

## 6. Связь формальных нейронных сетей с теорией перцептронов

В этой главе, чтобы сравнить формальные нейронные сети с теорией перцептронов рассмотрены примеры решения задачи классификации с точки зрения трёх разных подходов:

- 1) пороговый элемент (простейший случай);
- 2) формальные нейронные сети;
- 3) теория перцептронов.

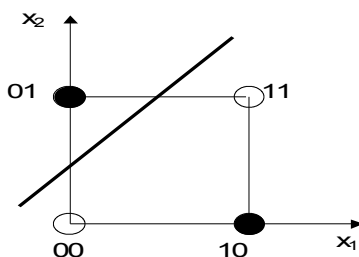
После этого сделаны некоторые качественные выводы о преимуществах и недостатках каждого из подходов.

### 6.1. Решение задачи классификации с помощью порогового элемента

Пусть требуется построить пороговый элемент, реализующий функцию:

$$F = x_1 \neg x_2 \vee \neg x_1 x_2$$

В геометрической интерпретации единичным значениям этой функции соответствуют четыре точки, а разделяющей функции соответствует прямая.



Видно, что как бы мы ни смещали и поворачивали прямую, невозможно её провести так, чтобы на одной линии находились только чёрные точки, а на другой – только светлые.

Это означает, что каким бы ни брать порог  $\theta$  и весовые коэффициенты  $w_1$  и  $w_2$ , на одном пороговом элементе невозможно реализовать данную функцию.

### 6.2. Решение задачи классификации с помощью формальных нейронных сетей

Реализуем для формального нейрона эту же функцию:

$$F(x) = x_1 \neg x_2 \vee \neg x_1 x_2$$

Чтобы синтезировать формальный нейрон нужно решить следующую систему уравнений (вывод этого уравнения можно найти в работе [2], а его обобщение на случаи, где количество переменных  $n$ , можно найти в работе [3]):

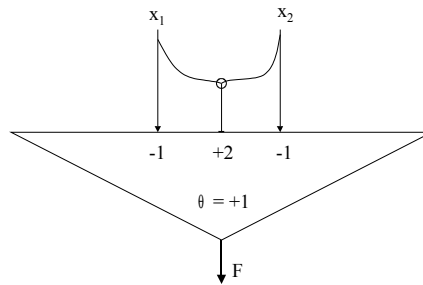
$$\left. \begin{aligned} X^o + X_y &= d \\ Y^o + Y_x &= b \\ X^o + Y^o &= a \end{aligned} \right\}$$

В общем случае система имеет бесконечное число решений, т.к. в ней одна переменная избыточна. При задании этой переменной система решается полностью и однозначно. Обычно это делается исходя из соображений минимизации числа волокон синтезируемого нейрона. Стоимость переменной показывает число волокон, необходимых для её реализации. Например, стоимость переменной  $X^o$  равна 1, а стоимость  $X_y$  равна 2 и т.д. Так как выбор задаваемой неизвестной является произвольным, то, с целью максимальной экономии волокон необходимо задать её равной нулю и выбрать из тех переменных, стоимости которых максимальны. В данном случае это  $X_y$  и

$Y_x$ . Задавая одну из них равной нулю и решая систему, получим нейрон, в котором число волокон равно 4. Ниже в таблице приведены все четыре варианта решений системы.

| Неизвестное | Стоимость | $X_y = 0$   | $Y_x = 0$   | $X^o = 0$ | $Y^o = 0$ |
|-------------|-----------|-------------|-------------|-----------|-----------|
| $X^o$       | 1         | d           | $a - b$     | 0         | a         |
| $X_y$       | 2         | 0           | $b + d - a$ | d         | $d - a$   |
| $Y^o$       | 1         | $a - d$     | b           | a         | 0         |
| $Y_x$       | 2         | $b + d - a$ | 0           | $b - a$   | b         |

Тогда для данной функции формальный нейрон будет выглядеть так:



Здесь  $\delta = 2$ ,  $n = 3$  и условие возбуждения нейрона записывается так:

$$\sum_{i=1}^3 w_i \alpha_i - \theta \geq 0,$$

где  $w_1 = -1$ ;  $w_2 = +2$ ;  $w_3 = -1$ ;  $\theta = +1$ .

Тогда разделяющая функция в пространстве синапсов будет иметь вид:

$$-\alpha_1 + 2\alpha_2 - \alpha_3 - 1 = 0$$

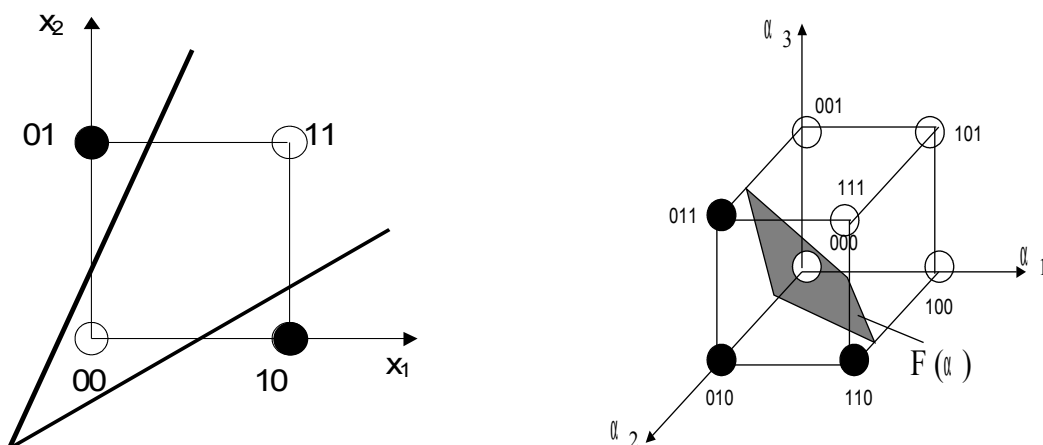
А в пространстве синапсов функция  $F(\mathbf{x}) = \mathbf{x}_1 \neg \mathbf{x}_2 \vee \neg \mathbf{x}_1 \mathbf{x}_2$  будет иметь вид:

$$F(\alpha) = \alpha_1 \alpha_2 \neg \alpha_3 \vee \neg \alpha_1 \alpha_2 \alpha_3 \vee \neg \alpha_1 \alpha_2 \neg \alpha_3$$

Точки на которых функция  $F(\alpha)$  равна 1 зачернены. Так как

$$\mathbf{x}_1 = \alpha_2 - \alpha_1 \text{ и } \mathbf{x}_2 = \alpha_2 - \alpha_3,$$

то при переходе к пространству входов ось  $\mathbf{x}_1$  будет находиться на плоскости  $\alpha_1 \alpha_2$ , а ось  $\mathbf{x}_2$  - на плоскости  $\alpha_2 \alpha_3$ . Линия пересечения разделяющей плоскости с этими плоскостями является ломаной линией, проекция которой в пространстве входов также будет ломанной. Ниже показана геометрическая интерпретация функции  $F(\mathbf{x}) = \mathbf{x}_1 \neg \mathbf{x}_2 \vee \neg \mathbf{x}_1 \mathbf{x}_2$  в пространстве синапсов и в пространстве входов.



Таким образом, формальный нейрон преобразует двухмерное пространство входов в трёхмерное пространство синапсов, где удаётся провести разделяющую плоскость, разделяя на нужные классы. Проекция этой плоскости в пространстве входов представляется ломаной линией. Излом

разделяющей линии соответствует взаимодействию волокон нейрона. Сколько узлов или точек взаимодействия имеется во входных волокнах нейрона, столько и изломов будет иметь разделяющая линия.

Таким образом, в общем случае, когда нейрон имеет  $\delta$  входов и  $n$  синапсов, нужное разделение на классы производится в  $n$  – мерном пространстве синапсов гиперплоскостью, проекция которой в  $\delta$  – мерном пространстве входов представляет собой разделяющую гиперплоскость со столькими изгибами, сколько точек взаимодействия имеется во входных волокнах нейрона.

### 6.3. Решение задачи классификации с помощью теории перцептронов

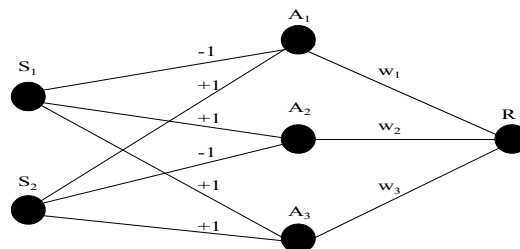
Реализуем для элементарного перцептрона эту же функцию:

$$F(x) = x_1 \neg x_2 \vee \neg x_1 x_2$$

Причём, следует отметить, что для элементарного перцептрона знать функцию в таком строгом виде нет необходимости, а нужно лишь требуемую задачу представить в виде набора стимулов, и указать какой из стимулов относится к какому классу. Например, для нашей функции будет эквивалентно следующее представление:

|          | Вход 1 | Вход 2 | Класс |
|----------|--------|--------|-------|
| Стимул 1 | 0      | 0      | -     |
| Стимул 2 | 1      | 1      | -     |
| Стимул 3 | 0      | 1      | +     |
| Стимул 4 | 1      | 0      | +     |

Стимул 1 означает в теории перцептронов, что стимула нету, и поэтому такой стимул в любом случае будет отнесён к отрицательному классу. Для решения этой задачи может быть использована следующая структура перцептрона:



Начальные значения весов и порог для всех элементов равен нулю. Тогда матрица SA будет такой:

|                | A <sub>1</sub> | A <sub>2</sub> | A <sub>3</sub> |
|----------------|----------------|----------------|----------------|
| S <sub>1</sub> | 0              | 0              | 1              |
| S <sub>2</sub> | 1              | 0              | 1              |
| S <sub>3</sub> | 0              | 1              | 1              |

В процессе нахождения (коррекции) весов  $\alpha$  - системой подкрепления веса будут изменяться следующим образом:

| Веса                  | Итерации |   |   |   |    |   |    |   |    |
|-----------------------|----------|---|---|---|----|---|----|---|----|
|                       | 1        |   |   | 2 | 3  |   | 4  |   | 5  |
| $w_1$                 | 0        | 1 | 1 | 1 | 1  | 2 | 2  | 2 | 2  |
| $w_2$                 | 0        | 0 | 1 | 1 | 1  | 1 | 1  | 2 | 2  |
| $w_3$                 | -1       | 0 | 1 | 0 | -1 | 0 | -1 | 0 | -1 |
| Корректирующий стимул | 1        | 2 | 3 | 1 | 1  | 2 | 1  | 3 | 1  |

Таким образом на пятой итерации найдены веса при которых происходит правильное разделение на классы. Геометрическая интерпретация решения задачи классификации для теории перцептронов будет аналогична геометрической интерпретации формальных нейронных сетей.

#### 6.4. Преимущества и недостатки в подходах к решению задачи классификации

Как видно из приведённых примеров решение задачи классификации для произвольной задачи решается только формальными нейронами или элементарным перцептроном, если же использовать один пороговый элемент, то например задача "XOR" не имеет решения.

Из описания перцептрона, видно, что он представляет собой сеть пороговых элементов, количество которых и определяет решается или нет поставленная задача. При этом следует отметить, что перцептрон состоит по крайней мере из 5 пороговых элементов, поэтому его функциональные возможности никоим образом нельзя сравнивать с функциональными способностями порогового элемента, что например было сделано Васерманом в работе [?].

Формальный нейрон в общем виде представляет собой более сложный элемент, чем элементарный перцептрон, и его можно сравнить с более сложной структурой перцептрона с перекрестными связями, которые в данной работе не рассматриваются (описание структуры перцептрона можно найти, например в работе [1]). Поэтому с помощью формального нейрона можно решать практически любые задачи классификации, но т.к. формальный нейрон трудно делится на части, которые можно было бы отдельно анализировать, он труднее подойдёт к изучению и настройке на конкретный класс задач. Сравнительно, у элементарного перцептрона отдельно подойдёт анализу любой из пороговых элементов.

Наиболее серьёзным отличием между формальными нейронами и элементарным перцептроном является процесс нахождения весов связей.

Для формальных нейронов этот процесс представляет собой процесс нахождения целевой функции с системой ограничений, т.е. представляет собой детерминированный процесс. При большом количестве входов эта задача существенно усложняется и с трудом подойдёт к решению.

Для элементарного перцептрона этот процесс представляет собой целенаправленный случайный процесс коррекции весов, т.е. является стохастическим процессом с элементами адаптации. При этом время нахождения решения для произвольной задачи классификации меньше зависит от количества стимулов, в следствии чего с помощью, например  $\alpha$  - системы подкрепления, можно быстро найти решение при большом количестве стимулов, что в случае формальных нейронов было бы затруднительно.

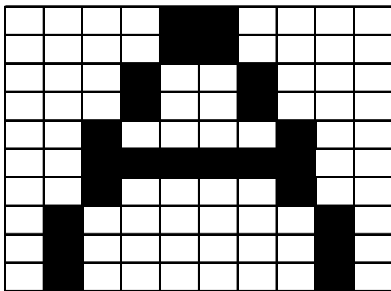
## 7. Выполнение первого и второго требования для элементарного $\alpha$ - перцептрона

### 7.1. Внешнее восприятие объектов $\alpha$ - перцептронами

Повторим, первое требование, которое было предъявлено для модели мозга:

**Исходными данными должны быть реальные объекты внешнего мира (изображения, звук, ощущения)**

Для начала рассмотрим, как для  $\alpha$  - перцептрона воспринимаются изображения.



Любое чернобелое изображение можно представить в виде сетчатки, в котором элементы сетчатки могут быть заполнены или нет. Тогда качество изображения определяется только разрешением этой сетчатки, а следовательно, количеством элементов сетчатки.

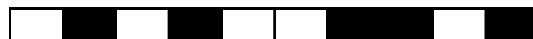
Именно для восприятия такого изображения наиболее пригодны перцептроны. В перцептроне сетчаткой уже являются  $S$  – элементы, значит, для работы перцептрона достаточно спроектировать любое чернобелое изображение на  $S$  – элементы, и больше исходных данных не требуется. Это соответствует реальному зрению человека.

Если же изображение цветное, перцептроны без дополнительной обработки не в состоянии его воспринять. Здесь есть несколько вариантов решения этой проблемы:

- 1) Использовать другие более сложные виды перцептронов или вообще другие модели восприятия;
- 2) Обработать цветное изображение, чтобы его представить в виде чернобелого изображения, а цвет подавать на сетчатку в закодированном виде как дополнительный параметр;
- 3) Закодировать цвет каким – либо образом, например в простейшем случае ввести 256 цветов, а в сложнейшем в виде RGB кодирования. Каждый элемент сетчатки разбить на более мелкую сетку. Тогда каждая точка цветного изображения представляется дополнительной сетчаткой из закодированного в двоичном формате цвета точки изображения.

В аналогичном аспекте рассмотрим как для  $\alpha$  - перцептрона воспринимаются звуки, музыка или голос человека.

Здесь аналогично чернобелому изображению может восприниматься последовательность звуков только двух видов, например принятое кодирование типа “точка - тире” в радиосвязи. В этом случае сетчатка представляется не двумерной матрицей, а двоичным вектором, размерность которого определяет одновременно воспринимаемую длину последовательности двоичного звука. Это выглядит так:



При моделировании восприятия реальной музыки или голоса человека, возникают проблемы аналогичные восприятию цветного изображения. Автор предлагает их решить аналогичным образом, а именно – в простейшем случае закодировать, например голос человека в виде последовательности разнообразных, которые человек может произносить, в наипростейшем случае можно ограничиться буквами алфавита. В более сложном случае голос или музыку можно закодировать учитывая длины и амплитуды звуковых волн. Если каждый элемент сетчатки разбить на более мелкую сетчатку, тогда каждая единица будет представляться дополнительной сетчаткой из закодированных в двоичном формате единицы звука. Тогда, например, слово “колонисты” может быть представлено так:

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| А |   |   |   |   |   |   |   |   |   |
| Б |   |   |   |   |   |   |   |   |   |
| В |   |   |   |   |   |   |   |   |   |
| Г |   |   |   |   |   |   |   |   |   |
| Д |   |   |   |   |   |   |   |   |   |
| Е |   |   |   |   |   |   |   |   |   |
| Ж |   |   |   |   |   |   |   |   |   |
| З |   |   |   |   |   |   |   |   |   |
| И |   |   |   |   |   | ■ |   |   |   |
| Й |   |   |   |   |   |   |   |   |   |
| К | ■ |   |   |   |   |   |   |   |   |
| Л |   |   | ■ |   |   |   |   |   |   |
| М |   |   |   |   |   |   |   |   |   |
| Н |   |   |   |   | ■ |   |   |   |   |
| О |   | ■ |   | ■ |   |   |   |   |   |
| П |   |   |   |   |   |   |   |   |   |
| Р |   |   |   |   |   |   |   |   |   |
| С |   |   |   |   |   |   | ■ |   |   |
| Т |   |   |   |   |   |   |   | ■ |   |
| У |   |   |   |   |   |   |   |   |   |
| Ф |   |   |   |   |   |   |   |   |   |
| Х |   |   |   |   |   |   |   |   |   |
| Ц |   |   |   |   |   |   |   |   |   |
| Ч |   |   |   |   |   |   |   |   |   |
| Ш |   |   |   |   |   |   |   |   |   |
| Щ |   |   |   |   |   |   |   |   |   |
| Ъ |   |   |   |   |   |   |   |   |   |
| Ы |   |   |   |   |   |   |   |   | ■ |
| Ь |   |   |   |   |   |   |   |   |   |
| Э |   |   |   |   |   |   |   |   |   |
| Ю |   |   |   |   |   |   |   |   |   |
| Я |   |   |   |   |   |   |   |   |   |

Как видно из приведённого выше анализа, проблема восприятия реальных объектов в  $\alpha$  - перцептронах решается для простейших объектов, но вводя некоторые дополнения можно сложные объекты разбить на простые, с которыми  $\alpha$  - перцептрона легко справляется.

## 7.2. Внутреннее восприятие объектов $\alpha$ - перцептронами

В теории перцептронов можно говорить не только о внешнем восприятии реальных объектов, но также и о внутреннем. Проводя аналогию с реальным мозгом человека, а конкретнее с восприятием человека можно сказать, что

Внешнее восприятие – это то как на самом деле выглядят проекции реальных объектов.

Внутреннее восприятие – это то как человек воспринимает проекции реальных объектов.

Внутреннее восприятие для  $\alpha$  - перцептронов осуществляется А – элементами посредством связей от S – элементов к А – элементам. Описанием внутреннего восприятия можно считать SA – матрицу:

|       | $A_1$    | $A_2$    | ... | $A_m$    |
|-------|----------|----------|-----|----------|
| $S_1$ | $q_{11}$ | $q_{12}$ | ... | $q_{1m}$ |
| $S_2$ | $q_{21}$ | $q_{22}$ | ... | $q_{2m}$ |
| ...   | ...      | ...      | ... | ...      |
| $S_n$ | $q_{n1}$ | $q_{n2}$ | ... | $q_{nm}$ |

где  $A_1, A_2, \dots, A_m$  – соответствующий А – элемент в перцептроне;

$S_1, S_2, \dots, S_n$  – стимул из внешней среды.

Значения элементов данной матрицы определяются следующим образом:

$$q_{ij} = \begin{cases} 1 & \text{если } A_j \text{ активизируется при стимуле } S_i \\ 0 & \text{если } A_j \text{ неактивизируется при стимуле } S_i \end{cases}$$

Следует заметить, что для количественного анализа сложности внутреннего восприятия, в смысле пересечения воспринятых понятий, имеет смысл, воспользоваться G – матрицей элементарного  $\alpha$  - перцептрона.

Имеет смысл, обратит внимание на следующую проблему. Рассмотрим следующую SA матрицу:

|       | $A_1$ | $A_2$ |
|-------|-------|-------|
| $S_1$ | 0     | 1     |
| $S_2$ | 0     | 1     |

Мы видим, что в данном случае стимул  $S_1$  воспринимается  $\alpha$  - перцептроном, точно также как стимул  $S_2$ . Это наталкивает нас на проблему, что существует вероятность того, что реально два разных объекта (стимула) человек будет воспринимать одинаково. А если существует такая вероятность,

то от  $\alpha$  - перцептрона, который является лишь моделью восприятия, нельзя ждать большего. Остаётся только показать, что эта вероятность в  $\alpha$  - перцептронах настолько же мала как и в мозге человека.

#### 7.4. Выполнение второго требования для элементарного $\alpha$ - перцептрона

Повторим, второе требование, которое было предъявлено для модели мозга:

**Модели должны уметь различать любые произвольные классификации (решать нелинейные задачи)**

Выполнение этого требования элементарным  $\alpha$  - перцептроном, считается доказанным. Поэтому остаётся только сослаться на главу 5, в которой формально показано, что элементарный  $\alpha$  - перцептрон теоретически способен полностью выполнить это требование.

Подтверждение этого на практике, видно из простого примера решение задачи произвольной классификации элементарным  $\alpha$  - перцептроном, приведенное в главе 6.

И как будет видно из экспериментов (3 часть данной работы) решение задачи произвольной классификации элементарным  $\alpha$  - перцептроном практически подтверждается и при достаточно сложных задачах.



## 8. Выполнение третьего требования для элементарного $\alpha$ - перцептрона

### 8.1. Временные затраты

Повторим, третье требование, которое было предъявлено для модели мозга:

**Модели должны требовать по возможности меньше аппаратных ресурсов**

Рассмотрим, на какие расчёты, тратится время (а соответственно, как долго используются аппаратные ресурсы) при моделировании задачи произвольной классификации элементарным  $\alpha$  - перцептроном.

Естественно – это нельзя рассмотреть, не конкретизируя программное обеспечение, используемое для моделирования таких решений. Поэтому мы рассмотрим затраты временных ресурсов, для программного обеспечения, написанного автором данной работы. При этом не будем рассматривать временные ресурсы, затраченные на инициализацию, например на создание исходной структуры  $\alpha$  - перцептрона. Итак, временные ресурсы тратятся на:

1. Расчёт, активирован  $i$ -й  $A$  - элемент или нет, при поступлении стимула  $S_i$  на сетчатку. (Это сводится к заполнению матрицы  $SA$ , рассмотренной в главе 7)
2. Расчёт, активирован или нет  $R$  - элемент.
3. Определение, правильна или нет реакция  $R$  - элемента.
4. В случае неправильной реакции  $R$ - элемента, происходит обучение, т.е. изменение весов  $A - R$  связей.

Если  $R$  – элементов  $N_r$ , то пункты 2, 3, 4 выполняются  $N_r$  раз. После этого снова нужно выполнить пункты 2, 3, 4  $N_r$  раз.

Исходя из такого анализа можно написать формулы временных затрат при решении задачи произвольной классификации:

$$T = n N_a t_{\text{актив.}}^a + N_r I_n (t_{\text{актив.}}^r + t_{\text{выхода}} + t_{\text{коррекции}}),$$

где  $n$  - количество стимулов;

$N_a$  - количество  $A$  - элементов;

$t_{\text{актив.}}^a$  - время расчёта активирован  $A$  - элемент или нет;

$t_{\text{актив.}}^r$  - время расчёта активирован  $R$  - элемент или нет;

$t_{\text{коррекции}}$  - время изменения весов  $A - R$  связей;

$I_n$  - количество итераций при обучении;

$t_{\text{выхода}}$  - время затрачиваемое на определение правильна реакция или нет (является константой).

Далее:

$$t_{\text{актив.}}^a = N_s t_{\text{сумм.}}; t_{\text{актив.}}^r = N_a t_{\text{произв.}} t_{\text{сумм.}}; t_{\text{коррекции}} = N_a t_{\text{сумм.}},$$

где  $N_s$  – число элементов сетчатки;

$t_{\text{сумм.}}$  – время выполнения операции прибавления или вычитания между двумя числами;

$t_{\text{произв.}}$  - время выполнения операции умножения между двумя числами.

Не вдаваясь в детальный анализ видно, что временные ресурсы сильно зависят от  $t_{\text{сумм.}}$  и  $t_{\text{произв.}}$ , т.е. от того какие математические действия требуются для расчёта функции активации.

Для сравнения рассмотрим модель, где используется в качестве функции активации нелинейная функция. Тогда:

$$t_{\text{актив.}}^a = N_s t_{\text{сумм.}} t_{\text{ф.}}; t_{\text{актив.}}^r = N_a t_{\text{произв.}} t_{\text{сумм.}} t_{\text{ф.}},$$

где  $t_{\text{ф.}}$  – время расчёта нелинейной функции;

Тогда приблизительная оценка насколько такие модели требуют больше временных расходов равна  $0,7 t_{\phi}$ .

Как видим  $\alpha$  - перцептрон является намного более оптимизированной системой, в плане потребления аппаратных ресурсов.

## 8.2. Затраты памяти

Кроме выше проанализированных временных затрат, важно учитывать затраты используемого объёма памяти.

При использовании целых чисел от 0 до 65535 требуется 2 байта. При использовании вещественных чисел одинарной точности – 4 байта.

Кроме того, следует учитывать, что работа с вещественными числами увеличивает время расчётов, т.е.

$$t_{\text{сумм.}} = t_{\text{сумм.}} + t_{\text{конвер.}}, \text{ причём } t_{\text{сумм.}} < t_{\text{конвер.}},$$

где  $t_{\text{конвер.}}$  – время требуемое для того, чтобы число можно было представить в формате вещественного числа.

## 8.3. Нелинейные функции активации

Выше было показано, что использование нелинейных функций активации значительно замедляет работу перцептрона. Здесь же будет показано, для каких целей вообще используются нелинейные функции активации.

Рассмотрим слабые стороны модели перцептрона. В теореме 1 (глава 5) показано, что существует хотя бы один перцептрон решающий задачу произвольной классификации.

Проблема состоит в том, что мы не знаем структуру этого перцептрона. Для этого в теореме 3 выдвигается достаточное условие, которое накладывает некоторые ограничения на структуру перцептрона. Если эти ограничения выполняются, то перцептрон способен решить задачу произвольной классификации.

Таких ограничений два (описаны в следствиях 1 и 2 теоремы 3):

- 1)  $G$  - матрица не должна быть особенной;
- 2) количество  $A$  - элементов должно быть равно или больше количества стимулов.

Исходя, из целей оптимизации по времени, мы всегда должны стремиться уменьшить число  $A$  - элементов. Но если мы хотим, чтобы перцептрон решал любую задачу классификации, то мы должны по минимуму, иметь столько же  $A$  - элементов, сколько будет стимулов.

Выполнение же первого ограничения связано с большими трудностями. Это связано с тем, что нельзя гарантировать того, что  $G$  - матрица не будет особенной. Но если использовать нелинейные функции активации, то можно попробовать уйти от этой проблемы. Использование нелинейных функций активации рассмотрим на следующих простейших примерах.

Случай 1. Пусть  $G$  - матрица при использовании линейной функции  $\alpha x$  активации имеет вид:

$$G = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

Эта матрица особенна, и, следовательно, существуют некоторые классификации, не имеющие решений. Изменим теперь функцию активации на  $\alpha^2 v$ , что приведёт к тому, что матрица примет вид:

$$G = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 4 & 9 \\ 5 & 9 & 16 \end{pmatrix}$$

Такая матрица уже не является особенной, а это означает, что при нелинейной функции активации все классификации будут иметь решения.

Случай 2. Пусть матрица  $G$  при использовании линейной функции активации  $\alpha v$  имеет вид:

$$G = \begin{pmatrix} 3 & 5 & 8 \\ 4 & 12 & 15 \\ 5 & 13 & 17 \end{pmatrix}$$

Эта матрица неособенная, так что здесь имеется решение для каждой классификации  $C(W)$ . Возьмём теперь функцию активации вида  $\alpha^2 v$ . Тогда матрица

$$G = \begin{pmatrix} 9 & 25 & 64 \\ 16 & 144 & 225 \\ 25 & 169 & 289 \end{pmatrix}$$

становится особенной, а это означает, что имеется классификация  $C(W)$ , для которой решения не существует.

Основываясь на работы Джозефа и А.Гамбы можно показать, как для конкретной задачи классификации найти оптимальную функцию активации, с точки зрения эффективности различения и скорости обучения.

Но в этом случае перцептрон приспособлен только для конкретной задачи, т.е. он теряет смысл универсальности. Кроме того, численная оценка фактической выгоды, по-видимому, практически невелика.