

## Раздел 7. Искусственные нейронные сети

### 7.1. Введение

Известно, что способ обработки информации человеческим мозгом в корне отличается от методов, применяемых обычными числовыми компьютерами. Мозг представляет собой чрезвычайно сложную, нелинейную, параллельную систему обработки информации. Он обладает способностью организовывать свои структурные компоненты, называемые нейронами, так, чтобы они могли выполнять такие задачи как распознавание образов, обработку сигналов органов чувств, моторные функции во много раз быстрее, чем могут позволить самые быстродействующие современные компьютеры. Например, на распознавание знакомого лица в незнакомом окружении человеческий мозг тратит 100-200 миллисекунд, в то время как выполнение аналогичной задачи на компьютере может занять несколько дней.

Нейронные сети можно рассматривать как современные вычислительные системы, которые преобразуют информацию в определенной степени по образу процессов, происходящих в мозгу человека. Обрабатываемая информация имеет численный характер, что позволяет использовать нейронную сеть, например, в качестве модели объекта с совершенно неизвестными характеристиками. Другие типовые приложения нейронных сетей охватывают задачи распознавания, классификации, анализа и сжатия образов.

Чтобы добиться высокой производительности, нейронные сети используют для хранения информации множество взаимосвязей между элементарными ячейками вычислений – нейронами. Для получения необходимой структуры взаимосвязей нейронов в нейронной сети применяется процедура, называемая алгоритмом обучения. Таким образом, сходство искусственной нейронной сети с человеческим мозгом состоит в том, что:

- Знания поступают в нейронную сеть из окружающей среды и используются в процессе обучения.
- Для накопления знаний применяются связи между нейронами, называемые синаптическими..

Использование нейронных сетей обеспечивает следующие преимущества:

- Нелинейность. Нейронные сети позволяют получить нелинейную зависимость выходного сигнала от входного.
- Адаптивность. Нейронные сети обладают способностью адаптировать свои синаптические веса к изменениям окружающей среды. Более того, для работы в нестационарной среде (где статистика изменяется со временем) могут быть созданы нейронные сети, изменяющие синаптические веса в реальном времени.
- Отказоустойчивость. Нейронные сети, реализованные на основе электронных компонентов, потенциально отказоустойчивы, так как контекстная информация распределена по всем связям нейронной сети, и выход из строя одного или нескольких узлов не приводит к отказу системы в целом.
- Единообразии анализа и проектирования. Нейронные сети позволяют решать сложные вычислительные задачи с помощью единообразных простых вычислительных узлов

– нейронов. Эта общность позволяет применять одни и те же алгоритмы обучения для решения различных прикладных задач.

## 7.2. Нейрон и его модели

Базовый элемент нервной системы – это нервная клетка, называемая нейроном. В нейроне можно выделить тело клетки (сома), а также исходящие из него два вида отростков: дендриты, по которым в нейрон поступает информация, и аксон, по которому нейрон передает информацию (рис. 7.1). Каждый нейрон имеет только один аксон, по которому он может передавать информацию нескольким другим нейронам.

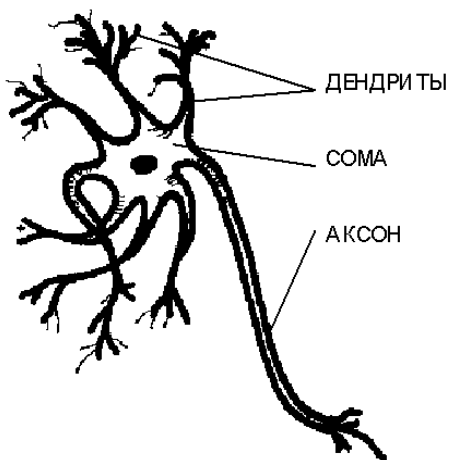


Рис. 7.1. Биологический нейрон

В нервной системе один нейрон принимает возбуждения от огромного количества нейронов (их число может достигать тысячи). Считается, что мозг человека состоит из порядка  $10^{11}$  нейронов, которые имеют между собой порядка  $10^{15}$  связей. Каждый нейрон передает возбуждение через нервные стыки (синапсы), при этом процесс передачи информации имеет сложную электрохимическую природу. Синапсы функционируют как репитеры информации, в результате работы которых возбуждение может усиливаться или ослабляться. В результате к нейрону приходят сигналы, оказывающие как тормозящее, так и возбуждающее воздействие.

### 7.2.1. Искусственный нейрон

Рассмотрим формализованную модель искусственного нейрона. Искусственный нейрон состоит из синапсов, каждому из которых соответствует определенный вес синаптической связи, сумматора и функции активации. Введем следующие обозначения:  $x_1, \dots, x_n$  – входные сигналы, приходящие от других нейронов;  $w_1, \dots, w_n$  – синаптические веса нейрона;  $b$  – пороговое значение (порог);  $y$  – выходной сигнал нейрона;  $\varphi(v)$  – функция активации. На рис. 7.2 схематически представлена модель искусственного нейрона.

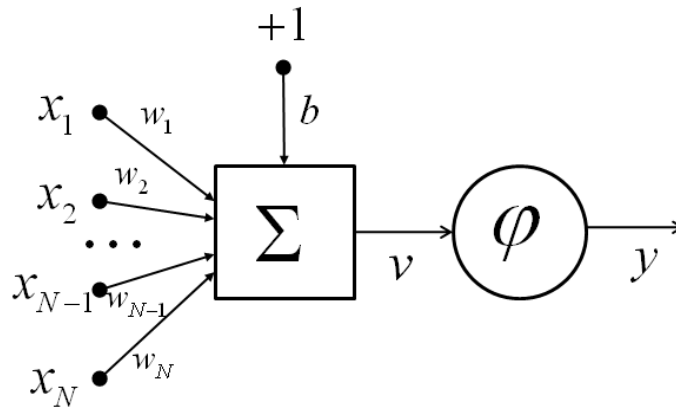


Рис. 7.2. Искусственный нейрон

Математически эту модель можно записать в виде:

$$y = \varphi\left(\sum_{i=1}^N w_i x_i + b\right) \quad (7.1)$$

Величину  $v = \sum w_i x_i + b$ , получаемую на выходе сумматора, называют индуцированным локальным полем нейрона.

На начальной стадии моделирования нейронных сетей применялись пороговые функции активации, например:

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad (\text{рис. 7.3 a}) \quad (7.2)$$

или

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ -1, & v < 0 \end{cases} \quad (\text{рис. 7.3 b}) \quad (7.3)$$

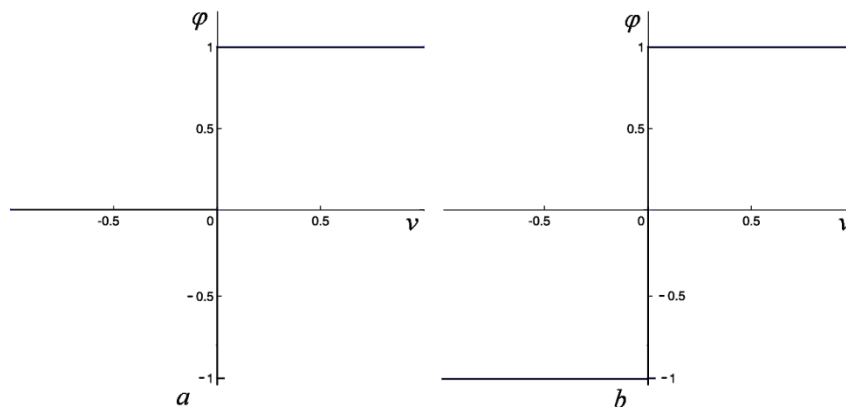


Рис. 7.3. Ступенчатые функции активации

В настоящее время чаще всего используется сигмоидальная функция активации, которая определяется как (см. рис. 7.4 а)

$$\varphi(v) = \frac{1}{1 + e^{-\alpha v}} \quad (7.4)$$

Следует отметить, что при  $\alpha \rightarrow \infty$  сигмоидальная функция активации стремится к

пороговой функции.

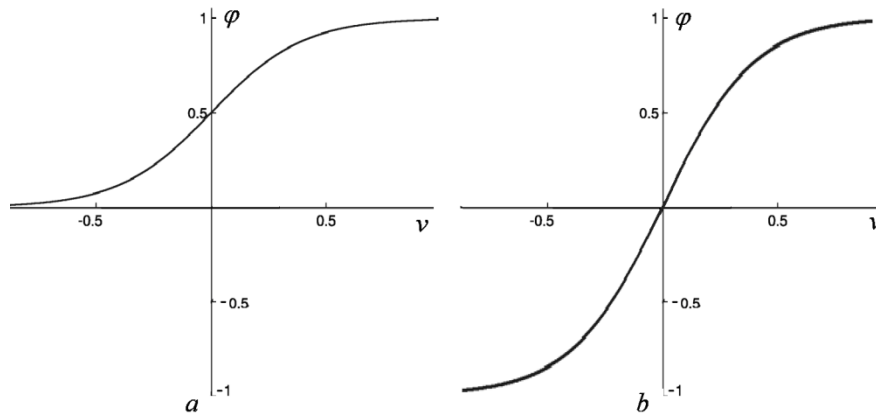


Рис. 7.4. Функции активации

В качестве альтернативы сигмоидальной функции активации иногда применяют функцию гиперболического тангенса (рис. 7.4 б):

$$\varphi(v) = th\left(\frac{\alpha v}{2}\right) = \frac{1 - e^{-\alpha v}}{1 + e^{-\alpha v}} . \quad (7.5)$$

### 7.2.2. Персептрон Розенблатта (модель МакКаллока-Питтса)

Персептрон Розенблатта представляет собой искусственный нейрон со ступенчатой функцией активации, который стал отправной точкой для построения первых искусственных нейронных сетей.

Задача, решаемая с помощью персептрона Розенблатта, состоит в классификации вектора  $x = [x_1 \dots x_N]^T$  в смысле отнесения его к одному из двух классов  $C_0$  и  $C_1$ . Вектор  $u$  относится к классу  $C_0$ , если выходной сигнал  $y$  принимает значение 0, и к классу  $C_1$ , если выходной сигнал равен 1. При этом персептрон разделяет  $N$ -мерное пространство входных сигналов  $u$  на два полупространства, разделяемых  $(N-1)$ -мерной гиперплоскостью, описываемой уравнением:

$$\sum_{i=1}^N w_i x_i + b = 0 . \quad (7.6)$$

Эта гиперплоскость называется решающей границей.

В случае, когда размерность пространства входных сигналов  $N = 2$ , решающая граница – это прямая, описываемая уравнением (рис 7.5):

$$w_1 x_1 + w_2 x_2 + b = 0 \quad (7.7)$$

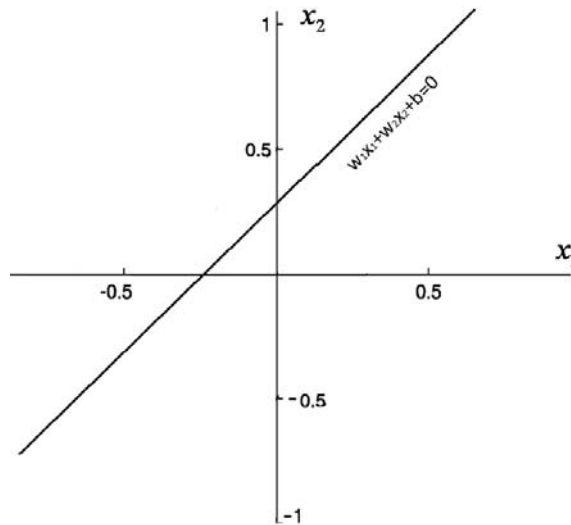


Рис. 7.5. Разделяющая прямая

Следует отметить, что персептрон Розенблатта можно использовать для решения весьма ограниченного класса задач, так как он может классифицировать только линейно разделимые входные сигналы.

### 7.2.3. Нейрон с сигмоидальной функцией активации

Рассмотрим нейрон с двумя входами (синапсами) и сигмоидальной функцией активации

$$\varphi(v) = \frac{1}{1 + e^{-\alpha v}} \quad (7.8)$$

Выходной сигнал такого нейрона описывается выражением:

$$y = \frac{1}{1 + e^{-\alpha(w_1 x_1 + w_2 x_2 + b)}} \quad (7.9)$$

На рис 7.6 показана зависимости выходного сигнала  $y$  от входного сигнала  $[x_1, x_2]$  для нейрона с сигмоидальной функцией активации  $y$ .

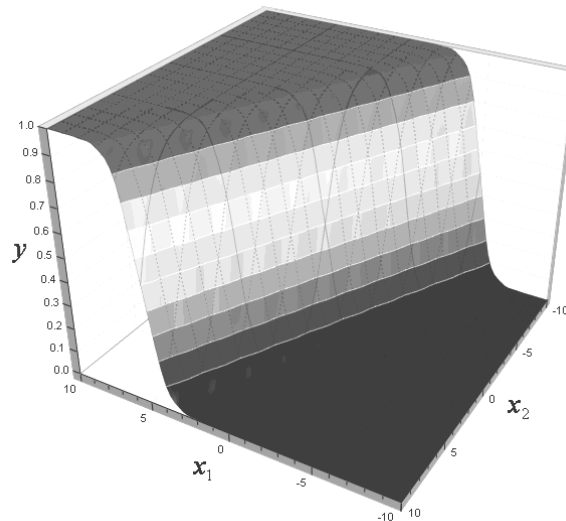


Рис. 7.6. Выходной сигнал нейрона с сигмоидальной функцией активации

Как отмечалось ранее, при  $\alpha \rightarrow \infty$ , сигмоидальная функция стремится к пороговой и формируется решающая граница, которая в данном случае будет прямой, описываемой уравнением  $w_1x_1 + w_2x_2 + b = 0$ .

#### **7.2.4. Стохастическая модель нейрона**

Рассмотренная ранее модель нейрона является детерминированной. Это значит, что преобразование входного сигнала в выходной задается некоторой однозначной функцией, определенной на всем множестве входных сигналов. Однако, в некоторых приложениях применяется стохастическая модель нейрона, в которой функция активации носит вероятностный характер.

В таких моделях выходной сигнал нейрона может быть +1 и -1 и определяется с учетом вероятности каждого из исходов. Таким образом, функция активации для стохастического нейрона будет иметь вид:

$$\varphi(v) = \begin{cases} +1, & \text{с вероятностью } P(v) \\ -1, & \text{с вероятностью } 1 - P(v) \end{cases}, \quad (7.10)$$

где  $P(v)$  – это вероятность активации нейрона, а  $v$  – это индуцированное локальное поле нейрона (сигнал, который формируется на выходе сумматора).

Вероятность активации нейрона  $P(v)$  может быть описана сигмоидальной функцией следующего вида:

$$P(v) = \frac{1}{1 + e^{-v/T}}, \quad (7.11)$$

где  $T$  – это аналог температуры, используемый для управления степенью неопределенности переключения. Следует отметить, что величина  $T$  не описывает физическую температуру нейронной сети.

Очевидно, что при  $T \rightarrow 0$  стохастический нейрон принимает детерминированную форму нейрона со ступенчатой функцией активации вида:

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ -1, & v < 0 \end{cases}. \quad (7.12)$$

### **7.3. Архитектура сетей**

#### **7.3.1. Однослойные сети прямого распространения**

Однослойная сеть прямого распространения представляет собой частный случай многослойной сети. Такая сеть состоит из узлов источника информации и одного слоя нейронов, который является одновременно входным слоем и выходным (рис. 7.7).

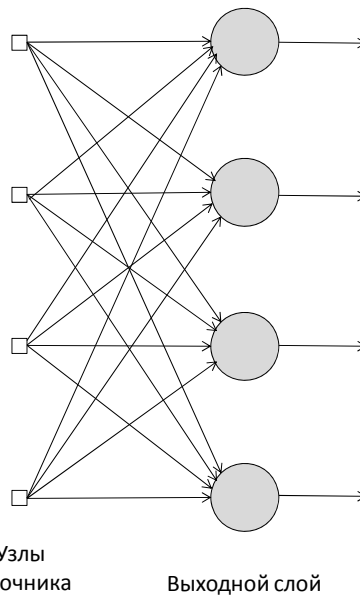


Рис. 7.7. Однослойная сеть прямого распространения

Такие сети называют сетями прямого распространения или ациклическими сетями.

### 7.3.2. Многослойные сети прямого распространения

Многослойные сети прямого распространения характеризуются наличием одного или нескольких скрытых слоев (рис. 7.8).

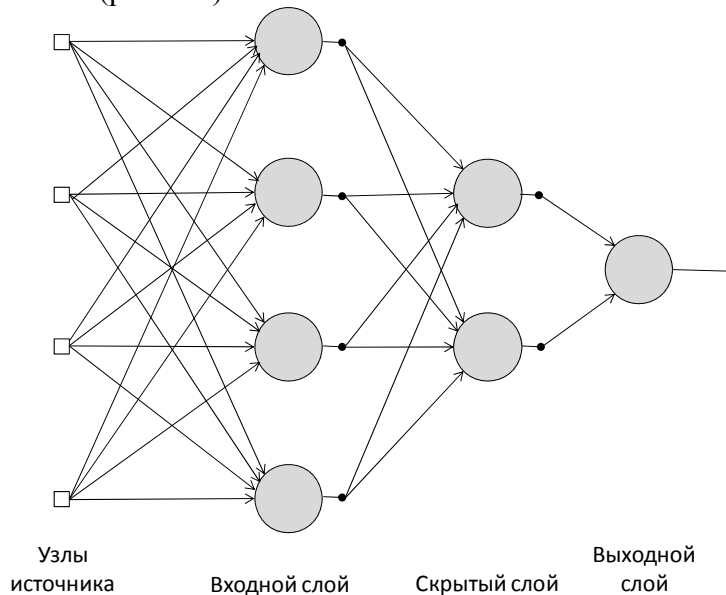


Рис. 7.8. Многослойная сеть прямого распространения

Узлы скрытых слоев называются скрытыми нейронами. Добавляя один или несколько скрытых слоев можно выделить глобальные закономерности в данных, что особенно существенно при большой размерности входных данных.

В многослойных сетях прямого распространения сигнал распространяется послойно: узлы источника формируют входной сигнал для первого (входного) слоя нейронов, выходной сигнал первого слоя нейронов используется в качестве входного сигнала для второго (скрытого) слоя нейронов и т. д. вплоть до последнего (выходного) слоя нейронов.

Сеть, в которой каждый нейрон скрытого слоя соединяется со всеми нейронами соседних слоев, называется полносвязной. Если некоторые синаптические связи отсутствуют, сеть называется неполносвязной.

### 7.3.3. Рекуррентные сети

Рекуррентные сети отличаются от сетей прямого распространения наличием, по крайней мере, одной обратной связи. На рис 7.9 показана архитектура простейшей рекуррентной сети.

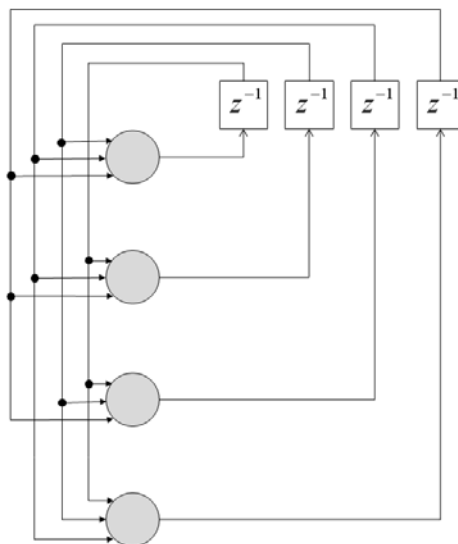


Рис. 7.9. Рекуррентная сеть

Здесь  $z^{-1}$  – это оператор единичной задержки. В общем случае оператор задержки определяется следующим образом:

$$z^{-l}(y_n) = y_{n-l}, \quad (7.13)$$

где  $y_n$  и  $y_{n-l}$  – значения  $y$  на  $n$  и  $n-l$  шаге соответственно.

## 7.4. Представление знаний

Под знаниями понимается хранимая информация и модели, используемые интерпретации, предсказания и реакции на внешние события.

Различают два вида знаний:

- априорная информация, состоящая из достоверных фактов об окружающей среде;
- наблюдения за окружающим миром.

Успех построения реального нейросетевого решения в значительной мере зависит от правильности представления знаний. Вопрос представления знаний в нейронных сетях является очень сложным, однако можно выделить несколько правил.

**Правило 1.** Сходные входные сигналы должны классифицироваться как относящиеся к одному классу. В качестве меры схожести двух входных сигналов  $x_i = [x_{i1} \dots x_{im}]^T$  и  $x_j = [x_{j1} \dots x_{jm}]^T$ , представляющих некоторые точки в пространстве  $\mathfrak{R}^m$  можно использовать эвклидово расстояние между ними. Эвклидово расстояние между парой векторов определяется как:

$$d(x_i, x_j) = \|x_i - x_j\| = \left[ \sum_{k=1}^m (x_{ik} - x_{jk})^2 \right]^{1/2}. \quad (7.14)$$



**Правило 2.** это правило полностью противоположно первому, оно гласит, что элементы, отнесенные к различным классам, должны иметь максимально различные представления.

**Правило 3.** Если некоторое свойство имеет большое значение, то для его представления следует использовать большее количество нейронов. Однако это правило следует применять с осторожностью, так как чрезмерное увеличение количества нейронов может привести к тому, что сеть запомнит некоторый набор примеров из множества доступных измерений, но не сможет выделить закономерности в данных.

**Правило 4.** В структуру сети должна быть встроена априорная информация об окружающей среде. Это обеспечивает специализацию сети и существенно ускоряет процесс обучения.

## 7.5. Процессы обучения

Самым важным свойством нейронных сетей является их способность обучаться на основании данных об окружающей среде и в результате повышать свою производительность.

**Обучение** – это итеративный процесс подстройки свободных параметров сети (синаптических весов и порогов) посредством моделирования окружающей среды. Тип обучения определяется способом подстройки этих параметров.

Процесс обучения предполагает следующую последовательность действий:

- на вход нейронной сети поступает стимул из внешней среды;
- в результате этого в соответствии с некоторым правилом изменяются свободные параметры сети;
- после изменения параметров сеть отвечает на входящие сигналы уже иным образом;
- процесс повторяется до тех пор, пока сеть не сможет решить поставленную задачу.

Эта последовательность действий называется **алгоритмом обучения**. Очевидно, что не существует универсального алгоритма обучения, подходящего для любой задачи и архитектуры сети. Существует лишь набор методов, каждый из которых имеет свои преимущества.

Алгоритмы обучения отличаются друг от друга по способу подстройки синаптических весов и по способу связи обучаемой сети с внешним миром (в этом смысле говорят о парадигме обучения).

### 7.5.1. Парадигмы обучения

#### Обучение с учителем

Участие учителя можно рассматривать, как наличие данных об окружающей среде, представленных в виде пар  $(x_i, d_i)$ , где  $x_i$  – это некоторый входной сигнал, а  $d_i$  – желаемый отклик сети на сигнал  $x_i$ . Схема обучения с учителем показана на рис. 7.10:

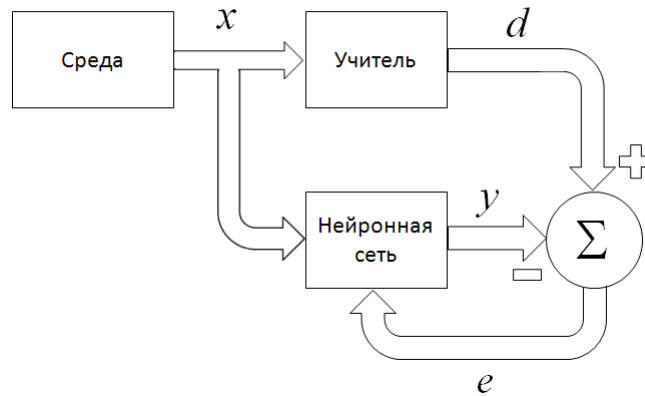


Рис. 7.10. Обучение с учителем

Параметры сети при обучении с учителем корректируются на основании сигнала ошибки  $e_i = d_i - y_i$ , который определяется как разность между желаемым и действительным откликом сети.

Эффективность обучаемой нейронной сети можно оценивать в терминах среднеквадратической ошибки на обучающей выборке (некотором наборе пар  $(x_i, d_i)$ ). Для среднеквадратической ошибки можно построить многомерную поверхность ошибки в координатах свободных параметров сети. Фактически процесс обучения сводится к поиску минимума на поверхности ошибки.

### Обучение без учителя

Обучение без учителя иногда называют обучением на основе самоорганизации. Процесс обучения осуществляется без участия внешнего учителя, существует лишь независимая от задачи мера качества представления. Нейронная сеть, обученная на основании статистических закономерностей во входных данных, способна формировать внутреннее представление признаков и самостоятельно формировать классы входных сигналов.

### Обучение, основанное на коррекции ошибок

Обучение, основанное на коррекции ошибок, является классическим примером обучения с учителем.

Рассмотрим один нейрон, который генерирует некоторый выходной сигнал  $y$  под управлением вектора входного сигнала  $x$ . Сравнивая полученный выходной сигнал  $y$  с ожидаемым сигналом  $d$ , получаем сигнал ошибки:

$$e = d - y. \quad (7.15)$$

На основании полученного сигнала ошибки осуществляется изменение синаптических весов нейрона с целью уменьшения функции стоимости  $E$ , которая определяется как:

$$E = \frac{1}{2} e^2. \quad (7.16)$$

Пошаговая корректировка синаптических весов нейрона продолжается до тех пор, пока нейрон не достигнет устойчивого состояния, то есть такого, при котором веса практически не меняются. Минимизация функции стоимости  $E$  осуществляется по так называемому дельта правилу или правилу Видроу-Хоффа.

Обозначим  $w_j(n)$  значение синаптического веса  $w_j$  на шаге  $n$ . Тогда, в соответствии с дельта правилом, изменение веса  $w_j$  будет задаваться выражением:

$$\Delta w_j = \eta e x_j, \quad (7.17)$$

где  $\eta$  – константа скорости обучения (обычно выполняется неравенство  $0 < \eta \leq 1$ ), а  $x_j$  – компонент входного вектора  $x$ , соответствующий синаптическому весу  $w_j$ . Таким образом, значение синаптического веса  $w_j$  после применения дельта правила будет определяться выражением:

$$w_j(n+1) = w_j(n) + \eta e x_j. \quad (7.18)$$

Дельта правило может быть сформулировано следующим образом: **корректировка, применяемая к синаптическому весу нейрона, пропорциональна произведению сигнала ошибки на входной сигнал, его вызвавший.**

Сформулированное дельта правило предполагает возможность прямого измерения сигнала ошибки, то есть наличие эталонного сигнала  $d$ .

Следует отметить, что константа скорости обучения  $\eta$  определяет эффективность и точность процесса обучения. Для обеспечения сходимости процесса обучения необходим тщательный подбор этого параметра.

## 7.5.2. Практики обучения

### Обучение Хебба

Алгоритм обучения Хебба назван так в честь нейрофизиолога, который в нейробиологическом контексте предложил следующее правило модификации синаптических весов:

*Если аксон клетки A находится на достаточно близком расстоянии от клетки B и постоянно или периодически участвует в ее возбуждении, наблюдается процесс метаболических изменений в одном или обоих нейронах, выражающийся в том, что эффективность нейрона A как одного из возбуждателей нейрона B возрастает.*

В контексте искусственных нейронных сетей данное правило может быть переформулировано в виде двух утверждений:

– если два нейрона, связанные синаптической связью возбуждаются одновременно (синхронно), то прочность этой связи (соответствующий синаптический вес) возрастает;

– если два нейрона по обе стороны синапса возбуждаются асинхронно, такой синапс ослабевает (синаптический вес уменьшается).

Можно сформулировать следующие свойства синапса Хебба:

– **Зависимость от времени.** Изменение синаптического веса зависит от точного времени возникновения пресинаптического и постсинаптического сигналов.

– **Локальность.** На изменение синаптического веса оказывают воздействие сигналы, находящиеся в пространственно-временной близости.

– **Интерактивность.** Изменение синаптического веса определяется сигналами на обоих его концах.

– **Корреляция.** Механизм изменения синаптического веса определяется наличием корреляции между пресинаптическим и постсинаптическим сигналом.

### Математические модели обучения Хебба

Рассмотрим синапс с синаптическим весом  $w_{kj}$ , который соединяет выходной сигнал  $y_j$  некоторого нейрона  $j$  с соответствующим входом нейрона  $k$ . Выходной сигнал  $y_j$  является

$j$ -ой компонентой вектора входного сигнала для нейрона  $k$ .

Таким образом, модификация веса  $w_{kj}$  нейрона  $k$ , будет определяться парой предсинаптического и постсинаптического сигналов  $(x_j, y_k)$ . Изменение синаптического веса  $w_{kj}$  в момент времени  $n$  будет определяться соотношением:

$$\Delta w_{kj}(n) = F(y_k(n), x_j(n)). \quad (7.19)$$

Это соотношение может быть записано в различных видах, каждый из которых все равно остается математической моделью алгоритма обучения Хебба.

**Гипотеза Хебба.** Является простейшей формой обучения Хебба. В соответствии с гипотезой Хебба выражение для модификации веса  $w_{kj}$  в момент времени  $n$  записывается в виде:

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n), \quad (7.20)$$

где  $\eta$  – константа скорости обучения.

Недостатком данной модели является то, что при частом приложении входного сигнала  $x_j$  наблюдается неограниченный рост синаптического веса  $w_{kj}$ , что приводит к потере избирательности связей и неспособности сети сохранить какую-либо информацию.

**Гипотеза ковариации.** Недостаток гипотезы Хебба может быть устранен, если переписать выражение для модификации синаптического веса  $w_{kj}$  в виде:

$$\Delta w_{kj}(n) = \eta (y_k(n) - \bar{y})(x_j(n) - \bar{x}), \quad (7.21)$$

где  $\bar{x}$  и  $\bar{y}$  – это усредненные по времени значения предсинаптического и постсинаптического сигналов.

Синапс, вес которого изменяется в соответствии с гипотезой ковариации, обладает следующими свойствами:

- синаптический вес увеличивается при высоком уровне предсинаптического и постсинаптического сигналов, то есть при  $x_j(n) > \bar{x}$  и  $y_k(n) > \bar{y}$ ;
- синаптический вес уменьшается при высоком уровне предсинаптического сигнала, который не вызывает существенной постсинаптической активности, то есть  $x_j(n) > \bar{x}$  и  $y_k(n) < \bar{y}$ ;
- синаптический вес уменьшается, если наблюдается существенная постсинаптическая активность при отсутствии значительного предсинаптического сигнала, то есть  $x_j(n) < \bar{x}$  и  $y_k(n) > \bar{y}$ .

### Конкурентное обучение

При конкурентном обучении, в отличие от, например, обучения Хебба, в фиксированный момент времени в возбужденном состоянии может находиться только один нейрон выходного слоя. В связи с этим конкурентные сети часто применяются для решения задач классификации входных образов.

Конкурентное обучение строится на основании следующих принципов:

- Все нейроны выходного слоя одинаковы и имеют в начальный момент времени случайные значения весов.
- Существует предельное значение выходного сигнала каждого нейрона.
- Существует механизм, определяющий нейрон, победивший в борьбе за право генерировать выходной сигнал. Принцип конкуренции можно сформулировать как «победитель получает все».

Механизм конкуренции может формироваться за счет введения отрицательных обратных связей, обеспечивающих латеральное торможение, когда каждый нейрон стремится затормозить связанные с ним нейроны (рис. 7.11).

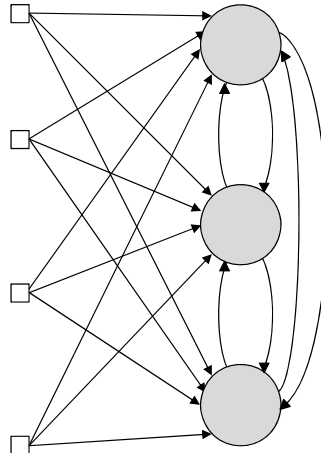


Рис. 7.11. Сеть конкурентного обучения

Для того, чтобы нейрон  $k$  выходного слоя победил в конкурентной борьбе, его индуцированное локальное поле  $v_k = \sum_{j=1}^n x_j w_{kj} + b_k$  должно быть максимальным среди всех нейронов выходного слоя. Таким образом, выходной сигнал нейрона  $k$  можно записать в виде:

$$\begin{cases} 1, & v_k > v_j, \forall j \neq k \\ 0, & \exists j \neq k : v_k < v_j \end{cases} \quad (7.22)$$

Если предположить, что синаптические веса конкурирующих нейронов положительны и выполняются следующие нормировки:

$$\sum_j w_{kj} = 1, \quad (7.23)$$

$$\sum_j x_j = 1. \quad (7.24)$$

Правило конкурирующего обучения можно записать в виде:

$$\begin{cases} \eta(x_j - w_{kj}), & y_k = 1 \\ 0, & y_k = 0 \end{cases} \quad (7.25)$$

Данное правило обеспечивает смещение вектора синаптических весов в сторону вектора входного сигнала.

### Обучение Больцмана

Нейронная сеть, состоящая из стохастических нейронов, в которой каждый нейрон связан обратной связью со всеми остальными, и обучаемая в соответствии с алгоритмом обучения Больцмана получила название машины Больцмана.

Алгоритм обучения Больцмана имеет стохастический характер, что способствует поиску абсолютного минимума целевой функции и выхода системы из состояния локального минимума целевой функции. Способность находить абсолютный минимум целевой функции обеспечила широкое применение машины Больцмана для решения задач классификации образов.

На рис. 7.12 приводится простейший пример архитектуры машины Больцмана.

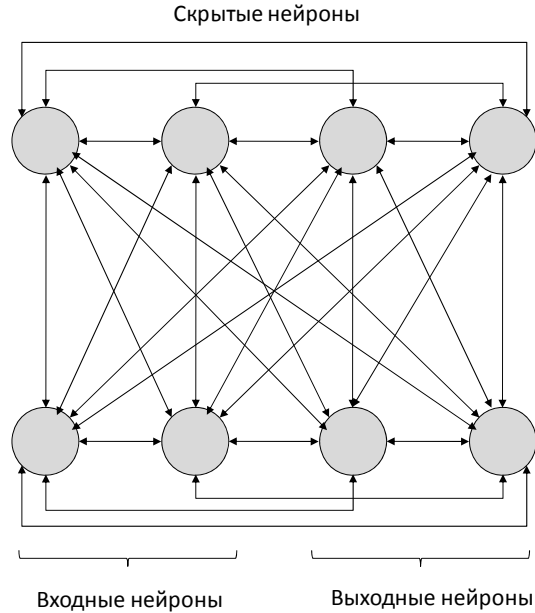


Рис. 7.12. Машина Больцмана

Подразумевается, что в такой сети применяются симметричные синаптические связи, то есть  $w_{ij} = w_{ji}$ .

Можно выделить два режима функционирования машины Больцмана:

- *скованное состояние*, в котором все видимые нейроны находятся в состояниях, определенных внешней средой;
- *свободное состояние*, в котором все нейроны могут свободно функционировать.

В процессе функционирования сети выбирается один из свободных нейронов, и его состояние изменяется с вероятностью:

$$P(y \rightarrow -y) = \frac{1}{1 + e^{-\Delta E/T}}, \quad (7.26)$$

где  $E$  – это энергия машины, определяемая соотношением:

$$E = \frac{1}{2} \sum_j \sum_{k (j \neq k)} w_{kj} x_k x_j, \quad (7.27)$$

а  $\Delta E$  – изменение энергии машины, вызванное переключением состояния выбранного нейрона. Многократное повторение этой процедуры приводит к достижению машиной состояния термального равновесия.

Процедура обучения для такой сети состоит из следующих шагов:

1. Вычислить закрепленные вероятности.
  - а) придать входным и выходным нейронам значения обучающего вектора;
  - б) предоставить сети возможность искать равновесие;
  - в) запомнить выходные значения (состояния) для всех нейронов;
  - г) повторить шаги от а до с для всех обучающих векторов;
  - д) вычислить вероятность  $P_{ij}^+$ , то есть по всему множеству обучающих векторов вычислить вероятность того, что состояния обоих нейронов равны единице.
2. Вычислить незакрепленные вероятности.

- а) предоставить сети возможность «свободного движения» без закрепления входов или выходов, начав со случайного состояния;
- б) повторить предыдущий много раз, запоминая состояния всех нейронов;
- в) вычислить вероятность  $P_{ij}^-$ , то есть вероятность того, что состояния обоих нейронов равны единице.

3. Скорректировать веса сети следующим образом:  $\Delta w_{ij} = \eta(P_{ij}^+ - P_{ij}^-)$ , где  $\Delta w_{ij}$  – изменение веса  $w_{ij}$ , а  $\eta$  – константа скорости обучения.

Обученная в соответствии с алгоритмом Больцмана нейронная сеть обладает способностью к дополнению входных образов. То есть, если входной вектор с неполной информацией поступает на вход сети, сеть дополнит недостающую информацию.

Основным недостатком алгоритма обучения Больцмана является большая длительность процесса обучения.

### **Алгоритм обратного распространения ошибки**

Алгоритм обратного распространения ошибки является классическим алгоритмом обучения многослойного персептрона. Этот алгоритм состоит из двух проходов вычислений:

- прямой проход, при котором вычисляется отклик сети на поданный входной сигнал;
- обратный проход, при котором, в соответствии с полученным сигналом ошибки, модифицируются веса всех нейронов от выходного слоя к входному.

Введем следующие обозначения:

- будем считать, что индекс  $i$  относится к нейрону во входном слое, индекс  $j$  – к нейрону в скрытом слое и индекс  $k$  – к нейрону в выходном слое;
- зададим обучающее множество  $S$  в виде набора примеров  $(x_n, d_n)$ , где  $n$  – это номер обучающего примера;
- обозначим ошибку на выходе нейрона  $k$  для обучающего примера  $n$  как  $e_k(n)$ ;
- введем обозначение  $E$  для половинной суммы квадратов ошибок  $e_k(n)$  на текущем шаге процесса обучения.

При прямом проходе вычислений определяются отклики  $y_k$  для нейронов выходного слоя. Сравнивая полученный отклик сети  $y$  с ожидаемым откликом  $d$ , получаем сигнал ошибки на выходе каждого из нейронов выходного слоя:

$$e_k = d_k - y_k \quad (7.28)$$

Энергию ошибки можно записать в виде:

$$E = \frac{1}{2} \sum_k e_k^2 \quad (7.29)$$

где суммирование осуществляется по всем нейронам выходного слоя, то есть по тем нейронам, для которых можно явно вычислить сигнал ошибки.

Модификация весов при обратном проходе вычислений должна быть направлена на уменьшение величины  $E$ , которая является функцией от всех свободных параметров сети.

Учитывая, что  $y_k = \varphi(v_k) = \varphi(\sum_{j=0}^m w_{kj} y_j)$ , производную сигнала ошибки по некоторому синаптическому весу в сети можно записать как:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial e_j} \frac{\partial e_j}{\partial y_j} \frac{\partial y_j}{\partial v_j} \frac{\partial v_j}{\partial w_{ji}}. \quad (7.30)$$

Очевидно, что выполняются следующие равенства:

$$\frac{\partial E}{\partial e_j} = \frac{1}{2} \frac{\partial \sum_j e_j^2}{\partial e_j} = e_j, \quad (7.31)$$

$$\frac{\partial e_j}{\partial y_j} = -1, \quad (7.32)$$

$$\frac{\partial y_j}{\partial v_j} = \phi'(v_j), \quad (7.33)$$

и наконец,

$$\frac{\partial v_j}{\partial w_{ji}} = y_i. \quad (7.34)$$

Таким образом, производную  $\frac{\partial E}{\partial w_{ji}}$  можно переписать в виде:

$$\frac{\partial E}{\partial w_{ji}} = -e_j \phi'(v_j) y_i. \quad (7.35)$$

Воспользовавшись методом градиентного спуска, запишем величину модификации синаптического веса  $w_{ji}$  в виде:

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}}, \quad (7.36)$$

где  $\eta$  – константа скорости обучения.

Величину модификации синаптического веса  $w_{ji}$  можно переписать в виде:

$$\Delta w_{ji} = \eta \delta_j y_i, \quad (7.37)$$

где  $\delta_j = -\frac{\partial E}{\partial v_j} = -\frac{\partial E}{\partial e_j} \frac{\partial e_j}{\partial y_j} \frac{\partial y_j}{\partial v_j} = e_j \phi'(v_j)$  – локальный градиент нейрона  $j$ .

Таким образом, для вычисления модификации синаптического веса  $w_{ji}$  достаточно вычислить локальный градиент  $\delta_j = e_j \phi'(v_j)$  нейрона  $j$ .

В случае, когда рассматриваемый нейрон находится в выходном слое, вычисление локального градиента не представляет труда, поскольку для нейронов выходного слоя можно явно определить сигнал ошибки  $e_j$ .

Рассмотрим локальный градиент нейрона  $j$ , находящегося в скрытом слое сети.

$$\delta_j = -\frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial v_j} = -\frac{\partial E}{\partial y_j} \phi'(v_j). \quad (7.38)$$

Поскольку  $E = \frac{1}{2} \sum_k e_k^2$ , производную  $\frac{\partial E}{\partial y_j}$  можно записать в виде:

$$\frac{\partial E}{\partial y_j} = \sum_k e_k \frac{\partial e_k}{\partial y_j} = \sum_k e_k \frac{e_k}{v_k} \frac{v_k}{y_j}. \quad (7.39)$$

Используя выражения:



$$e_k = d_k - y_k = d_k - \varphi_k(v_k) \quad (7.40)$$

а также

$$v_k = \sum_{j=0}^m w_{kj} y_j, \quad (7.41)$$

Можно получить следующие очевидные соотношения:

$$\frac{\partial e_k}{\partial v_k} = -\varphi'_k(v_k), \quad (7.42)$$

и, кроме того,

$$\frac{\partial v_k}{\partial y_j} = w_{kj}. \quad (7.43)$$

И переписать выражение для производной  $\frac{\partial E}{\partial y_j}$  в виде:

$$\frac{\partial E}{\partial y_j} - \sum_k e_k \varphi'_k(v_k) w_{kj} = -\sum_k \delta_k w_{kj}. \quad (7.44)$$

А следовательно, мы получаем выражение для локального градиента скрытого нейрона:

$$\delta_j = \varphi'_j(v_j) \sum_k \delta_k w_{kj}. \quad (7.45)$$

Подводя итоги, правило модификации весов для алгоритма обратного распространения ошибки можно сформулировать в виде:

$$\Delta w_{ji} = \eta \delta_j y_i, \quad (7.46)$$

где локальный градиент для нейронов выходного слоя описывается выражением:

$$\delta_j = e_j \varphi'_j(v_j). \quad (7.47)$$

А для нейрона в скрытом слое локальный градиент определяется как:

$$\delta_j = \varphi'_j(v_j) \sum_k \delta_k w_{kj}. \quad (7.48)$$

Модификация весов выполняется после вычисления отклика сети на каждый из обучающих примеров. Сначала вычисляются локальные градиенты для всех нейронов, начиная с выходного слоя, а потом, в соответствии с полученными градиентами, вычисляются модификации всех весов сети.

Процесс обучения прекращается, если ошибка  $E$  для всех обучающих примеров не превышает некоторого приемлемого значения, или если превышено максимально допустимое количество итераций.

### Вопросы для самоконтроля

1. Нейрон получает входной сигнал от четырех других нейронов, уровни возбуждения которых равны 10; -20; 4; -2. Соответствующие синаптические веса равны 0.8; 0.2; -1.0; 0.9. Вычислите выходной сигнал такого нейрона, если нейрон представлен моделью МакКаллока-Питтса.
2. Решите предыдущую задачу, если нейрон имеет сигмоидальную функцию активации.
3. Постройте схему полностью связанной сети прямого распространения, в которой входной и скрытый слои содержат по 3 нейрона, а выходной слой содержит 2 нейрона.
4. Постройте схему рекуррентной сети, состоящей из трех нейронов, не имеющих обратных связей с самими собой.

5. Что такое обучение нейронной сети?
6. Какие алгоритмы обучения можно отнести к парадигме обучения с учителем?
7. Какие алгоритмы обучения можно отнести к парадигме обучения без учителя?
8. Поясните, что такое решающая граница.
9. Сформулируйте дельта правило, применяемое при обучении, основанном на коррекции ошибок.
10. Чем отличается гипотеза Хебба от гипотезы ковариации?
11. Сформулируйте принципы, лежащие в основе конкурентного обучения.
12. Назовите основные достоинства и недостатки статистических методов обучения на примере обучения Больцмана.
13. Сформулируйте правила изменения синаптических весов в соответствии с алгоритмом обратного распространения ошибки.
14. Каким свойством должна обладать функция активации нейронов сети для того, чтобы можно было применить алгоритм обратного распространения ошибки?

## **Литература к разделу 7**

1. Саймон Хайкин, «Нейронные сети полный курс», 2-е изд., испр. : пер. с англ. – М.:ООО «И.Д. Вильямс», 2006. 1104 с.
2. Рутковская Данута, Пилиньский Мачей, Рутковский Лешек «Нейронные сети, генетические алгоритмы и нечеткие системы»: пер. с польск. И.Д. Рудинского. – М. : Горячая линия – Телеком, 2007. – 452 с.
3. Каширина И.Л. Искусственные нейронные сети: Учебное пособие. – Воронеж: Изд-во ВГУ, 2005. – 51 с.
4. Уоссермен Ф. Нейрокомпьютерная техника. – М. : Мир, 1992 г. – 240с.