Лекция 7.

Тема: “**Динамические рекурсивные ИНС**”

Отличие рассматриваемых в данном разделе сетей от сетей Хопфилда состоит в том, что

1. Рекурсивная сеть может иметь несколько слоев.

2. Нейроны могут иметь собственные обратные связи ($w_{ij} \neq 0$)*.*

3. Матрица весов может быть несимметричной.

4. Может осуществляться контролируемое обучение путем использования алгоритма обратного распространения.

Наличие обратных связей между нейронами различных слоев, включая и нейроны выходного слоя, обеспечивает *динамическим рекурсивным сетям* (ДРС) дополнительные положительные свойства, которые не могут быть достигнуты в статических многослойных сетях прямого распространения. К числу таких свойств относится, например, возможность работы с образами, параметры которых изменяются во времени.

8.1 Структура ДРС

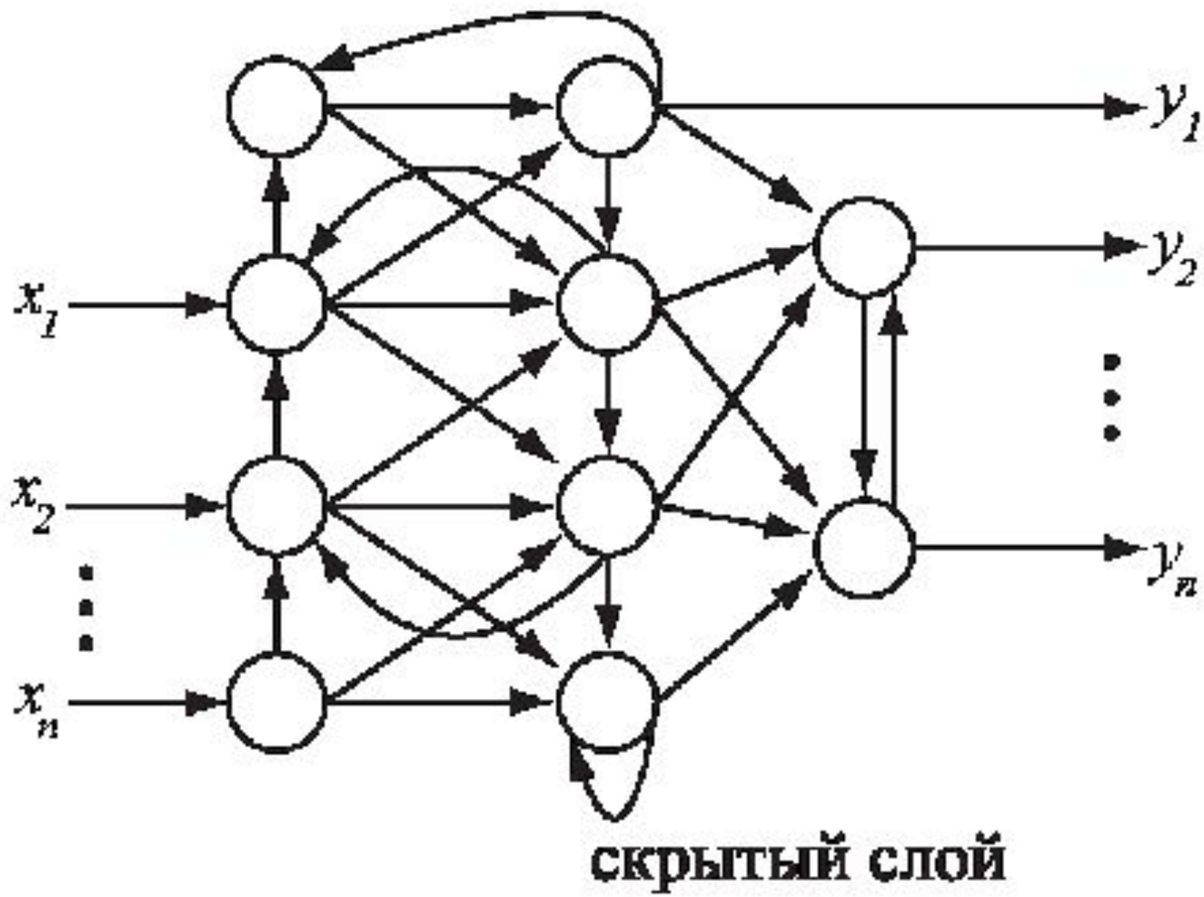

Рис.8.1 – Пример структуры ДРС

В данной сети выходными нейронами могут быть любые (например, на рис.8.1 один нейрон скрытого слоя является также выходным).

Сигналы, поступающие на входы нейронов входного слоя в некоторый момент времени t>0, преобразуется последними с помощью соответствующих функций активации $f_a(x,w)$ и в следующий момент времени передаются по имеющимся связям нейронам, которые их также преобразуют. Далее преобразованный сигнал по прямым и обратным связям поступает на входы нейронов и процесс повторяется. Наличие обратных связей приводит к тому, что в зависимости от значений входного сигнала и весовых параметров сеть может *а) достичь некоторого устойчивого состояния; б) осциллировать, т.е. периодически повторять значения выходных сигналов; в) хаотически изменять свое состояние*. Таким образом, динамика ДРС аналогична динамике сети Хопфилда (см. раздел 5).

Наличие обратных связей в данной сети не позволяет использовать для ее описания столь простые соотношения, которые применялись в рассмотренных ранее сетях. Поэтому их динамика, как и динамика сетей Хопфилда, являющихся частным случаем ДРС, описывается нелинейными дифференциальными (в непрерывном случае) или разностным (в дискретном случае) уравнениями первого порядка.

8.2 Непрерывные ДРС

Динамика i-го нейрона непрерывной ДРС описывается уравнениями, аналогичными уравнению (5.16),

$$k_i \frac{dx_i(t)}{dt} = -x_i(t) + f_i\left(\sum_{j=1}^{M} w_{ij} x_j(t) + v_i\right), \quad i = \overline{1, M}, \tag{8.1}$$

где k_i - постоянная времени i-го нейрона; $x_i(t)$ - состояние i-го нейрона в момент времени t; $f_i(t)$ - нелинейная функция активации; v_i - внешний входной сигнал i-го нейрона; M - количество нейронов в сети.

Элементы w_{ij} весовой матрицы W определяются путем решения уравнений

$$\frac{dx(t)}{dt} = 0. \tag{8.2}$$

В зависимости от вида матрицы W различают три типа сетей:

- симметричная матрица весов с нулевыми диагональными элементами описывает ДРС типа сети Хопфилда;
- треугольная матрица W характеризует ДРС прямого распространения без обратных связей;
- весовая матрица произвольного вида характеризует ДРС общего вида.

В последнем случае сеть в зависимости от значений ее параметров может либо достигать некоторого устойчивого состояния, либо осциллировать, либо хаотически изменять свое состояние. В частности, для достижения устойчивого состояния весовые коэффициенты w_{ij} должны удовлетворять следующему условию:

$$\sum_i \sum_j w_{ij}^2 < (\max_i |f_i'|)^{-2}, \qquad (8.3)$$

где $f_i' = \dfrac{df}{dx}$ - производная функции активации.

8.3 Дискретная ДРС

Уравнение, описывающее динамику дискретной ДРС, получается из (8.1) путем перехода к конечным разностям, а элементы весовой матрицы являются решением системы уравнений

$$x_i(k+1) = f\left(\sum_{j=1}^{M} w_{ij} x_j(k) + v_i\right). \tag{8.4}$$

При $\Delta t \to 0$ поведение непрерывной и дискретной ДРС идентично.

В общем же случае поведение обеих сетей даже при одинаковых значениях элементов весовой матрицы различно (например, если одна достигает устойчивого состояния, то другая может осциллировать).

От описания (8.4) можно перейти к описанию в пространстве состояний вида

$$\begin{aligned} z(k+1) &= f(z(k), x(k)), \\ y(k) &= g(z(k), x(k)), \end{aligned} \tag{8.5}$$

где $z(i)$ - состояние сети в момент времени i.

8.3.1 Полносвязные ДРС

Структура полносвязной ДРС приведена на рис.8.2.
ДРС данного типа изучались в работах [ ].

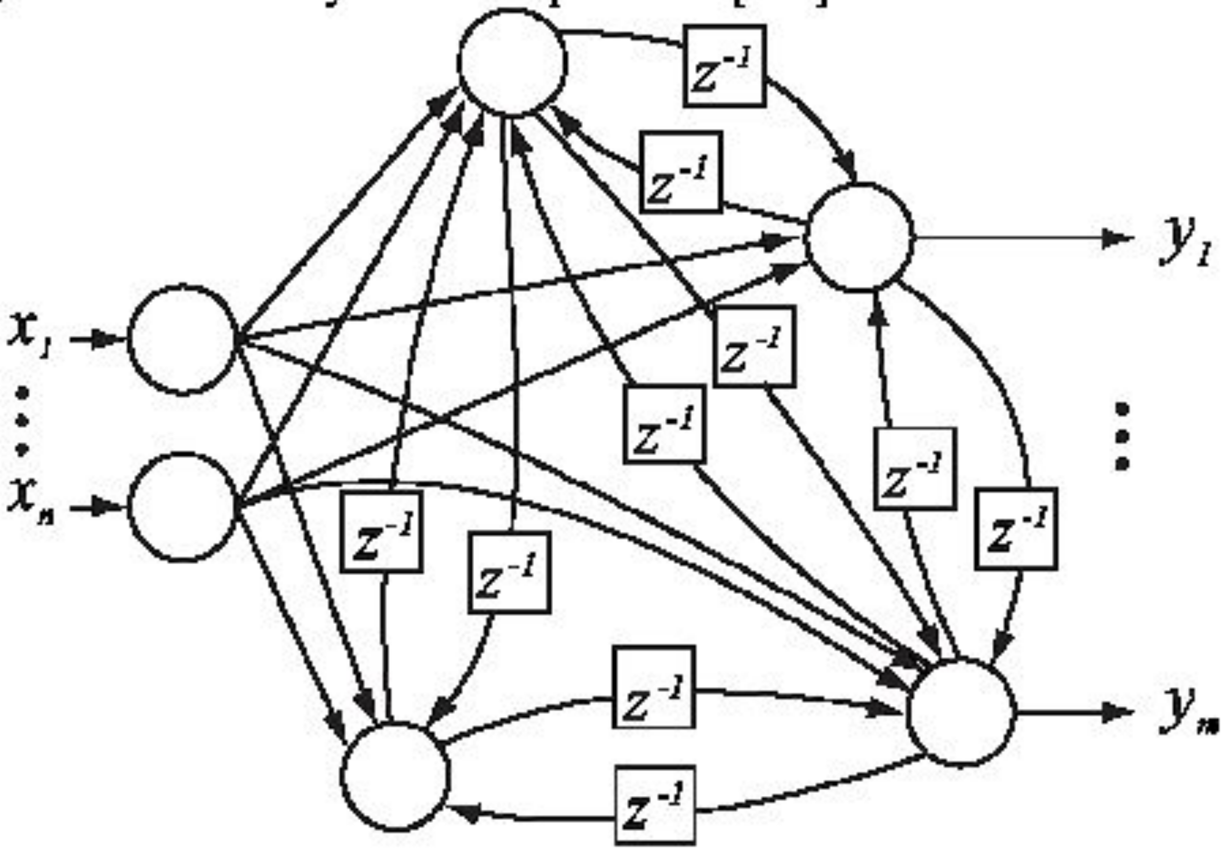

Рис.8.2 – Полносвязная ДРС

Данная архитектура была первоначально предложена для решения задач, связанных с анализом и обработкой последовательностей, но в последствии была использована также для идентификации нелинейных динамических объектов. Однако данной сети свойственен серьезный недостаток – медленная сходимость (существенная длительность процесса обучения) и возникающие при этом проблемы устойчивости [ ].

8.3.2 Частично-рекурсивные сети

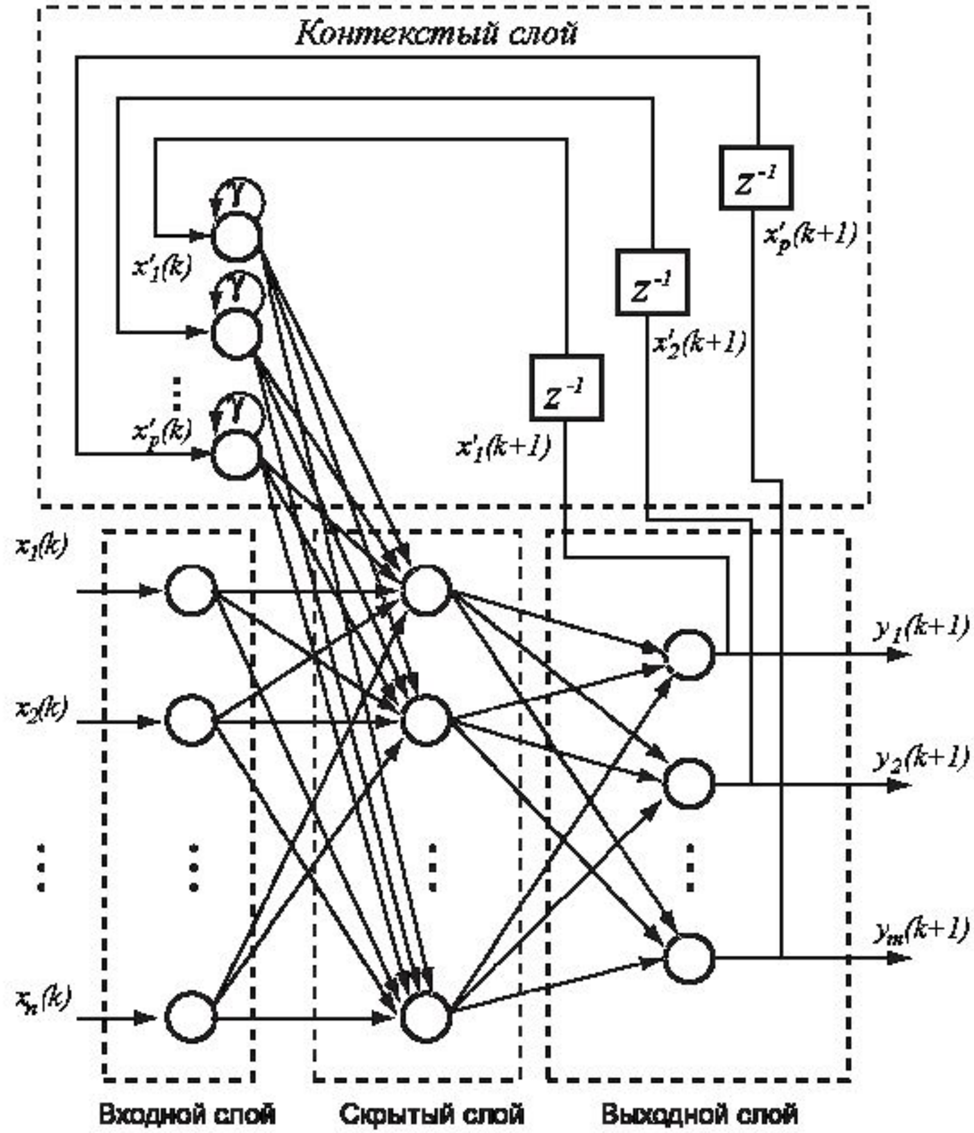

Рис.8.3 – Сеть Джордана

Выходной сигнал сети в каждый момент времени зависит от значения входного сигнала и состояния, запомненного в контекстном слое. Поэтому сеть описывается уравнениями (8.5)

$$z(k+1) = f(z(k), x(k));$$
$$y(k) = g(z(k), x(k)),$$

где $z(k)$ - состояние в момент времени k; g,f -функции, определяющие соответственно выход и состояние сети на следующем такте (в теории цифровых автоматов их называют функциями выходов и переходов соответственно).

То, что состояние $z(k+1)$ зависит от $z(k)$ и $x(k)$, видно из топологии сети

$$z(k+1) = \hat{f}(z(k), y(k)) = \hat{f}(z(k), g(z(k), x(k)) = f(z(k), x(k)).$$

Последовательное изменение состояний нейронов при некотором начальном состоянии z_0 описывается формулой

$$z(k) = \begin{cases} z_0, & \text{если } k = 1; \\ \gamma z(k-1) + \lambda y(k-1), & \text{если } k > 1, \end{cases} \tag{8.6}$$

т.е.

$$z(k) = \gamma^{k-1} z_0 + \lambda \sum_{n=1}^{k-1} \gamma^{n-1} y(k-n), \tag{8.7}$$

где $\lambda \in (0,1]$ - весовой параметр обратной связи (обычно $\lambda = 1$).

Если $z_0 = 0$ и $\lambda = 1$, то

$$Z(k) = \sum_{n=1}^{k-1} \gamma^{n-1} y(t-n), \tag{8.8}$$

т. е. состояние представляет собой экспоненциально взвешенную сумму всех имевшихся до этого времени выходных сигналов.

При малых значениях γ влияние прошлых состояний мало, а при $\gamma \to 1$ велико. При $\gamma = 1$ все состояния имеют одинаковый вес (использование параметра $\gamma < 1$ с одной стороны уменьшает влияние ранее поступивших образов на текущий, а с другой- позволяет учесть это влияние). Обычно принимают $\gamma \approx 0{,}5$.

Во время обучения сети веса γ_i не изменяются. Хотя принципиально изменение γ возможно, как показывают исследования, это лишь приводит к затягиванию процесса обучения без существенного улучшения работы ИНС. Однако вопрос выбора γ, малые значения которого позволяет лучше реагировать на вновь поступающий образ, а большие - обеспечивать запоминание всех пришедших ранее образов, остается открытым.

Сеть Джордана способна ассоциировать различные входные образы с различными выходными последовательностями.

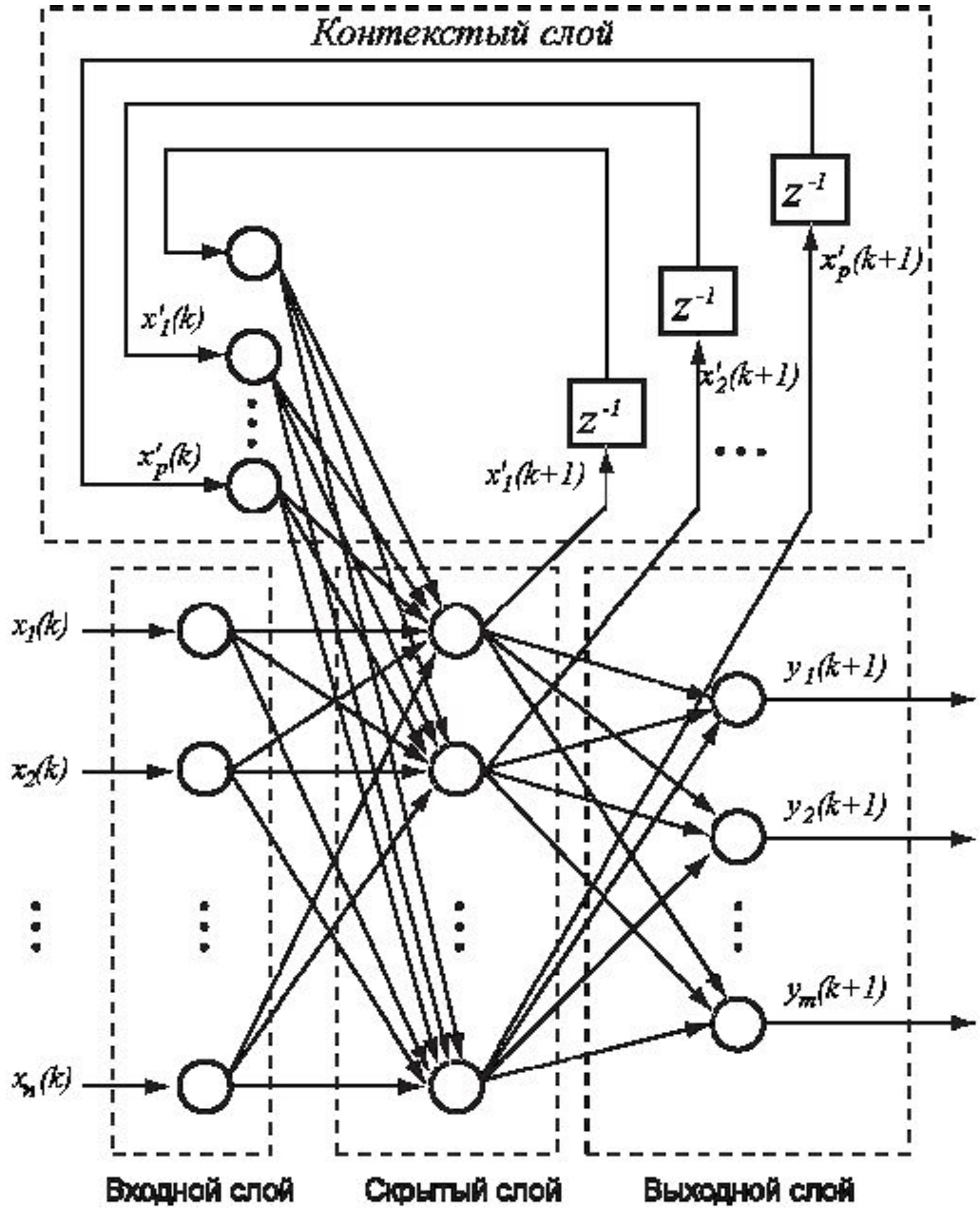

Рис.8.4 – Сеть Элмана

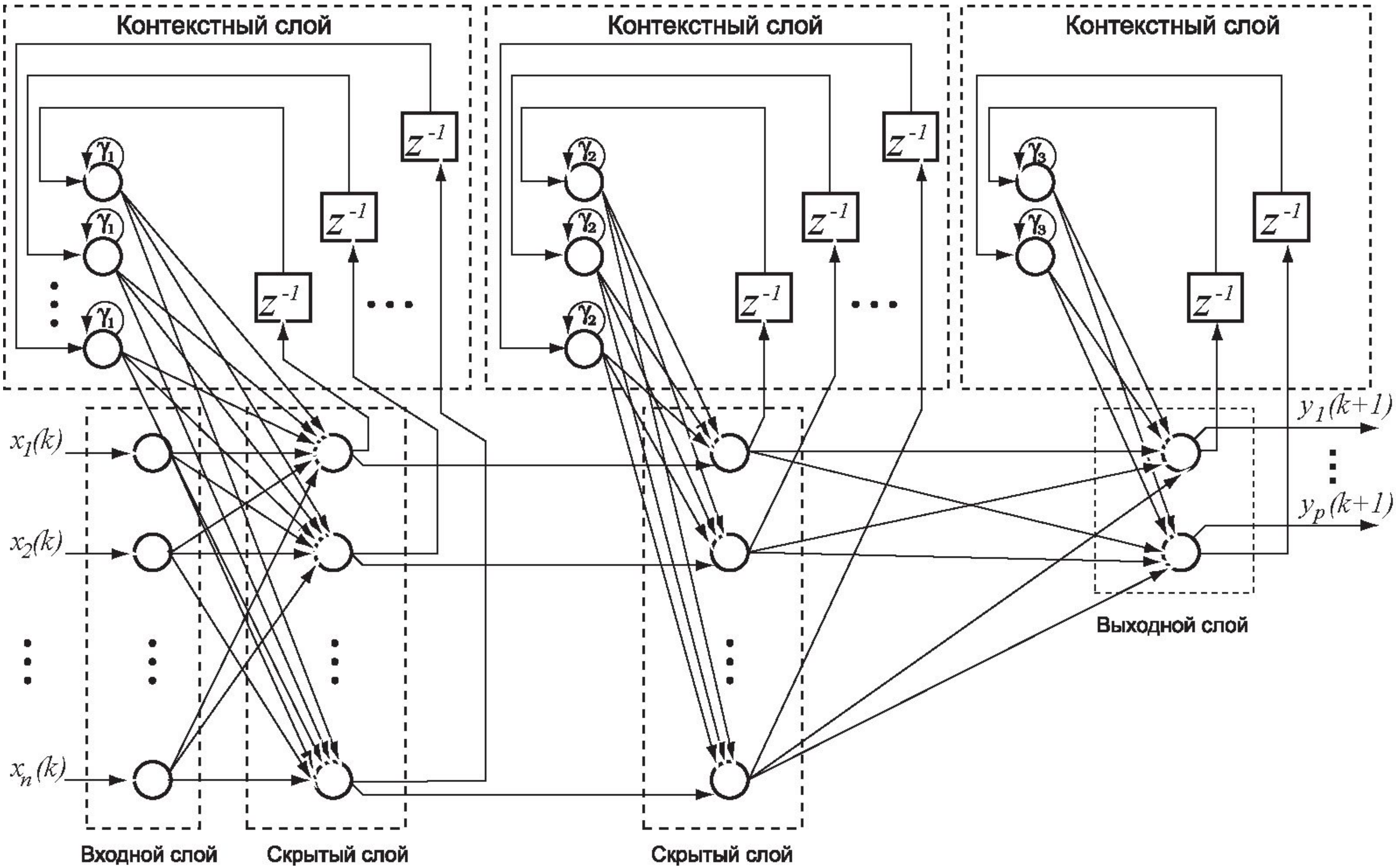

Рис.8.5 – Иерархическая частично-рекурсивная сеть

8.3.3 Локально-рекурсивные сети прямого распространения

Сети данного типа не используют ни обратную связь между нейронами соседних слоев, ни латеральные связи между нейронами одного слоя. Рекурсивность в них всегда ограничивается одним нейроном. Получаемые при этом структуры являются линейными, а вводимые обратные связи интерпретируются как фильтры с конечной или бесконечной импульсной характеристикой (КИХ или БИХ). Как показано на рис.8.6, существует три разных способа получения локальной рекурсивности или, другими словами, локального введения в сеть динамики:

- динамика синапсов, использующая локальные обратные связи нейронов;
- «активационная» динамика, являющаяся частным случаем динамики синапсов и позволяющая в случае одинаковых передаточных функций всех синапсов одного нейрона существенно упростить структуру сети; в этом случае фильтры, используемые в синапсах, могут быть заменены одним фильтром, имеющим те же нули и полюсы и стоящим после операции суммирования [ ];
- динамика обратной связи, реализуемая путем введения линейной обратной связи с выхода нейрона на его вход.

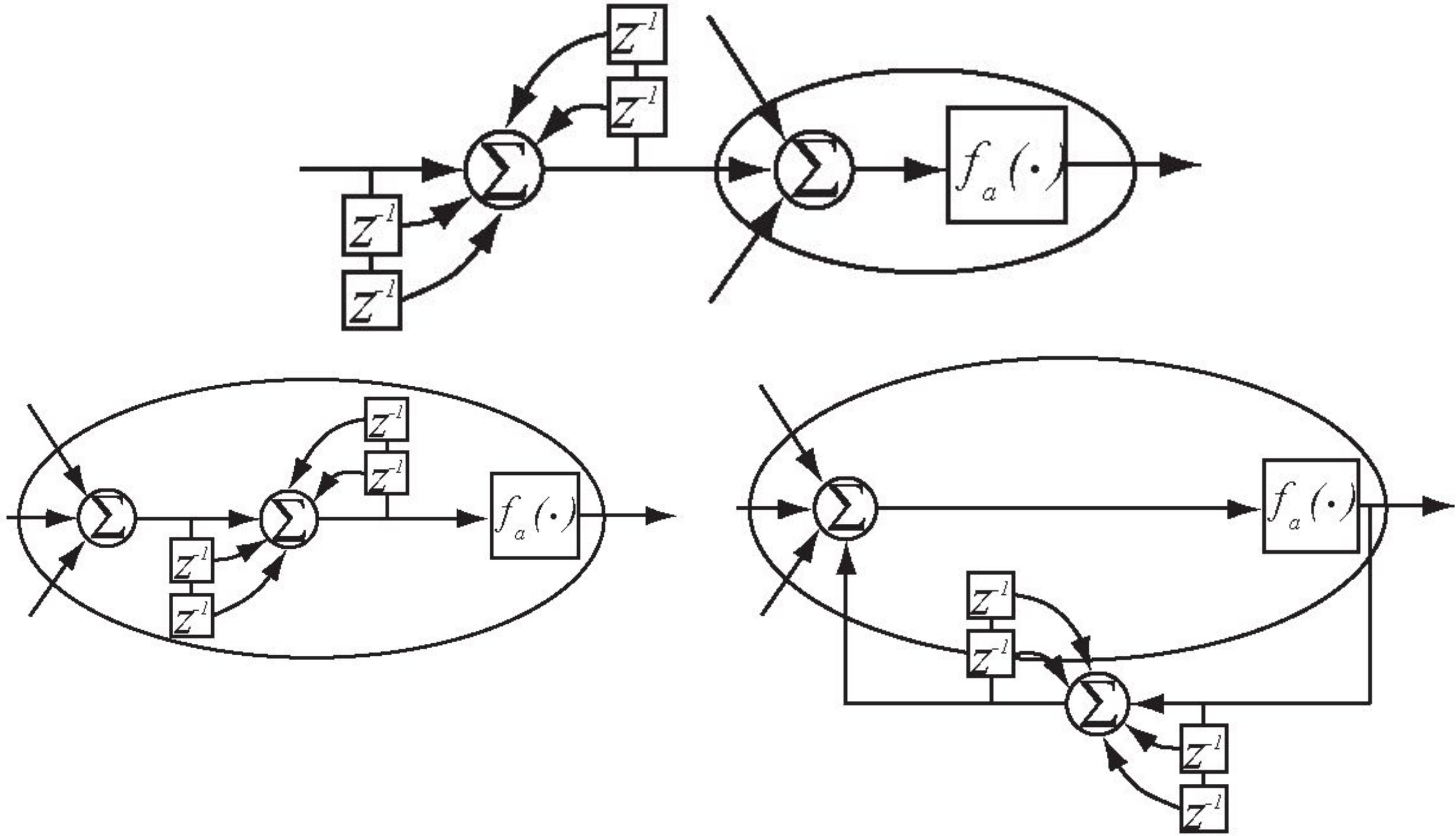

Рис.8.6 – Типы локальной рекурсивности

8.4 Обучение ДРС

Так как ДРС описываются нелинейными уравнениями, для настройки параметров этих сетей используют методы нелинейной оптимизации, среди которых чаще всего применяют градиентные методы. Наличие в ДРС обратных связей приводит к тому, что значение градиента зависит от прошлых состояний сети. В связи с этим различают два подхода к решению задачи обучения сети: применение *алгоритма обратного распространения ошибки* и применение *адаптивных алгоритмов*, в основе которых лежат рекуррентные процедуры. Оба эти подхода используют градиентные схемы минимизации, т.е. представляют собой обычно градиентные методы первого порядка (хотя могут быть использованы и методы более высоких порядков), отличающиеся количеством используемой и хранимой в памяти информации.

8.4.1 Алгоритм обратного распространения ошибки

Этот алгоритм был предложен *Д.Румельхардтом*, *Г.Хинтоном* и *Р.Уильямсом* в работе [ ] и использует представление ДРС в виде многослойной сети прямого распространения (МСПР), у которой на каждом такте работы происходит увеличение слоев на единицу. Каждый же слой МСПР состоит из того же количества нейронов, имеющих те же связи, что и исходная ДРС. На рис. 8.7 б) приведена эквивалентная полносвязной ДРС, содержащей два нейрона (рис. 8.7 а), развивающаяся многослойная сеть прямого распространения.

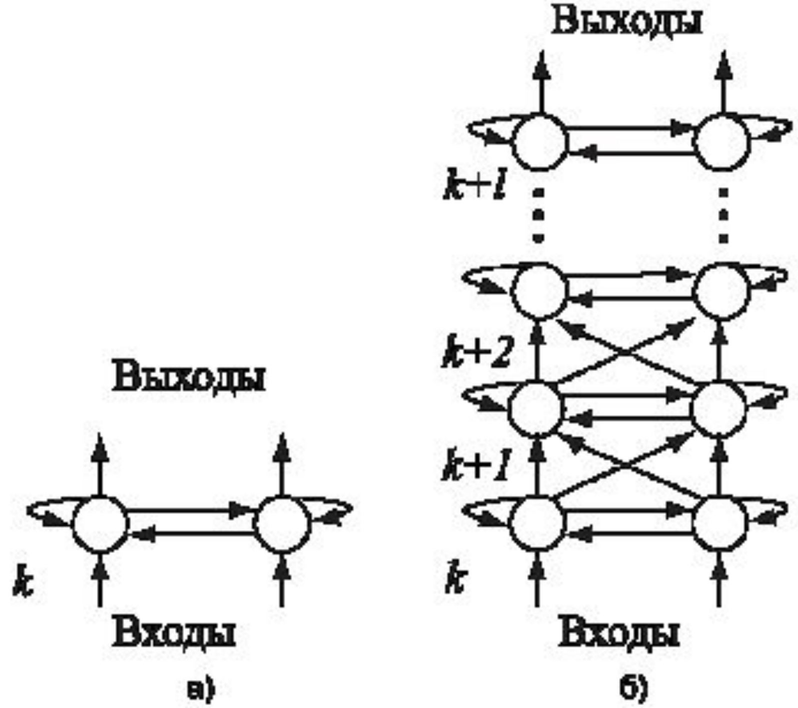

Рис.8.7 – Простая полносвязная ДРС (а); эквивалентная, развивающаяся во времени МСПР(б)

Первый слой такой сети соответствует состоянию ДРС в момент времени $k = 1$.

Очевидно, для обучения сети может быть применен алгоритм обратного распространения ошибки, использующий в любой момент времени $k \quad (k = 1,2,\ldots)$ вычисления градиента

$$\frac{\partial I(k)}{\partial w} = g(w, z(k), x(k)), \qquad (8.9)$$

где $\mathbf{w}$ - искомые параметры сети; $\mathbf{z}(k)$ - матрица состояний сети размерности $n \times k$ (n - число состояний сети, k - количество предшествующих тактов); $X(k)$ - матрица входных сигналов $N \times k$ (N - число входов).

Из (8.9) видно, что для реализации алгоритма настройки весов $\mathbf{w}$ необходимо запоминание большого обьема информации, растущего с увеличением количества тактов обучения k. В связи с этим чаще применяют модификации алгоритма обратного распространения, заключающиеся в распространении выходной ошибки не на все слои, начиная с выходного, а на ограниченное их число.

8.4.2 Адаптивный алгоритм обучения

Данный алгоритм был предложен и исследован *Р.Уильямсом* и *Д.Ципсером* [ ] и использует вычисление градиента в момент времени k на основе частных производных состояний нейронов, вычисленных в момент времени $k-1$, т.е.

$$\frac{\partial I(k)}{\partial w} = g(w, \frac{\partial z(k-1)}{\partial w}, z(k-1), x(k)) \qquad (8.10)$$

В этом случае обьем требуемой памяти существенно уменьшается и зависит только от числа нейронов в сети и количества настраиваемых параметров. Хотя и здесь необходимо принимать во внимание все предшествующие состояния сети, в данном алгоритме происходит учет информации лишь об одном последнем такте, т.е. входящие в (8.10) $\mathbf{z}(k-1)$ и $\mathbf{x}(k)$ имеют размерности $n\times 1$ и $N\times 1$ соответственно.

При использовании данного алгоритма в реальном времени предполагают, что

$$\frac{\partial \mathbf{z}(k-1)}{\partial \mathbf{w}} \approx \frac{\partial \mathbf{z}(k-1)}{\partial \mathbf{w}(k-1)}, \qquad (8.11)$$

что и учитывается в потактовой коррекции весов.