Тема 4

Основные понятия теории нейронных сетей

Для описания нейронных сетей в нейроинформатике существует специальная "схемотехника", в которой элементарные устройства – это простые функциональные элементы (сумматоры, синапсы, нейроны и т. п.). Они объединяются в вычислительные сети, предназначенные для решения задач. Поэтому естественным для описания процессов в нейроподобных системах является язык функциональных схем или язык системного анализа.

В аппаратной реализации нейронных сетей, в профессиональном программном обеспечении, а также в математических моделях, описывающих нейросети, все эти элементы присутствуют необязательно как отдельные блоки. Используемая в нейроинформатике идеальная схемотехника - это особый язык для представления структуры нейронных сетей. При программной и аппаратной реализации, выполненные на этом языке описания, переводятся на языки другого уровня, более пригодные для воплощения. Традиционно используемым для описания нейронных сетей математическим языком является аппарат векторной и матричной алгебры.

Основными функциональными элементами традиционно считаются: *синапс, адаптивный сумматор, активатор, точка ветвления.* Из этих простых элементов строятся более сложные – *формальные нейроны*.

Определение 4.1. *Линейная связь (синапс)* – элемент, умножающий входной сигнал *x* на «вес синапса» α (см. рис. 4.2).

$$x \xrightarrow{\alpha} \alpha x$$

Рис. 4.1. Синапс

Адаптивный сумматор – элемент, вычисляющий скалярное произведение входного сигнала $\overline{x}$ на вектор параметров $\overline{\alpha}$. Адаптивным называем его из-за наличия вектора настраиваемых параметров $\overline{\alpha}$. Очевидно, что синапсы являются элементами структуры адаптивного сумматора (см. рис. 4.3).

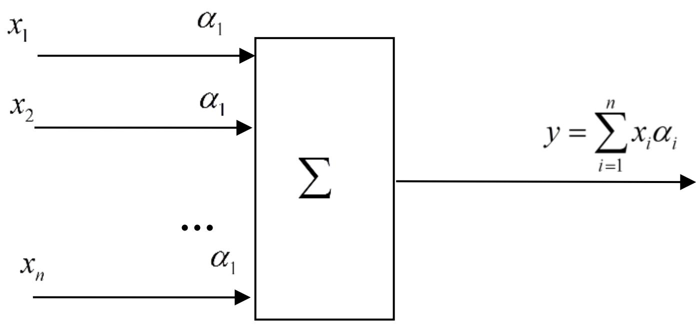

Рис. 4.2 Адаптивный сумматор

Нелинейный преобразователь сигнала (функция активации нейрона, активатор) – элемент, получающий скалярный входной сигнал x и переводящий его в $\varphi(x)$ (см. рис. 4.4).

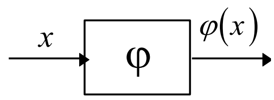

Рис. 4.3. Нелинейный преобразователь сигнала

Точка ветвления - получает скалярный входной сигнал x и передает его всем своим выходам. Такие выходы называют *аксонами* (см. рис. 4.5).

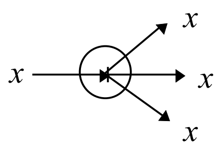

Рис. 4.4. Точка ветвления.

Преобразователь Паде – получает два скалярных сигнала x_1 и x_2 и выдает произведение $x_{1*} x_{2}$.

Стандартный формальный нейрон – элемент, составленный из входного сумматора, нелинейного преобразователя и точки ветвления на выходе (см. рис.).

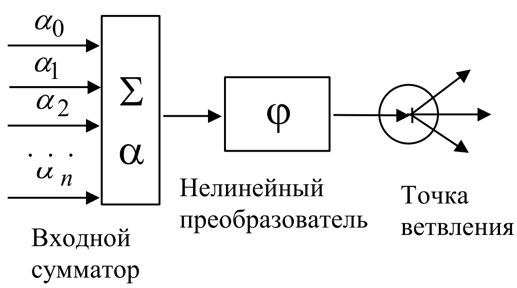

Рис. 4.5. Формальный нейрон стандартного вида

Классификация искусственных нейронных сетей

Приведем классификацию искусственных нейронных сетей в зависимости от различных характеристик, так как она приведена в [5]:

1. По типу входной информации:
   1.1. Аналоговые нейронные сети, которые используют информацию в форме действительных чисел.
   1.2. Цифровые нейронные сети. Они оперируют с информацией, предоставленной в цифровом виде.
2. По характеру обучения:
   2.1. С учителем, когда известно выходное пространство решений нейронной сети.
   2.2. Без учителя. В этом случае нейронная сеть формирует выходное пространство решений только на основе входных воздействий. Такие сети называются самоорганизующимися.
3. По характеру настройки синапсов:
   3.1. Сети с фиксированными связями. В этом случае весовые коэффициенты нейронной сети выбираются сразу, исходя из условия задачи и далее не меняются.
   3.2. Сети с динамическими связями. Для них в процесс обучения происходит настройка синаптических связей.
4. По методу обучения:
   4.1. Нейронные сети с алгоритмом обратного распространения ошибки.
   4.2. Нейронные сети с конкурентным обучением.
   4.3. Нейронные сети, использующие правило Хебба.
   4.4. Нейронные сети с гибридным обучением, в которых применяются различные алгоритмы обучения.
5. По характеру связей:
   5.1. Нейронные сети с прямыми связями (feedforward networks). В таких сетях информация распространяется только в одном направлении от слоя к слою. К ним относятся персептронные сети: однослойные, многослойные, гомогенные и гетерогенные.
   5.2. Нейронные сети с обратным распространением информации (feedback networks). Они характеризуются как прямым, так и обратным распространением информации между слоями нейронной сети. К сетям с такими связями относятся следующие:

5.2.1. Релаксационные, в которых циркуляция информации происходит до тех пор, пока не перестанут изменяться выходные значения нейронной сети (состояние равновесия). К ним относятся нейронные сети Хопфолда, Хэмминга и двунаправленная ассоциативная память (сеть Коско).

5.2.2. Многослойные сети, в которых отсутствует процесс релаксации. К ним относится, в частности, многослойный персептрон.

6. По архитектуре и обучению различают следующие известные нейронные сети:
   6.1. Персептронные (многослойные) сети с прямыми связями.
   6.2. Самоорганизующиеся нейронные сети. К ним относятся:
      6.2.1. Нейронные сети Кохонена.
      6.2.2. Нейронные сети адаптивного резонанса.
      6.2.3. Рециркуляционные нейронные сети.
   6.3. Нейронные сети с обратными связями:
      6.3.1. Нейронные сети Хопфилда.
      6.3.2. Нейронные сети Хемминга.
      6.3.3. Двунаправленная ассоциативная память.
      6.3.4. Рекуррентные нейронные сети.
   6.4. Гибридные нейронные сети:
      6.4.1. Нейронные сети встречного распространения.
      6.4.2. Нейронные сети с радиально-базисной функцией активации.

Формальное описание искусственных нейронных сетей

Определим математическую модель формального нейрона. Под формальным нейроном будем понимать объект следующего вида:

$$F = (X, Y, S, g)\,, \qquad (4.1)$$

где $S = \{\bar{s}\} \subseteq R^{N_S}$ - множество состояний, $X = \{\bar{x}\} \subseteq \{R \cup e\}^{N_X}$ - множество входных сигналов, $Y = \{\bar{y}\} \subseteq R^{N_Y}$ - множество выходных сигналов, $e \notin R$ - элемент, используемый для обозначения отсутствия сигнала, $g : S \times X \to Y$ - функция преобразования $\bar{y} := g(\bar{s}, \bar{x})$. Как правило, функция g определяется следующими формулами

$$y = \varphi(z)\,, \qquad (4.2)$$

$$z = \sum_{i=1}^{N_X} x_i s_i \qquad (4.3)$$

где

$N_X = N_S$ - число входов нейрона;

s_i - вес синапса ($i = 1, \ldots, n$);

x_i - компонента вектора входного сигнала ($i = 1, \ldots, n$);

z - результат суммирования;

$\varphi(z)$ - нелинейное преобразование;

y - выходной сигнал нейрона.

Как правило, рассматриваются формальные нейроны, функционирующие в непрерывном или дискретном времени. В этом случае, входной вектор, выходной вектор, а также вектор состояния зависят от параметра t. В дальнейшем, мы будем рассматривать нейроны, функционирующие в дискретном времени и, там где не возникает двусмысленности, для краткости, будем опускать параметр t.

Функции активации

Пусть функционирование формального нейрона описывается функцией

$$y = \varphi\left(\sum_{i=1}^{n} x_i\, s_i\right),$$

где функция φ - некоторая функция активации. Выражение в скобках – это скалярное произведение вектора весов нейрона на входной вектор. То есть его можно переписать так:

$$z = \sum_{i=1}^{n} x_i\, s_i = (\overline{x}, \overline{s}) = |\overline{x}||\overline{s}| \cos\alpha,$$

где α – угол между векторами $\overline{x}$ и $\overline{s}$.Так как после обучения $|\overline{s}| = \mathrm{const}$,

$$z = |\overline{s}| \cdot \overline{x}_s,$$

где $\overline{x}_s$ – проекция входного вектора на вектор весов. Это означает, что значение z будет максимальным при коллинеарности входного вектора и вектора весов.

Теперь рассмотрим возможные варианты функции активации нейрона. В общем случае, выбор функции активации ничем не ограничен. Однако при решении конкретных задач выбор вида функции ограничен исходными условиями задачи, алгоритмом обучения, вычислительными ресурсами ЭВМ и некоторыми другими факторами, о которых будет сказано в следующих разделах. В таблице 4.1 приведены некоторые, часто используемые виды функций активации.

Таблица 4.1. Виды функций активации.

Название	**Формула**	**Применение**	**Преимущества**	**Недостатки**
Линейная	$y = k \cdot z$	Линейная ассоциативная память, авторегрессионные модели.	Быстрое вычисление, непрерывность, монотонность.	Применима только в некоторых типах НС.
Пороговая	$y = sign(z)$	Персептроны, автоассоциативные НС, самоорганизующиеся НС.	Быстрое вычисление, нелинейность.	Разрыв в нулевой точке не позволяет дифференцировать, из-за чего неприменимы градиентные методы обучения.
Сигмоидальная	$y = \frac{1}{1 + e^{-cz}}$	Широкое.	Нелинейность, монотонность, дифференцируемость в области определения, быстрое вычисление производной.	Сравнительно невысокая скорость вычисления.
Гиперболический тангенс	$y = th(cz) = \frac{\sin cz}{\cos cz}$	Широкое.	Нелинейность, монотонность, дифференцируемость в области определения, быстрое вычисление производной.	Медленное вычисление.
Рациональная сигмоида	$y = \frac{z}{\lvert z \rvert + c}$	Широкое.	Нелинейность, монотонность, дифференцируемость в области определения, быстрое вычисление производной.	Существенных недостатков нет.
Гауссиан	$y = \exp\left(-\frac{z^2}{2\sigma^2}\right)$	Сети с радиально-бизисной функцией активации.	Нелинейность, дифференцируемость, быстрое вычисление производной.	Сравнительно невысокая скорость вычисления.

Связь двух нейронов

Определим связи между двумя нейронами. Два формальных нейрона $F^1 = (S^1, X^1, Y^1, g^1)$ и $F^2 = (S^2, X^2, Y^2, g^2)$ можно соединить следующим образом. Пусть i - номер компоненты выходного вектора $\overline{y}^1(t) \in Y^1$, а j - номер компоненты входного вектора $\overline{x}^2(t) \in X^2$, где $t \in \mathrm{N}$. Будем говорить, что i-я компонента выхода функционального элемента F^1 соединена с j-й компонентой входа функционального элемента F^2, если

$$\overline{x}_j^2(t) := \overline{y}_i^1(t). \tag{4.4}$$

Теперь предположим что у нас множество, состоящее из m нейронов - $E = \{F^i\}_{i=1}^m$. Связь между двумя любыми нейронами этого множества можно задать при помощи специальной матрицы. Пусть два нейрона $F^i = (X^i, Y^i, S^i, g^i)$ и $F^j = (X^j, Y^j, S^j, g^j)$, $1 \leq i, j \leq m$, связаны между собой каким либо образом. Пусть также размерности множеств равны $\dim X^i = N_X^i, \dim Y^i = N_Y^i, \dim S^i = N_S^i$ и $\dim X^j = N_X^j, \dim Y^j = N_Y^j, \dim S^j = N_S^j$ соответственно. Определим квадратную матрицу M_{ij} размерности $p_{ij} \times p_{ij}$, где $p_{ij} = \max(N_X^i, N_Y^j)$. Элементы этой матрицы M_{ij}^{kl} будут равны 1 только для тех k, l, для которых k-й выход i-го элемента равен l-му входу j-го элемента. В остальных случаях элементы матрицы равны нулю.

Это определение, а также некоторые другие понятия, приведённые дальше, соответствуют определениям аналогичных понятий в системном анализе. Такое определение связи удобно использовать, так как поступление выходного вектора одного элемента на вход другого элемента можно записать с помощью равенства:

$$\overline{x}^i(t) = \overline{y}^j(t) \cdot M_{ij}.$$

Пример. Пусть элементы F^1, F^2 соединены, так как это показано на рисунке.

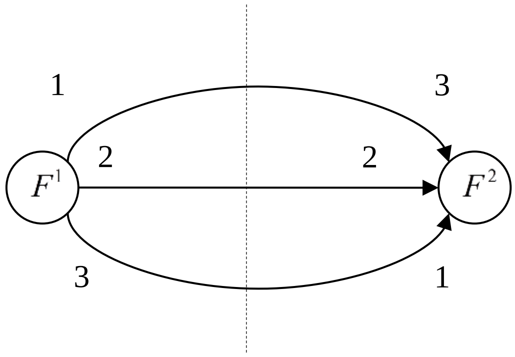

Рис. 4.6. ИНС из двух нейронов

Тогда матрица связи этих нейронов будет следующей:

$$M_{12} = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{Vmatrix}.$$

Для двух не связанных между собой элементов матрица $M_{ij} = 0$. Теперь можно рассмотреть матрицу M , элементами которой являются матрицы M_{ij}:

$$M = \left\| M_{ij} \right\|, \; i, j = \overline{1, m}. \qquad (4.5)$$

Такая матрица задаёт связи для всех элементов множества E. Мы будем называть эту матрицу *матрицей внутренних связей множества* E. Если не все матрицы M_{ij} являются матрицами одного порядка, они дополняются нулями справа и снизу там, где это нужно. Это соответствует добавлению «фиктивных» входов и выходов функциональным элементам множества E.

Для всех нейронов множества E мы можем рассматривать общий входной вектор $\overline{x} \in X^1 \times X^2 \times ... \times X^m$, общий вектор состояния $\overline{s} \in S^1 \times S^2 \times ... \times S^m$ и общий вектор выходов $\overline{y} \in Y^1 \times Y^2 \times ... Y^m$.

Теперь можно ввести понятие нейронной сети, как структуры, состоящей из простых элементов связанных между собой.

Определение 4.2. *Искусственная нейронная сеть (нейросеть, ИНС)* – структура, состоящая из связанных между собой нейронов, определяемая следующим образом

$$N = (E, G, M, H, \overline{x}_0, \overline{s}_0, T), \qquad (4.6)$$

где

$E = \{F^i\}_{i=1}^m$ - множество формальных нейронов;

$G(\overline{s},\overline{x})=(g^1(\overline{s}^1,\overline{x}^1),g^2(\overline{s}^2,\overline{x}^2),...,g^m(\overline{s}^m,\overline{x}^m))$ - вектор-функция сети определяющая общий выходной вектор по общему входному вектору и вектору состояния;

M – матрица внутренних связей множества E;

$H:(\overline{s},\overline{x})\mapsto\overline{s}\,'$ - некоторая функция, позволяющая изменять вектор состояния сети;

$\overline{x}_0,\overline{s}_0$ - начальные значения общего входного вектора и вектора состояния;

T - некоторый критерий останова функционирования нейронной сети (в простейшем случае – число тактов).

Большинство современных работ посвящено нейронным сетям, функционирующим в дискретном времени. Мы, в дальнейшем, будем иметь ввиду именно такие нейронные сети.

Пример. Рассмотрим ИНС состоящую из трёх нейронов. Матрицу связей определим следующим образом

$$M=\begin{pmatrix}0&0&0&0&0&1&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0\\1&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0\end{pmatrix}. \qquad (4.7)$$

Эта ИНС изображёна на рисунке 4.7 и имеет некоторую схожесть с изображением графа. Отличие в том, что некоторые дуги остаются как бы «повисшими в воздухе», т. е. не имеют вершинного окончания.

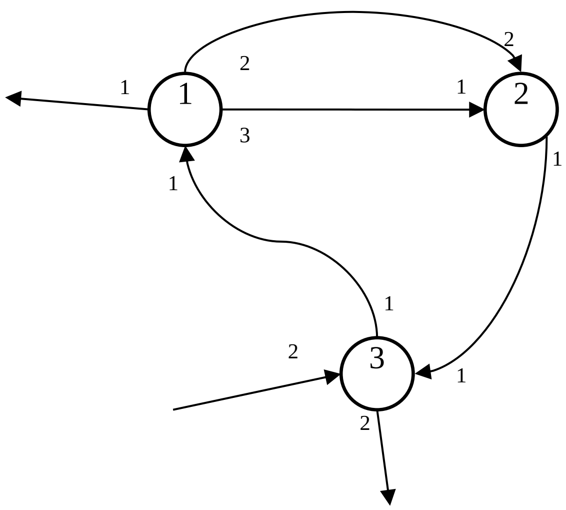

Рис. 4.7. Здесь входы и выходы пронумерованы своими номерами внутри каждого функционального элемента

Модель функционирования нейронной сети.

Функционирование нейронной сети происходит следующим образом:

- в момент времени $t = 0$, входной сигнал поступает на входные нейроны сети;
- в момент времени $t = 1$, сигнал, преобразованный входными нейронами, через связи (аксонов) передается на другие нейроны в соответствии со структурой связей и так далее;
- в некоторый момент $t = k$, сигнал поступает на выходные нейроны сети.

Результатом работы сети является сигнал, поступивший на выходные аксоны в момент $t = T$ (если критерий останова работы сети задан явно в виде количества тактов функционирования T).

Определение 4.3. Нейроны, часть синапсов которых являются входными, будем называть *входными нейронами сети*. Нейроны, аксон которых является выходным, будем называть *выходным нейроном сети*.

Из определения видно, что нейронная сеть вычисляет линейные функции, нелинейные функции одного переменного, а также всевозможные суперпозиции – функции от функций, получаемые при каскадном соединении сетей. Из этого, а также из теоремы 3.2 следует утверждение.

Утверждение 4.1. С помощью нейронной сети можно задать любую непрерывную функцию, с любой, заданной точностью.

Это утверждение позволяет надеяться, что для любой задачи, в основе решения которой лежит построение некоторой непрерывной функции, мы можем построить нейронную сеть, реализующую искомую функцию, и, следовательно, являющуюся решением задачи.

Лемма 4.1. Класс функций, вычислимый с помощью нейронных сетей, замкнут относительно линейных операций.

Доказательство. Действительно, пусть есть нейронные сети $S_1, S_2, \ldots, S_k$, которые вычисляют функции $F_1, F_2, \ldots, F_k$ от вектора входных сигналов x. Линейная комбинация $\alpha_0 + \alpha_1 F_1 + \alpha_2 F_2 + \ldots + \alpha_k F_k$ вычисляется сумматором (рис. 4.3) с весами $\alpha_0, \alpha_1, \ldots \alpha_k$, на вход которого подаются выходные сигналы сетей $S_1, S_2, \ldots, S_k$. Разница в числе тактов функционирования этих сетей до получения ответа легко компенсируется «линиями задержки», составленными из связей (рис. 4.2) с единичным весом. Лемма доказана.

Лемма 4.2. Класс функций, вычислимый с помощью нейронных сетей, замкнут относительно унарной операции, осуществляемой нелинейным преобразователем сигнала, входящим в состав нейрона (см. рис. 4.4, 4.6).

Доказательство. Если сеть S вычисляет функцию F, то, подавая выход этой сети на вход нелинейного преобразователя (рис. 4.4), получим на его выходе функцию $\varphi(F)$. Лемма доказана.

Таким образом, из теоремы 3.2, а также из лемм 4.1–4.2, следует теорема:

Теорема 4.1. Множество функций, вычислимых нейронными сетями с заданной непрерывной нелинейной функцией активации, плотно в пространстве непрерывных функций от входных сигналов.

Соединение двух ИНС

Рассмотрим ИНС следующего вида $N^1 = (E^1, G^1, {}_{11}M, H^1, \bar{x}_0^1, \bar{s}_0^1, T^1)$ и $N^2 = (E^2, G^2, {}_{22}M, H^2, \bar{x}_0^2, \bar{s}_0^2, T^2)$. Пусть в E^1 - m^1 элементов $F^i, i = \overline{1, m^1}$, а в E^2 - m^2 элементов $F^j, j = \overline{m^1+1, m^1+m^2}$. Матрицы внутренних связей - ${}_{11}M$ и ${}_{22}M$ соответственно.

Определим сеть $N = (E, G, M, H, \bar{x}_0, \bar{s}_0, T)$ следующим образом:

- множество нейронов $E := E^1 \times E^2$,
- начальные значения $\bar{x}_0 := (\bar{x}_0^1, \bar{x}_0^2)$, $\bar{s}_0 := (\bar{s}_0^1, \bar{s}_0^2)$.
- матрицу внутренних связей M сети N определим следующим образом:

$$M := \begin{Vmatrix} {}_{11}M_{1,1} & \cdots & {}_{11}M_{1,m^1} & {}_{12}M_{1,m^1+1} & \cdots & {}_{12}M_{1,m^1+m^2} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ {}_{11}M_{m^1,1} & \cdots & {}_{11}M_{m^1,m^1} & {}_{12}M_{m^1,m^1+1} & \cdots & {}_{12}M_{m^1,m^1+m^2} \\ {}_{21}M_{m^1+1,1} & \cdots & {}_{21}M_{m^1+1,m^1} & {}_{22}M_{m^1+1,m^1+1} & \cdots & {}_{22}M_{m^1+1,m^1+m^2} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ {}_{21}M_{m^1+m^2,1} & \cdots & {}_{21}M_{m^1+m^2,m^1} & {}_{22}M_{m^1+m^2,m^1+1} & \cdots & {}_{22}M_{m^1+m^2,m^1+m^2} \end{Vmatrix},$$

где ${}_{21}M$ - матрица, определяющая связь элементов ИНС N^2 с элементами N^1, а ${}_{12}M$ матрица, определяющая связь элементов ИНС N^1 с элементами N^2. Матрицу M можно представить в компактном виде:

$$M = \begin{Vmatrix} {}_{11}M & {}_{12}M \\ {}_{21}M & {}_{22}M \end{Vmatrix}.$$

- Операции H и G

$$H((\overline{s}^1,\overline{s}^2),(\overline{x}^1,\overline{x}^2)) := (H^1(\overline{s}^1,\overline{x}^1), H^2(\overline{s}^2,\overline{x}^2)),$$
$$G((\overline{s}^1,\overline{s}^2),(\overline{x}^1,\overline{x}^2)) := (G^1(\overline{s}^1,\overline{x}^1), G^2(\overline{s}^2,\overline{x}^2)).$$

- Количество тактов функционирования может определяться по разному, так как в общем случае T – это критерий останова работы сети. Но в простейшем случае параметр T может быть определен как $T := T^1 + T^2$, либо как $T := \max(T^1, T^2)$.

Таким образом, мы определили ИНС N, являющуюся соединением сетей N^1 и N^2, при помощи матриц внешних связей ${}_{12}M$ и ${}_{21}M$.

Замечание. Матрица M является матрицей внутренних связей для ИНС N. В то же время, элементы этой матрицы ${}_{12}M$ и ${}_{21}M$ являются матрицами внешних связей, а элементы ${}_{11}M$ и ${}_{22}M$ являются матрицами внутренних связей сетей N^1 и N^2 соответственно.

Пример. В качестве примера рассмотрим соединение сети из предыдущего примера с сетью N',изображённой на рисунке 4.8, при помощи матриц внешних связей

$${}_{12}M = \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Vmatrix}, \quad {}_{21}M = \begin{Vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{Vmatrix}.$$

В результате соединения этих сетей получим $N = (E, G, K, H, \overline{x}_0, \overline{s}_0, T)$.

Определим матрицу M:

$$M = \begin{Vmatrix} {}_{11}M & {}_{12}M \\ {}_{21}M & {}_{22}M \end{Vmatrix}.$$

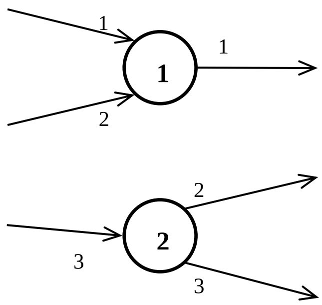

Рис. 4.8. Простая ИНС с тремя входами и тремя выходами

В результате получим ИНС изображённую на рисунке 4.9.

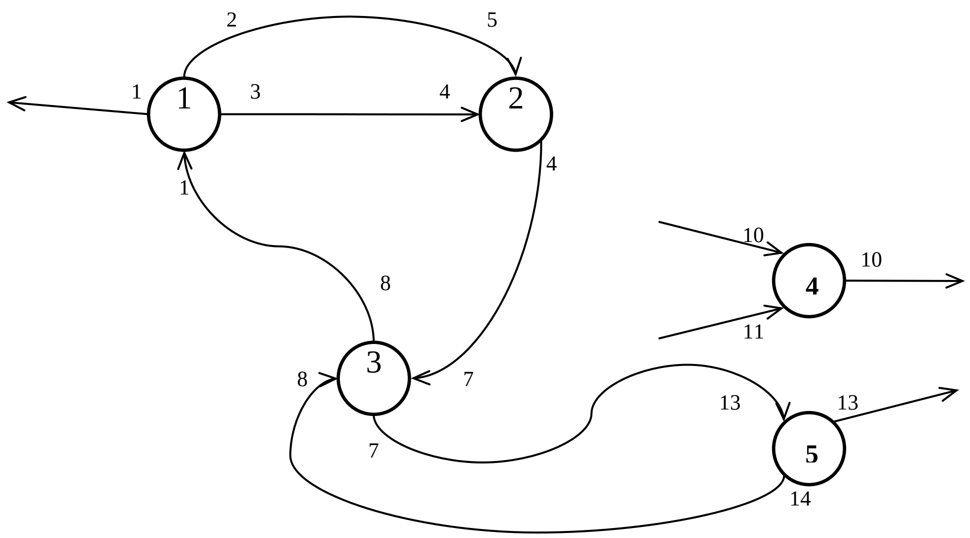

Рис. 4.9. ИНС полученная соединением

Многослойные ИНС

Аналогичным способом мы можем определить соединение сетей трёх и большего количества сетей. Например, если мы имеем три ИНС N^1, N^2, N^3 то, определив связи нейроподобных элементов с помощью матрицы

$$M = \begin{Vmatrix} {}_{11}M & {}_{21}M & {}_{31}M \\ {}_{21}M & {}_{22}M & {}_{23}M \\ {}_{31}M & {}_{32}M & {}_{33}M \end{Vmatrix},$$

мы получим соединение трёх сетей, которое также является ИНС.

Необходимо отметить особый класс сетей – так называемые *многослойные нейронные сети*. Такие сети будет удобно определить как соединение простых по структуре сетей. Для этого рассмотрим соединение сетей $\Pi = \{N^i\}_{i=1}^T$, $N^i = (E^i, G^i, {}_{ii}M, H^i, \bar{x}_0^i, \bar{s}_0^i, 1)$. В каждой из этих сетей матрица внутренних связей $M_{ii} = 0$, т. е. нейроны в такой ИНС не связаны между собой. Соединение сетей определим так, что для любого $1 \le i < T$, N^i связана посредством матрицы ${}_{ii+1}M$ с ИНС N^{i+1} и других внешних связей не существует. Матрица внутренних связей, полученной сети $N = (E, G, M, H, \bar{x}_0, \bar{s}_0, T)$ будет выглядеть следующим образом:

$$M = \begin{Vmatrix} 0 & {}_{12}M & 0 & \dots & 0 & 0 \\ 0 & 0 & {}_{23}M & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & {}_{T-1T}M \\ 0 & 0 & 0 & 0 & 0 & 0 \end{Vmatrix}.$$

Т. е. в этой матрице ненулевыми являются только элементы, расположенные непосредственно над элементами главной диагонали. Такую сеть будем называть многослойной ИНС [3,4]. В литературе нейронную сеть с такой внутренней структурой называют *многослойным персептроном*.

Если не оговорено противное, то каждый выходной сигнал i-го слоя подается на вход всех нейронов (i+1)-го. Число нейронов в каждом слое может быть любым и никак заранее не связано с количеством нейронов в других слоях. Стандартный способ подачи входных сигналов: все нейроны первого слоя получают каждый входной сигнал.

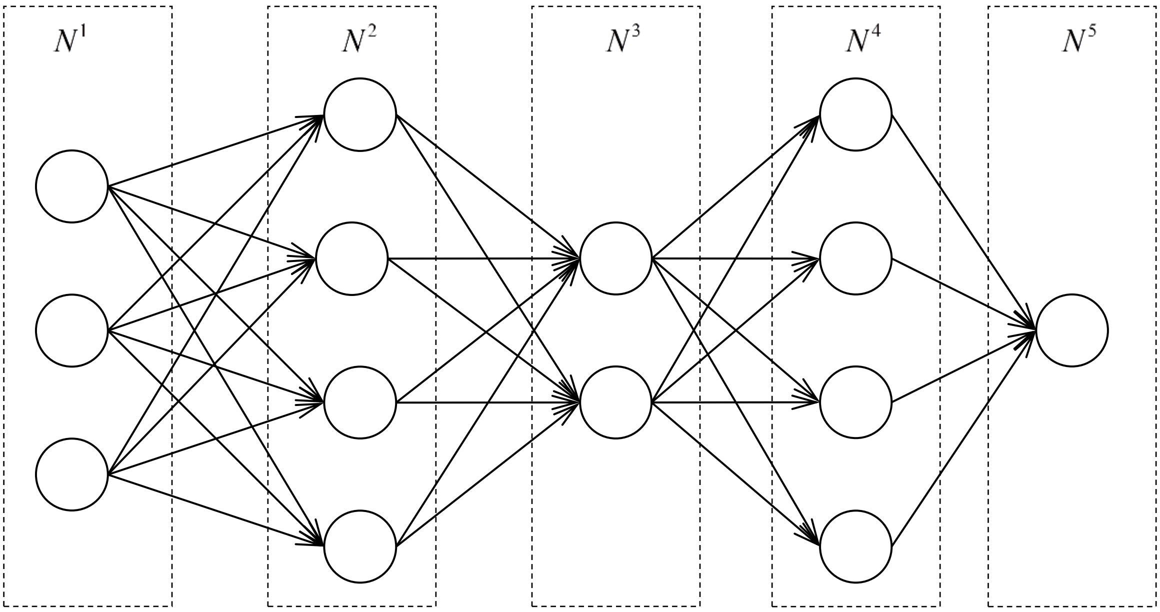

Рис 4.10. Многослойная ИНС являющаяся последовательным соединением пяти сетей

Прямое произведение ИНС

Еще один примечательный класс нейронных сетей – это ИНС, которые могут быть получены соединением с нулевыми матрицами внешних связей.

Рассмотрим множество сетей $\Pi = \{N^i\}_{i=1}^m$, произвольного вида $N^i = (E^i, G^i, {}_{ii}M, H^i, \bar{x}_0^i, \bar{s}_0^i, T^i)$. Матрицы внешних связей ${}_{ij}M = 0, i \neq j$. ИНС N, полученная в результате такого соединения этих сетей будет иметь такую матрицу внутренних связей

$$M = \begin{Vmatrix} {}_{11}M & 0 & 0 & \dots & 0 & 0 \\ 0 & {}_{22}M & 0 & \dots & 0 & 0 \\ 0 & 0 & {}_{33}M & \dots & 0 & 0 \\ \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & 0 & {}_{m-1m-1}M & 0 \\ 0 & 0 & 0 & 0 & 0 & {}_{mm}M \end{Vmatrix}.$$

В этой матрице ненулевыми являются только элементы, расположенные непосредственно на главной диагонали.

Условие для останова работы сети может быть определено как логическое произведение условий останова всех m сетей. В частности, когда критерий останова работы задаётся количеством тактов функционирования, формула может выглядеть так:

$$T := \max(T^1, T^2, ..., T^m).$$

Такая ИНС примечательна тем, что каждая из составляющих её сетей N^i функционирует независимо от остальных и, следовательно, результат работы N^i также не зависит от работы других сетей. Поэтому результат работы N состоит из отдельных результатов работы сетей N^i. ИНС имеющую описанные свойства будем называть *прямым произведением сетей* N^i. Если такую сеть рассматривать как автомат специального вида, то определение прямого произведения нейронных сетей становится эквивалентным определению прямого произведения автоматов, описанному, например, в [2].

Отметим также еще один класс ИНС – *полносвязные сети.*

Полносвязная сеть – нейронная сеть, каждый нейрон которой связан со всеми нейронами этой сети, включая самого себя. Все входные сигналы подаются всем нейронам. Выходными сигналами сети могут быть все или некоторые выходные сигналы нейронов после нескольких тактов функционирования сети.

Существенное различие между полносвязной и слоистой сетями возникает тогда, когда число тактов функционирования заранее не ограничено – слоистая сеть так работать не может.

Определение 4.6. Если функция активации нейронов φ - одна и та же для всех нейронов сети, то такая сеть называется *однородной*. Если же φ зависит еще от одного или нескольких параметров, значения которых меняются от нейрона к нейрону, то сеть называют *неоднородной.*

*О*пределения 4.4 и 4.5 дают два крайних случая архитектуры ИНС. Однако существуют и промежуточные варианты: ИНС, нейроны которой связаны не со всеми остальными нейронами сети, но и не в строгой послойной последовательности. Связи, порождающие рекуррентные зависимости выхода нейронов от значений выхода на предыдущих тактах функционирования, называются рекуррентными.

Определение 4.7. Нейронная сеть, с рекуррентными внутренними связями называется рекуррентной ИНС.

Выше были определены основные понятия нейроинформатики, используя которые можно строить практические модели формальных нейронных сетей.

Литература

1. Митрофанов Ю. И. Системный анализ: учеб. пособие. – Саратов: Изд-во «Научная книга», 2000. 232 с.
2. Богомолов А. М., Салий В. Н. Алгебраические основы теории дискретных систем. М., Наука, 1997.
3. Тархов Д. А. Нейронные сети. Модели и алгоритмы. – М.: изд. Радиотехника, 2005.
4. Горбань А.Н. Обучение нейронных сетей. – М.: изд. СССР-США СП "ПараГраф", 1990.
5. Головко В. А. Нейронные сети: обучение, организация и применение. – М.: ИПРЖР, 2001.