \dots, u_n^0$ . Отметим, что входные сигналы  $u_1^1, \dots, u_n^0$  могут интерпретироваться как признаки  $X_1, \dots, X_n$  в общей постановке задачи распознавания.

Предположим, что для  $i$ -го нейрона 1-го внутреннего слоя связь с рецепторами осуществляется с помощью весовых коэффициентов  $w_1^{i0}, \dots, w_n^{i0}$ . Сумматор  $i$ -го нейрона первого внутреннего слоя вычисляет взвешенную сумму  $\xi^{i0} = \sum_{t=0}^n w_t^{i0} u_t^0$ .

Сигнал на выходе  $i$ -го нейрона первого внутреннего слоя вычисляется по формуле  $u_i^1 = \Phi(\xi^{i0})$ . Аналогичным образом вычисляются сигналы на выходе нейронов второго внутреннего слоя. Сигналы  $g_1, \dots, g_L$  рассчитываются с помощью той же самой процедуры, которая используется при вычислении сигналов на выходе нейронов из внутренних слоёв. То есть при вычислении  $g_i$  на первом шаге соответствующий сумматор вычисляет взвешенную сумму

$$\xi^{iH} = \sum_{t=0}^n w_t^{iH} u_t^H,$$

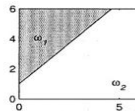
где  $w_1^{iH}, \dots, w_{r(H)}^{iH}$  - весовые коэффициенты, характеризующие связь  $i$ -го реагирующего нейрона с нейронами последнего внутреннего слоя  $H$ ,  $u_1^H, \dots, u_{r(H)}^H$  - сигналы на выходе внутреннего слоя  $H$ . Сигнал на выходе  $i$ -го реагирующего нейрона вычисляется по формуле  $g_i = \Phi(\xi^{iH})$ . Очевидно, что вектор выходных сигналов является функцией вектора входных сигналов  $u^0$  (вектора признаков  $x$ ) и матрицы весовых коэффициентов связей между нейронами.

Один реагирующий нейрон позволяет аппроксимировать области, являющиеся полупространствами, ограниченными гиперплоскостями. Нейронная сеть с одним внутренним слоем позволяет аппроксимировать произвольную выпуклую область в многомерном признаковом пространстве (открытую или закрытую). Было доказано также, что МП с двумя внутренними слоями позволяет аппроксимировать произвольные области многомерного признакового пространства. Аппроксимирующая способность многослойного перцептрона с различным числом внутренних слоёв проиллюстрирована на рисунке 3. На рисунке области, соответствующие классам  $\omega_1$  и  $\omega_2$  разделяются с помощью простого нейрона, а также с помощью многослойных перцептронов с одним и двумя внутренними слоями.

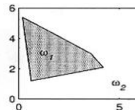
конфигурация НС



полупространства,  
ограниченные  
гиперплоскостями



выпуклые области  
(открытые или закрытые)



произвольные области

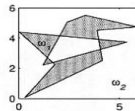


Рис.3

Наиболее распространённым способом обучения нейросетевых алгоритмов является метод обратного распространения ошибки. Обозначим через  $\alpha_* = (\alpha_{*1}, \dots, \alpha_{*L})$  вектор индикаторных функций классов  $K_1, \dots, K_L$  на объекте  $s_*$  с описанием  $\mathbf{x}_*$ . То есть  $\alpha_{*i} = 1$ , если  $s_* \in K_i$  и  $\alpha_{*i} = 0$  в противном случае. Пусть на выходе  $i$ -го реагирующего нейрона вычисляется оценка  $g_i(\mathbf{x}_*)$  за класс  $K_i$ , принадлежащая отрезку  $[0, 1]$ . Отметим, что оценка  $g_i(\mathbf{x}_*)$  вычисляется активационной функцией реагирующего нейрона. Далее будет предполагаться, что данная активационная функция является сигмоидной. Такое же предположение делается для активационных функций каждого из внутренних нейронов. Потери, связанные с классификацией объекта  $s_*$  естественно оценивать с помощью функционала

$$\sum_{i=1}^L [\alpha_{*i} - g_i(\mathbf{x}_*)]^2.$$

Качество аппроксимации на обучающей выборке

$\tilde{S}_t = \{(\alpha_1, \mathbf{x}_1), \dots, (\alpha_m, \mathbf{x}_m)\}$  оценивается с помощью функционала

$$E(\tilde{S}_t, \tilde{\mathbf{w}}) = \sum_{j=1}^m \sum_{i=1}^L [\alpha_{ji} - g_i(\mathbf{x}_j)]^2.$$

Где  $\tilde{\mathbf{w}} = \{w_t^{ih} \mid h = 0, \dots, H; t = 1, \dots, r(h); i = 1, \dots, (r_{h+1})\}$  - множество весовых коэффициентов связей между нейронами. Обучение заключается в поиске значений коэффициентов из  $\tilde{\mathbf{w}}$ , при которых достигает минимума функционал  $E(\tilde{S}_t, \tilde{\mathbf{w}})$ . В основе обучения лежит метод градиентного спуска. Метод градиентного спуска является итерационным методом оптимизации произвольного функционала  $F$ , зависящего от параметров  $(\theta_1, \dots, \theta_r)$  и дифференцируемого по каждому из параметров в произвольной точке  $\mathbb{R}^n$ . Новые значения вектора параметров на  $k$ -ой итерации  $\theta^{(k)}$  вычисляется через вектор  $\theta^{(k-1)}$ , полученный на предыдущей итерации.



При этом используется формула

$$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + \eta \times \mathbf{grad}[F(\theta_1, \dots, \theta_r)],$$

где  $\eta > 0$  - вещественный параметр, задающий размер каждого шага. На предварительном этапе обучения весовым коэффициентам из  $\tilde{\mathbf{w}}$  случайным образом присваиваются исходные значения. На обучение подаётся некоторый объект обучающей выборки  $s_j = (\boldsymbol{\alpha}_j, \mathbf{x}_j)$ , по описанию которого вычисляются входные и выходные сигналы внутренних нейронов сети, а также выходные сигналы реагирующих нейронов  $g_1(\mathbf{x}_j, \tilde{\mathbf{w}}), \dots, g_L(\mathbf{x}_j, \tilde{\mathbf{w}})$ . Проведём коррекцию весовых коэффициентов связей  $i$ -го реагирующего нейрона с нейронами предшествующего внутреннего слоя:

$$(w_0^{iH}, \dots, w_{r(H)}^{iH}).$$

Для упрощения формул  $g_i(\mathbf{x}_j, \tilde{\mathbf{w}})$  будем обозначать  $g_i(\mathbf{x}_j)$  или просто  $g_i$ .

От весовых коэффициентов  $(w_0^{iH}, \dots, w_{r(H)}^{iH})$  зависит только компонента  $[\alpha_{ji} - g_i(\mathbf{x}_j)]^2$  ошибки прогнозирования для объекта  $s_j$ , равная  $E(s_j, \tilde{\mathbf{w}}) = \sum_{i=1}^L [\alpha_{ji} - g_i(\mathbf{x}_j)]^2$ . Поэтому

$$\begin{aligned} \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{iH}} &= \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial g_i(\mathbf{x}_j)} \frac{\partial g_i}{\partial w_t^{iH}} = \\ &= -2[\alpha_{ji} - g_i(\mathbf{x}_j)] \frac{\partial g_i(\mathbf{x}_j)}{\partial w_t^{iH}} \end{aligned}$$

Однако

$$\frac{\partial g_i(\mathbf{x}_j)}{\partial w_t^{iH}} = \frac{\partial g_i(\mathbf{x}_j)}{\partial \xi^{iH}} \frac{\partial \xi^{iH}}{\partial w_t^{iH}},$$

где  $\xi^{iH} = \sum_{t=1}^{r(H)} w_t^{iH} u_t^H$ ,  $u_t^H$  - сигнал на выходе нейрона с номером  $t$  из слоя  $H$ .

## Метод обратного распространения ошибки

Поскольку  $g_i$  является сигмоидной функцией от  $\xi^{iH}$ , то

$$\frac{\partial g_i}{\partial w_t^{iH}} = (1 - g_i)g_i u_t^H.$$

Таким образом

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{iH}} = \delta^{iH} u_t^H,$$

где

$$\delta^{iH} = \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial \xi^{iH}} = -2[\alpha_{ji} - g_i(\mathbf{x}_j)][1 - g_i(\mathbf{x}_j)]g_i(\mathbf{x}_j).$$

Воспользовавшись методом градиентного спуска, запишем новые значения весовых коэффициентов  $w_t^{iH}(k)$ , вычисляемые на  $k$  итерации  $k$  в виде

$$w_t^{iH}(k) = w_t^{iH}(k-1) + \eta \times \delta^{iH} u_t^H$$

Рассмотрим теперь коррекцию весовых коэффициентов  $[w_0^{i(H-1)}, \dots, w_r^{i(H-1)}]$ , соответствующих связям нейрона  $i$  из слоя  $H$  с нейронами предшествующего внутреннего слоя  $(H - 1)$ . Вклад этих коэффициентов в величину ошибки осуществляется только через сигнал  $u_i^H$  на выходе нейрона  $i$  из слоя  $H$ . Поэтому

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{i(H-1)}} = \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial u_i^H} \frac{\partial u_i^H}{\partial w_t^{i(H-1)}}$$

Однако

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial u_t^H} = \sum_{l=1}^L \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial \xi^{lH}} \frac{\partial \xi^{lH}}{\partial u_t^H}$$

Принимая во внимание, что

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial \xi^{lH}} = \delta^{lH},$$

а также, что  $\frac{\partial \xi^{lH}}{\partial u_t^H} = w_t^{lH}$ ,  
получаем

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial u_t^H} = \sum_{l=1}^L \delta^{lH} w_t^{lH}$$

Исходя из предположения о том, что активационная функция каждого из нейронов является сигмоидной, нетрудно показать также, что

$$\frac{\partial u_i^H}{\partial w_t^{i(H-1)}} = u_i^H (1 - u_i^H) u_t^{H-1}$$

В итоге

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{i(H-1)}} = \left( \sum_{l=1}^L \delta^{lH} w_i^{lH} \right) u_i^H (1 - u_i^H) u_t^{H-1} = \delta^{i(H-1)} u_t^{H-1},$$

где

$$\delta^{i(H-1)} = \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial \xi^{i(H-1)}} = \left( \sum_{l=1}^L \delta^{lH} w_i^{lH} \right) u_i^H (1 - u_i^H)$$

Воспользовавшись методом градиентного спуска, запишем новые значения весовых коэффициентов  $w_t^{i(H-1)}(k)$ , вычисляемые на итерации  $k$  в форме

$$w_t^{i(H-1)}(k) = w_t^{i(H-1)}(k-1) + \eta \times \delta^{i(H-1)} u_t^{H-1}$$

Рассмотрим теперь процедуру коррекции весовых коэффициентов  $w$  для связей между искусственными нейронами из слоя  $h$  с искусственными нейронами из слоя  $h + 1$  при  $h < H - 1$ . Пусть

$$[w_0^{i(H-h)}, \dots, w_{r(H-h)}^{i(H-h)}]$$

- весовые коэффициенты, связывающие нейрон с номером  $i$  из слоя  $H - h + 1$  с нейронами из слоя  $H - h$ . Очевидно, что справедливо равенство:

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{i(H-h)}} = \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial u_i^{H-h+1}} \frac{\partial u_i^{H-h+1}}{\partial w_t^{i(H-h)}}$$

Нетрудно показать, что

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial u_i^{H-h+1}} = \sum_{l=1}^{r(H-h+1)} \frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial \xi^{H-h+1}} \frac{\partial \xi^{H-h+1}}{\partial u_i^{H-h+1}} = \sum_{l=1}^{r(H-h+1)} \delta^{l(H-h+1)} w_i^{l(H-h+1)}$$

Учитывая, что активационная функция для каждого внутреннего нейрона является сигмоидной, и принимая во внимание определение  $\xi^{i(H-h)}$ , получаем

$$\frac{\partial u_i^{H-h+1}}{\partial w_t^{i(H-h)}} = \frac{\partial u_i^{H-h+1}}{\partial \xi^{i(H-h)}} \frac{\partial \xi^{i(H-h)}}{\partial w_t^{i(H-h)}} = u_i^{H-h+1} (1 - u_i^{H-h+1}) u_t^{H-h}$$

В итоге получаем

$$\frac{\partial E(s_j, \tilde{\mathbf{w}})}{\partial w_t^{i(H-h)}} = \delta^{i(H-h)} u_t^{H-h},$$

где

$$\delta^{i(H-h)} = \left[ \sum_{l=1}^{r(H-h+1)} \delta^{l(H-h+1)} w_i^{l(H-h+1)} \right] u_i^{H-h+1} (1 - u_i^{H-h+1})$$



Коррекция согласно процедуре градиентного спуска производится по формуле:

$$w_t^{i(H-h)}(k) = w_t^{i(H-h)}(k-1) + \eta \times \delta^{i(H-1)} w_t^{H-h}$$

Таким образом может быть представлен общая схема метода обратного распространения ошибки для многослойного перцептрона. На предварительном этапе выбирается архитектура сети: задаётся число внутренних слоёв и количества нейронов в каждом слое. Случайным образом задаются исходные весовые коэффициенты. На вход многослойного перцептрона поочерёдно подаются векторные описания объектов обучающей выборки. С использованием описанного выше способа производится коррекция весовых коэффициентов. С использованием новых скорректированных весовых коэффициентов  $\tilde{\mathbf{w}}$  вычисляется значение функционала  $E(s_j, \tilde{\mathbf{w}})$ .

Обучение заканчивается при выполнении одного из заранее заданных условий:

- а) Величина функционала ошибки оказывается меньше выбранного порогового значения:  $E(s_j, \tilde{\mathbf{w}}) < \varepsilon$ ;
- б) Изменения функционала ошибки на протяжении нескольких последних итераций оказывается меньшим некоторого порогового значения.
- в) общее время обучения превышает допустимый предел;