3. Структура технического нейрона

Впервые формализованная математическая модель нейрона была разработана У.С. Мак-Каллоком и У. Питтсом [66]. Мак-Каллок и Питтс предложили использовать в качестве модели нейрона бинарный пороговый элемент, вычисляющий взвешенную сумму входных сигналов и формирующий на выходе сигнал величины 1, если эта сумма превышает определенное пороговое значение, и 0 – в противном случае. К настоящему времени модель искусственного нейрона не претерпела существенных изменений, за исключением, быть может, введения различных типов активационных функций. Структурная схема искусственного нейрона представлена на рис. 1.2.

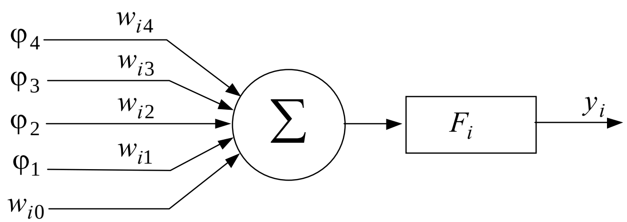

Рис. 1.2. Формальная модель искусственного нейрона

На вход искусственного нейрона поступает некоторое множество сигналов $\varphi_i\ (i=\overline{1,n})$, каждый из которых является выходом другого нейрона или входным сигналом нейросетевой модели. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе, все произведения суммируются, определяя уровень активации нейрона s. Данное преобразование с математической точки зрения эквивалентно скалярному произведению вектора входов φ и вектора весовых коэффициентов нейрона ω. Далее скалярный сигнал s преобразуется активационной (передаточной) функцией нейрона *F* в выходной сигнал *y*. Таким образом, формальный нейрон реализует отображение $R^n \to R^1$ в соответствии с соотношением:

$$y = F\left(\sum_{i=1}^{n} w_i \varphi_i + w_0\right) = F\left(\sum_{i=0}^{n} w_i \varphi_i\right), \qquad (1.1)$$

где $\varphi_i, i = \overline{1,n}$ – входы нейрона;

n – размерность вектора входов;

$w_i,\ i = \overline{1,n}$ – весовые коэффициенты нейрона, настраиваемые в процессе обучения;

w_0 – «нейронное смешение», вводимое для инициализации сети, – подключается к неизменяемому входу $\varphi_0 = +1$;

$F(*)$ – активационная функция нейрона;

Наибольшее распространение получили следующие активационные функции (рис. 1.3):

1) линейная (рис. 1.3*а*):

$$F(x) = k \cdot x\,; \qquad (1.2)$$

2) функция гиперболического тангенса (рис. 1.3*б*):

$$F(x) = \mathrm{th}(x) = (e^x - e^{-x})/(e^x + e^{-x})\,; \qquad (1.3)$$

3) сигмоидальная (рис. 1.3*в*):

$$F(x) = 1/(1 + e^{-x})\,; \qquad (1.4)$$

4) бинарные функции различного определения, например (рис. 1.3*г*):

$$F(x) = \text{sign}(x). \qquad (1.5)$$

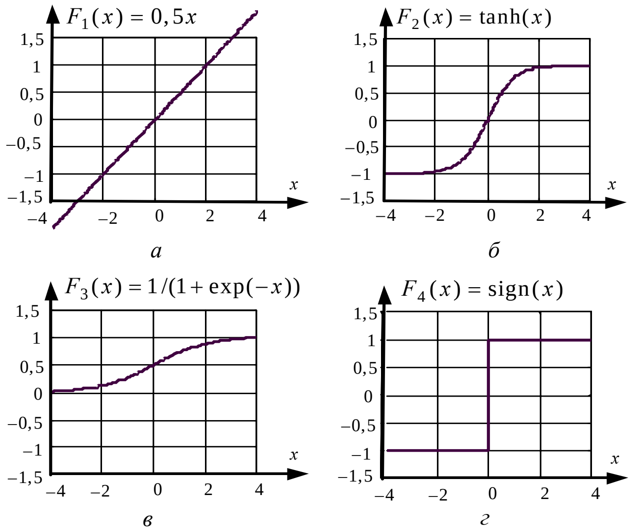

Рис. 1.3. Активационные функции искусственных нейронов: *а*) *линейная; б*) *гиперболический тангенс; в*) *сигмоидальная*; *г*) *пороговая* (*ступенчатая*)

Рассмотренная простая модель искусственного нейрона игнорирует многие свойства своего биологического прототипа. Например, она не принимает во внимание задержки по времени, которые воздействуют на динамику системы: входные сигналы сразу порождают выходной сигнал. Также не учитывается влияние функции частотной модуляции или синхронизирующей функции биологического нейрона, которые с биологических позиций считаются решающими.

Несмотря на эти несоответствия, сети, построенные из формальных нейронов, обнаруживают свойственные биологическим системам особенности. Более того, при подобранных соответствующим образом весовых коэффициентах совокупность параллельно функционирующих нейронов подобного типа способна выполнять универсальные вычисления.