# ПОНЯТИЕ О МАТЕМАТИЧЕСКОМ НЕЙРОНЕ

## 1. Биологический прототип

Развитие искусственных нейронных сетей вдохновляется биологией. То есть, рассматривая сетевые конфигурации и алгоритмы, исследователи применяют термины, заимствованные из принципов организации мозговой деятельности. Но на этом аналогия заканчивается. Наши знания о работе мозга столь ограничены, что мало бы нашлось точно доказанных закономерностей для тех, кто пожелал бы руководствоваться ими. Поэтому разработчикам сетей приходится выходить за пределы современных биологических знаний в поисках структур, способных выполнять полезные функции. Во многих случаях это приводит к необходимости отказа от биологического правдоподобия, мозг становится просто метафорой, и создаются сети, невозможные в живой материи или требующие неправдоподобно больших допущений об анатомии и функционировании мозга.

Несмотря на то, что связь с биологией слаба и зачастую несущественна, искусственные нейронные сети продолжают сравнивать с мозгом. Их функционирование часто имеет внешнее сходство с человеческим познанием, поэтому трудно избежать этой аналогии. К сожалению, такие сравнения неплодотворны и создают неоправданные ожидания, неизбежно ведущие к разочарованию.

Нервная система человека, построенная из элементов, называемых нейронами, имеет ошеломляющую сложность. Около $10^{11}$ нейронов участвуют в примерно $10^{15}$ передающих связях, имеющих длину метр и более. Каждый нейрон обладает многими свойствами, общими с другими органами тела, но ему присущи абсолютно уникальные способности: принимать, обрабатывать и передавать электрохимические сигналы по нервным путям, которые образуют коммуникационную систему мозга.

***Нейро́н (от др.-греч. – волокно, нерв) — структурно-функциональная единица нервной системы, представляющая собой электрически возбудимую***

***клетку, которая обрабатывает, хранит и передает информацию посредством электрических и химических сигналов.***


***Аксон – это нейрит (длинный цилиндрический отросток нервной клетки), по которому нервные импульсы идут от тела клетки (сомы) к другим нервным клеткам.***

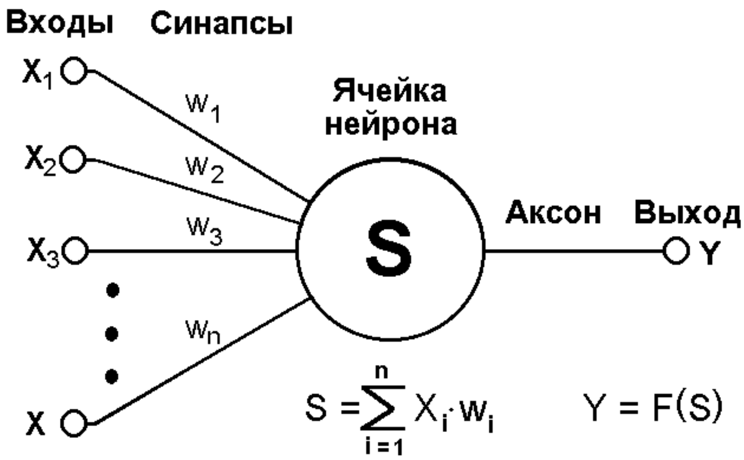
***Дендрит – разветвлённый отросток нейрона, который получает информацию через химические (или электрические) синапсы от аксонов других нейронов и передаёт её через электрический сигнал телу нейрона, из которого вырастает.***

***Си́напс (греч. — соединение, связь) — место контакта между двумя нейронами***.

Дендриты идут от тела нервной клетки к другим нейронам, где они принимают сигналы в точках соединения, называемых синапсами. Принятые синапсом входные сигналы передаются к телу нейрона. Здесь они суммируются, причем одни входы стремятся возбудить нейрон, другие — воспрепятствовать его возбуждению.

Когда суммарное возбуждение в теле нейрона превышает некоторый порог, нейрон возбуждается, посылая по аксону сигнал другим нейронам. У этой основной функциональной схемы много усложнений и исключений, тем не менее, большинство искусственных нейронных сетей моделируют лишь эти простые свойства.

Проведя аналогию с биологическим можно построить искусственный нейрон.




Наиболее часто нейронные сети используются для решения следующих задач:

***Классификация образов*** – указание принадлежности входного образа, представленного вектором признаков, одному или нескольким предварительно определенным классам.

***Кластеризация*** – классификация образов при отсутствии обучающей выборки с метками классов.

***Прогнозирование*** – предсказание значения $y(t_{n+1})$ при известной последовательности $y(t_1),y(t_2),\ldots,y(t_n)$.

***Оптимизация*** – нахождение решения, удовлетворяющего системе ограничений и максимизирующим или минимизирующим целевую функцию.

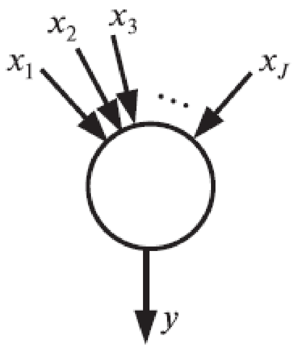
***Память, адресуемая по содержанию (ассоциативная память)*** –память, доступная при указании заданного содержания.

***Управление*** – расчет такого входного воздействия на систему, при котором она следует по желаемой траектории.

## 2. Математический нейрон Мак-Каллока – Питтса. Простой персептрон.

Первой работой, которая заложила теоретический фундамент для создания интеллектуальных устройств, моделирующих человеческий мозг на самом низшем—структурном—уровне, принято считать опубликованную в 1943 г. статью Уоррена Мак-Каллока и Вальтера Питтса «*Идеи логических вычислений в нервной деятельности*». Они предложили математическую модель нейрона мозга человека, назвав ее ***математическим*** или ***модельным нейроном***.

Авторы математического нейрона предложили изображать нейрон в виде кружочка со стрелочками.




Стрелки означают входы и выход нейрона. Через входы математический нейрон принимает входные сигналы $\mathbf{x_1, x_2, \ldots, x_j, \ldots, x_J}$ и суммирует их, умножая каждый входной сигнал на некоторый весовой коэффициент $\mathbf{w_j}$:

$$S = \sum_{j=1}^{J} w_j x_j$$

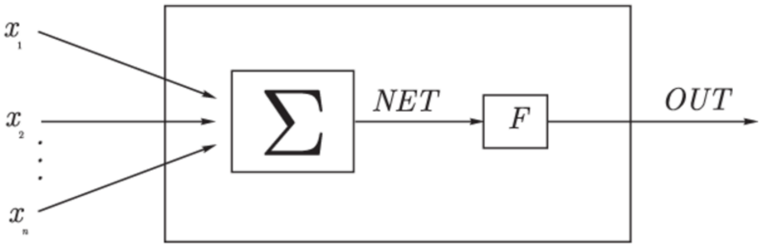
После выполнения операции суммирования математический нейрон формирует выходной сигнал ***y*** согласно следующему правилу:

$$y = \begin{cases} 1, & если \quad S \geq \theta \\ 0, & если \quad S < \theta \end{cases}$$

где θ – порог чувствительности нейрона.

Таким образом, математический нейрон может существовать в двух состояниях. Если взвешенная сумма входных сигналов ***S*** меньше порога ***θ***, то его выходной сигнал ***y*** равен нулю. В этом случае говорят, что нейрон не возбуждён. Если же входные сигналы достаточно интенсивны и их взвешенная сумма достигает порога чувствительности ***θ***, то нейрон переходит в возбуждённое состояние, и на его выходе, согласно правилу, образуется сигнал ***y=1***.

Иногда встречается следующее схематичное изображение, которое более точно отображает процессы происходящие в математическом нейроне.




Тогда входные сигналы ($x_1, x_2, \ldots, x_n$), можно в совокупности обозначить вектором $X$, а веса ($w_1, w_2, \ldots, w_n$) совокупностью $\mathbf{W}$. Выход из суммирующего блока обозначается ***NET***.

$$NET = S = X\mathbf{W}$$

Сигнал ***NET*** далее преобразуется активационной функцией ***F*** и дает выходной нейронный сигнал ***OUT***. Активационная функция может быть обычной линейной функцией

$$OUT = F(NET)$$

где ***F*** – константа, пороговой функцией

$$OUT = \begin{cases} 1, & \text{если } NET > T; \\ 0, & \text{если } NET \leqslant T \end{cases}$$

где ***T*** – некоторая постоянная пороговая величина, или же функция, более точно моделирующая нелинейную передаточную характеристику биологического нейрона и предоставляющей нейронной сети большие возможности.

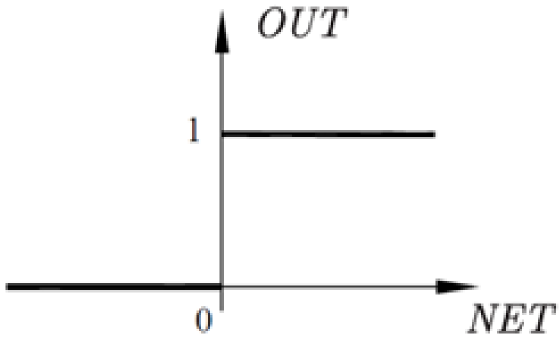
**3. Активационная (пороговая) функция**

***Активационная функция нейрона*** определяет нелинейное преобразование, осуществляемое нейроном и вычисляющая выходной сигнал формального нейрона. Выбор активационной функции определяется спецификой поставленной задачи либо ограничениями, накладываемыми некоторыми алгоритмами обучения.

Существует множество видов активационных функций, но более всего распространены следующие три:

**1. ПОРОГОВАЯ ФУНКЦИЯ.**

Другое название — ***функция Хевисайда***. Представляет собой перепад. До тех пор пока взвешенный сигнал на входе нейрона не достигает некоторого уровня $\boldsymbol{T}$ – сигнал на выходе равен нулю. Как только сигнал $\boldsymbol{NET}$ на входе $\boldsymbol{F}$ превышает указанный уровень выходной сигнал $\boldsymbol{OUT}$ скачкообразно изменяется на единицу.




Математическую запись функции я вам уже приводил:

$$OUT = \begin{cases} 1, \text{если } NET \geq T \\ 0, \text{если } NET < \mathrm{T} \end{cases}$$

Таким образом, математический нейрон представляет собой пороговый элемент с несколькими входами и одним выходом. Каждый математический нейрон имеет свое определенное значение порога чувствительности $\boldsymbol{T}$.

***Пример расчета с порогом.***

Авторы математического нейрона в своей статье  также показали, что с помощью математического нейрона с помощью пороговой функции можно моделировать различные логические функции, например функцию логического умножения ***«И» («AND» конъюнкция),*** функцию логического сложения ***«ИЛИ» («OR»)*** и функцию логического отрицания ***«НЕТ» («NOT»)***.

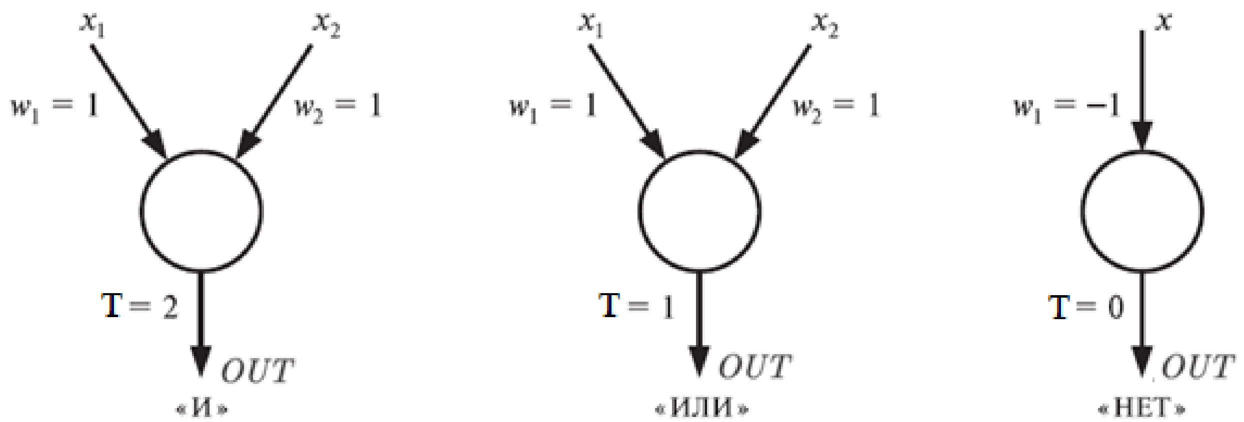
*Таблицы истинности логических функций.*

| **И (⋀)** | | |
|---|---|---|
| $x_1$ | $x_2$ | $x_1 \wedge x_2$ |
| **0** | **0** | **0** |
| **0** | **1** | **0** |
| **1** | **0** | **0** |
| **1** | **1** | **1** |

| **ИЛИ (⋁)** | | |
|---|---|---|
| $x_1$ | $x_2$ | $x_1 \vee x_2$ |
| **0** | **0** | **0** |
| **0** | **1** | **1** |
| **1** | **0** | **1** |
| **1** | **1** | **1** |

| **НЕТ(¬)** | |
|---|---|
| $x$ | $\neg x$ |
| **0** | **1** |
| **1** | **0** |

Представим математические нейроны




Рассчитываем ***логическое «И»***. Сигнал ***NET*** может быть :

*При* $x_1 = 0,\ x_2 = 0 \rightarrow x_1 * w_1 + x_2 * w_2 = 0*1 + 0*1 = NET_1 = 0$

*При* $x_1 = 0,\ x_2 = 1 \rightarrow x_1 * w_1 + x_2 * w_2 = 0*1 + 1*1 = NET_2 = 1$

*При* $x_1 = 1,\ x_2 = 0 \rightarrow x_1 * w_1 + x_2 * w_2 = 1*1 + 0*1 = NET_3 = 1$

*При* $x_1 = 1,\ x_2 = 1 \rightarrow x_1 * w_1 + x_2 * w_2 = 1*1 + 1*1 = NET_4 = 2$

Устанавливаем пороговую функцию

$$OUT = \begin{cases} 1, \text{если } NET \geq 2 \\ 0, \text{если } NET < 2 \end{cases}$$

Тогда на выходе: $OUT_1 = 0,\ OUT_2 = 0,\ OUT_3 = 0,\ OUT_4 = 1$

Рассчитываем ***логическое «ИЛИ».*** Сигналы ***NET*** может такие же как у *логического «И»*, но пороговая иная.

$$OUT = \begin{cases} 1, \text{если } NET \geq 1 \\ 0, \text{если } NET < 1 \end{cases}$$

Тогда на выходе $OUT_1 = 0,\ OUT_2 = 1,\ OUT_3 = 1,\ OUT_4 = 1$

Рассчитываем ***логическое «НЕТ».*** Тогда сигнал ***NET*** может быть :

*При* $x = 0 \rightarrow x * w = 0*(-1) = NET_1 = 0$

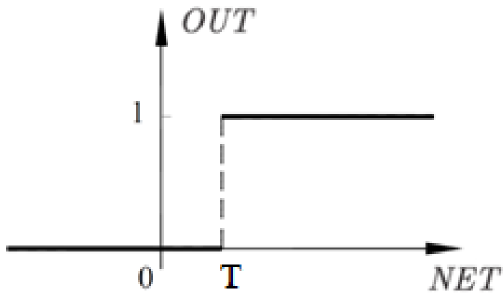
*При* $x = 1 \rightarrow x * w = 1*(-1) = NET_2 = -1$

Устанавливаем пороговую функцию для логического отрицания.

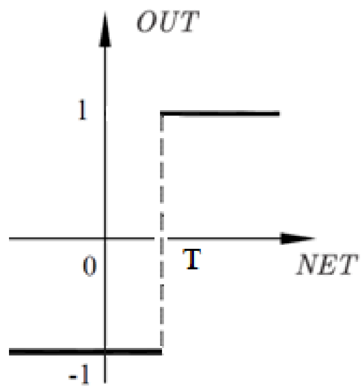
$$OUT = \begin{cases} 1, \text{если } NET \geq 0 \\ 0, \text{если } NET < 0 \end{cases}$$

Тогда на выходе: $OUT_1 = 1,\ OUT_2 = 0$

Таким образом, если рассматривать пороговую функцию, то в ней происходит смещение на величину ***T***.




Пороговая функция может иметь и другую интерпретацию.




Математическую запись функции будет выглядеть:

$$OUT = \begin{cases} 1, \text{если } NET \geq T \\ -1, \text{если } NET < T \end{cases}$$

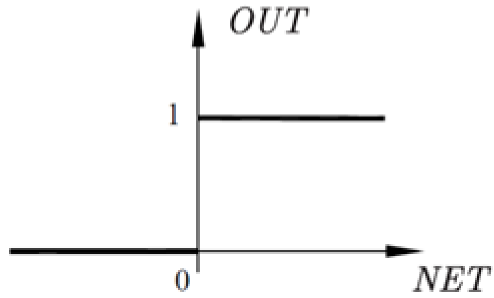
Поэтому для упрощения расчета в современном исполнении нейрона вводят понятия смещения ***b***, которое отличается от ***T*** только знаком: ***b* = - *T***. При этом нейронное смещение ***b*** можно рассматривать как вес $\boldsymbol{w_0}$ некоторого дополнительного входного сигнала $\boldsymbol{x_0 = 1}$, величина которого всегда равна единице.

$$\mathbf{NET = S} = \sum_{j=1}^{J} \boldsymbol{w_j x_j} + \boldsymbol{w_0 \, x_0}$$

Таким образом, дополнительный вход $\boldsymbol{x_0}$ и соответствующий ему вес $\boldsymbol{w_0}$ используются для ***инициализации нейрона***. Под инициализацией подразумевается смещение активационной функции нейрона по горизонтальной оси, то есть формирование порога чувствительности нейрона.

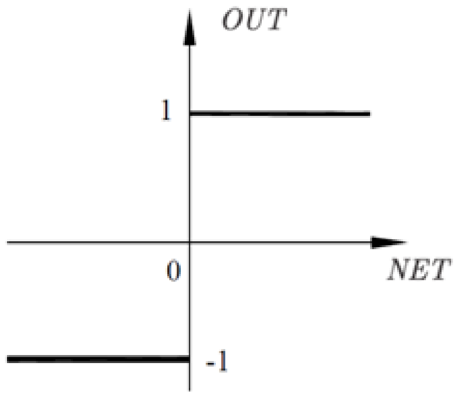
Тогда пороговая активационная функция нейрона примет вид:

$$OUT = \begin{cases} 1, \text{если } NET \geq 0 \\ 0, \text{если } NET < 0 \end{cases}$$




Или более симметричный вид

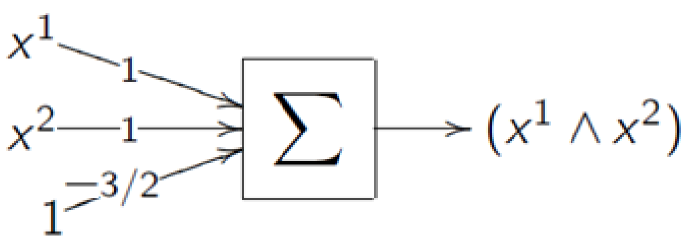
$$OUT = \begin{cases} 1, \text{если } NET \geq 0 \\ -1, \text{если } NET < 0 \end{cases}$$




***Пример расчета со смещением.***

<u>Рассчитываем ***логическое «И»***</u>

Установим $\boldsymbol{x_0 = 1}$, а $\boldsymbol{w_0 = -3/2}$




*При* $x_1=0$, $x_2=0 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 0*1 + 0*1 - 1*3/2 = NET_1 = -1{,}5$

*При* $x_1=0$, $x_2=1 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 0*1 + 1*1 - 1*3/2 = NET_2 = -0{,}5$

*При* $x_1=1$, $x_2=0 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 1*1 + 0*1 - 1*3/2 = NET_3 = -0{,}5$

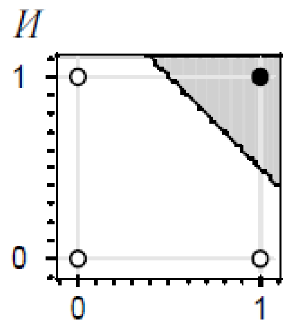
*При* $x_1=1$, $x_2=1 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 1*1 + 1*1 - 1*3/2 = NET_4 = 0{,}5$

Знаем, что пороговая функция для смещения равна

$$OUT = \begin{cases} 1, \text{если } NET > 0 \\ 0, \text{если } NET \leq 0 \end{cases}$$

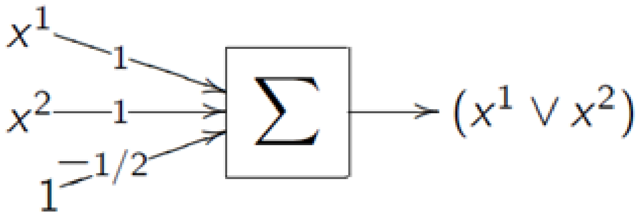
Тогда на выходе: ***OUT₁ = 0, OUT₂ = 0, OUT₃ = 0, OUT₄ = 1***

Пороговая функция делит входное векторное пространство на две части гиперплоскостью ***($x_1 + x_2 = 1,5$)***, классифицируя входные векторы как относящиеся к первому классу, если выходной сигнал ***OUT > 0*** или ко второму классу, если выходной сигнал ***OUT ≤ 0.***




<u>Рассчитываем ***логическое «ИЛИ»***</u>

Установим ***$x_0 = 1$***, а ***$w_0 = -1/2$***




*При* $x_1=0,\ x_2=0 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 0*1 + 0*1 - 1*1/2 = NET_1 = -0,5$

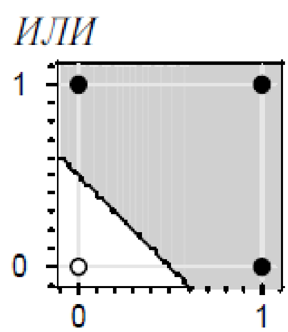
*При* $x_1=0,\ x_2=1 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 0*1 + 1*1 - 1*1/2 = NET_2 = 0,5$

*При* $x_1=1,\ x_2=0 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 1*1 + 0*1 - 1*1/2 = NET_3 = 0,5$

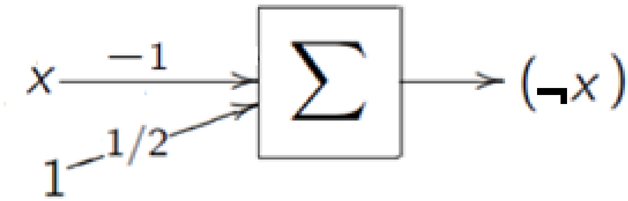
*При* $x_1=1,\ x_2=1 \rightarrow x_1*w_1 + x_2*w_2 + x_o*w_0 = 1*1 + 1*1 - 1*1/2 = NET_4 = 1,5$

Тогда согласно активационной функции на выходе получаем ***$OUT_1 = 0$, $OUT_2 = 1$, $OUT_3 = 1$, $OUT_4 = 1$***

Графическая интерпретация будет следующей при гиперплоскости ***($x_1 + x_2 = 0,5$)***.



Аналогично рассчитываем логическое отрицание при $x_0 = 1$, а $w_0 = 1/2$

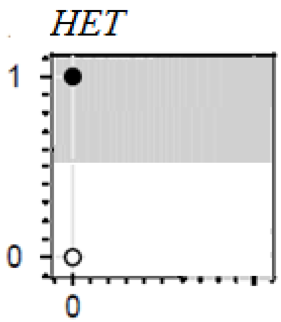



При x = 0 → $x^* w + x_0^* w_0 = 0*(-1) + 1* 1/2 = NET_1 = 1/2$

При x = 1 → $x^* w + x_0^* w_0 = 1*(-1) + 1* 1/2 = NET_2 = -1/2$

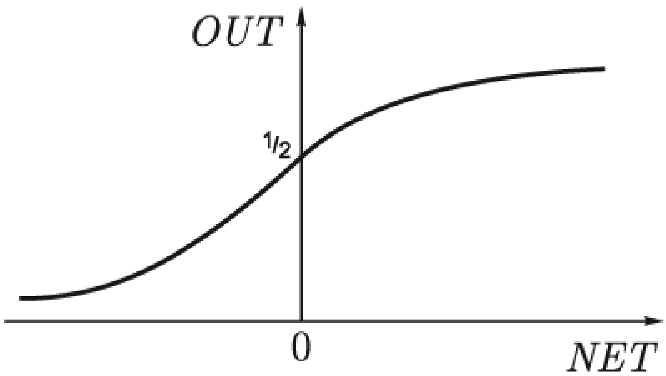
Тогда на выходе: ***$OUT_1$ = 1, $OUT_2$ = 0***

Графическая интерпретация будет следующей при гиперплоскости ***(x = 0,5)***.




**2.** СИГМОИДАЛЬНАЯ ФУНКЦИЯ ИЛИ СИГМОИД

Если блок ***F*** сужает диапазон изменения величины ***NET*** так, что при любых значениях ***NET*** значения ***OUT*** принадлежат некоторому конечному интервалу, то ***F*** называется «сжимающей» функцией. В качестве «сжимающей» функции часто используется ***"сигмоидальная"*** (S-образная) функция, показанная. Часто под сигмоидой понимают логистическую функцию.




Эта функция математически выражается как

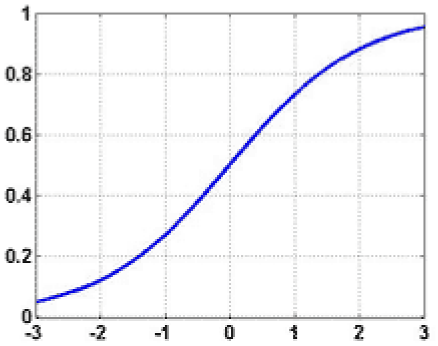
$$F(x) = 1/(1 + e^{-x}).$$

Таким образом,

$$OUT = \frac{1}{1+e^{-NET}}.$$

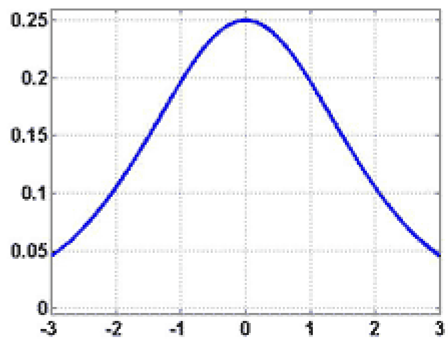
Одна из причин, по которой сигмоид используется в нейронных сетях, это простое выражение его производной через саму функцию (которое и позволило существенно сократить вычислительную сложность *метода обратного распространения ошибки, сделав его применимым на практике*).

Функция




$$f(n) = \frac{1}{1+\exp(-n)}$$

Производная




$$f(n) = \frac{exp(-n)}{(1+exp(-n))^2}$$

где **exp** = $\mathbf{e^n}$ (экспонента)

Тогда

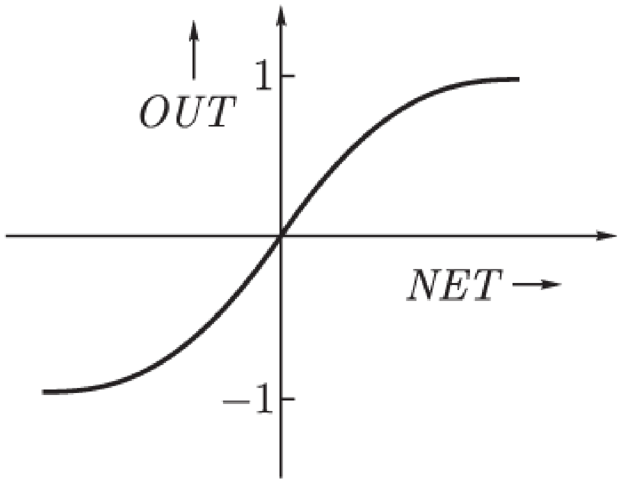
$$\frac{\partial OUT}{\partial NET} = OUT(1 - OUT)$$

Его дополнительное преимущество состоит в автоматическом контроле усиления. Для слабых сигналов (т.е. когда ***OUT*** близко к нулю) кривая вход-выход имеет сильный наклон, дающий большое усиление. Когда величина сигнала становится больше, усиление падает. Таким образом, большие сигналы воспринимаются сетью без насыщения, а слабые сигналы проходят по сети без чрезмерного ослабления. Другими словами центральная область логистической функции, имеющая большой коэффициент усиления, решает проблему обработки слабых сигналов, в то время как области с падающим усилением на положительном и отрицательном концах подходят для больших возбуждений.

## 3. ГИПЕРБОЛИЧЕСКИЙ ТАНГЕНС

Другой широко используемой активационной функцией является ***гиперболический тангенс***. По форме она сходна с логистической функцией и часто используется биологами в качестве математической модели активации нервной клетки. В качестве активационной функции искусственной нейронной сети она записывается следующим образом:

$$\mathbf{OUT = th(x)}.$$




Подобно логистической функции гиперболический тангенс является ***S-образной функцией***, но он симметричен относительно начала координат, и в точке ***NET = 0*** значение выходного сигнала ***OUT*** равно нулю. В отличие от логистической функции, гиперболический тангенс принимает значения различных знаков, и это его свойство применяется для целого ряда сетей.

Рассмотренная простая модель искусственного нейрона игнорирует многие свойства своего биологического двойника. Например, она не принимает во внимание задержки во времени, которые воздействуют на динамику системы. Входные сигналы сразу же порождают выходной сигнал. И, что более важно, она не учитывает воздействий функции частотной модуляции или синхронизирующей функции биологического нейрона, которые ряд исследователей считают решающими в нервной деятельности естественного мозга.