

4. МНОГОСЛОЙНЫЕ НЕЙРОННЫЕ СЕТИ И ИХ АППРОКСИМИРУЮЩИЕ СВОЙСТВА

Нейроны могут группироваться в сетевую структуру различным образом. Функциональные особенности нейронов и способ их объединения в сетевую структуру обуславливают ту или иную парадигму нейронной сети. Для решения задач идентификации и управления наиболее адекватными, без сомнения, являются многослойные нейронные сети (МНС) прямого действия или многослойные перцептроны (МСП). При проектировании МНС нейроны объединяются в слои, каждый из которых обрабатывает вектор сигналов от предыдущего слоя (или входной вектор). Минимальной реализацией является двухслойная нейронная сеть, состоящая из входного (распределительного), промежуточного (скрытого) и выходного слоя. При подсчете числа слоев входной слой обычно не учитывается, так как служит лишь для распределения входных сигналов по нейронам последующего слоя.

Реализация модели двухслойной нейронной сети прямого действия имеет следующее математическое представление:

$$g_i(\theta) = \hat{y}_i(\theta) = \hat{y}_i(w, W) = F_i \left(\sum_{j=1}^{n_h} W_{ij} f_j \left(\sum_{l=1}^{n_\phi} w_{jl} \phi_l + w_{j0} \right) + W_{i0} \right), \quad (1.6)$$

где n_ϕ – размерность вектора входов ϕ нейронной сети;

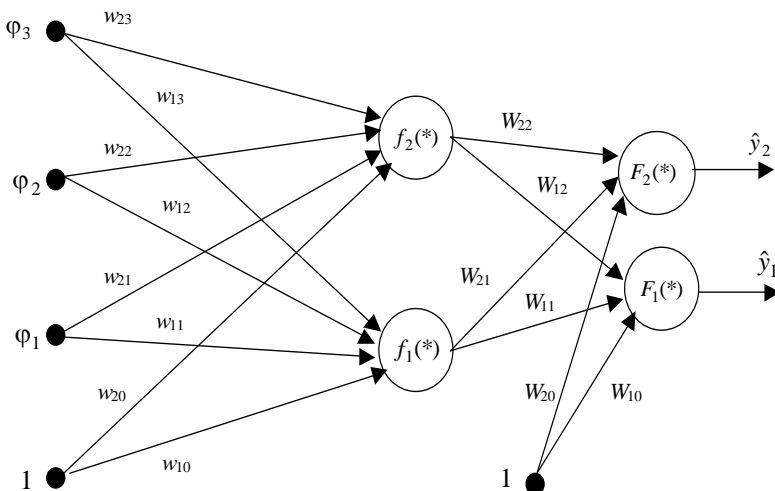
n_h – число нейронов в скрытом слое;

θ – вектор настраиваемых параметров нейронной сети, включающий весовые коэффициенты и нейронные смещения (w_{jl}, W_{ij}) ;

$f_j(x)$ – активационная функция нейронов скрытого слоя;

$F_i(x)$ – активационная функция нейронов выходного слоя.

На рис. 1.4 представлена структурная схема двухслойной НС прямого действия. Сигналы в сети распространяются от входа к выходу, связи между нейронами одного слоя и обратные связи отсутствуют.



**Рис. 1.4. Структурная схема двухслойной нейронной сети
(число входов – 3; выходов – 2)**

Необходимо показать, что МНС, имеющая математическое представление в форме (1.6), при условии соответствующего выбора активационных функций и весовых коэффициентов может быть использована в качестве модельной структуры для решения задачи идентификации. Предположим, что дискретная динамическая система может быть представлена как некоторая функция (в общем случае, нелинейная) от предыдущих значений входов u и выходов y :

$$y(t) = \bar{f}(y(t-1), \dots, y(t-n), u(t-1), \dots, u(t-m)). \quad (1.7)$$

Естественно предположить, что МНС может аппроксимировать функцию (1.7) при условии, что в качестве вектора входов сети φ выбираются n предыдущих значений выходов системы и m предыдущих входов.

Рассмотрим функционирование МНС как совокупности взаимосвязанных элементарных элементов (нейронов) с математической точки зрения. Каждый структурный элемент МНС получает на входе вектор сигналов φ , вычисляет его скалярное произведение на вектор весовых коэффициентов нейрона ω и некоторую функцию F в выходной сигнал y . Результат поступает на входы других нейронов или на выход. Таким образом, нейронные сети вычисляют суперпозиции функций одного переменного и их линейные комбинации. Для обоснования возможности использования МНС в качестве моделей динамических систем нужно получить ответ на вопрос: можно ли произвольную непрерывную функцию n переменных получить с помощью операций сложения, умножения и суперпозиции функций одного переменного?

В серии работ А.Н. Колмогорова и В.И. Арнольда решена следующая математическая проблема (составляющая существо тринадцатой проблемы Гильберта): любую непрерывную функцию n переменных можно получить с помощью операций сложения, умножения и суперпозиции из непрерывных функций одного переменного. На основе этих работ (суть которых изложена в [4]) доказан ряд теорем [27, 33, 39, 47] об аппроксимации непрерывных функций многих переменных нейронными сетями с использованием практически произвольной функции одного переменного. Помимо подтверждения общих аппрок-

симирующих свойств МНС необходимо получить ответы на ряд частных вопросов, касающихся структуры сети:

- сколько скрытых слоев должна содержать нейронная сеть?
- сколько нейронов должно быть включено в каждый слой?
- какой тип активационной функции должен быть выбран?

В работе [33] показано, что любая непрерывная нелинейная функция может быть аппроксимирована с достаточной точностью нейронной сетью с одним скрытым слоем, содержащим нейроны с сигмоидальными (или типа «гиперболический тангенс») функциями активации, и выходным слоем, содержащим нейроны с линейной активационной функцией. Попытка исследования влияния числа нейронов в скрытом слое на аппроксимирующие свойства сети сделана в работе [27], однако полученный результат практически невозможно применить на практике.

Тем не менее результаты исследований, представленные в работах [27, 33, 39, 47], подтверждают универсальные аппроксимирующие свойства нейронных сетей, что позволяет сделать вывод о возможности использования МНС в качестве модельных структур при реализации процедуры идентификации.

В настоящей работе рассматривается минимальная реализация МНС в соответствии с выражением (1.6) и активационными функциями типа «гиперболический тангенс» (1.2) для нейронов в скрытом слое и линейными активационными функциями (1.3) нейронов выходного слоя. Возможно, репрезентативные способности МНС могут быть улучшены путем введения дополнительных скрытых слоев, особенно в случае моделирования сложных взаимосвязей. Однако усложнение

структуры нейросети приводит к значительным трудностям при практической реализации, параметрической оптимизации (обучении) и последующем анализе МНС. Это объясняет факт использования именно минимальной реализации МНС в большинстве технических приложений.