

Лекция 2.

Тема: “Основные понятия искусственных
нейронных сетей”

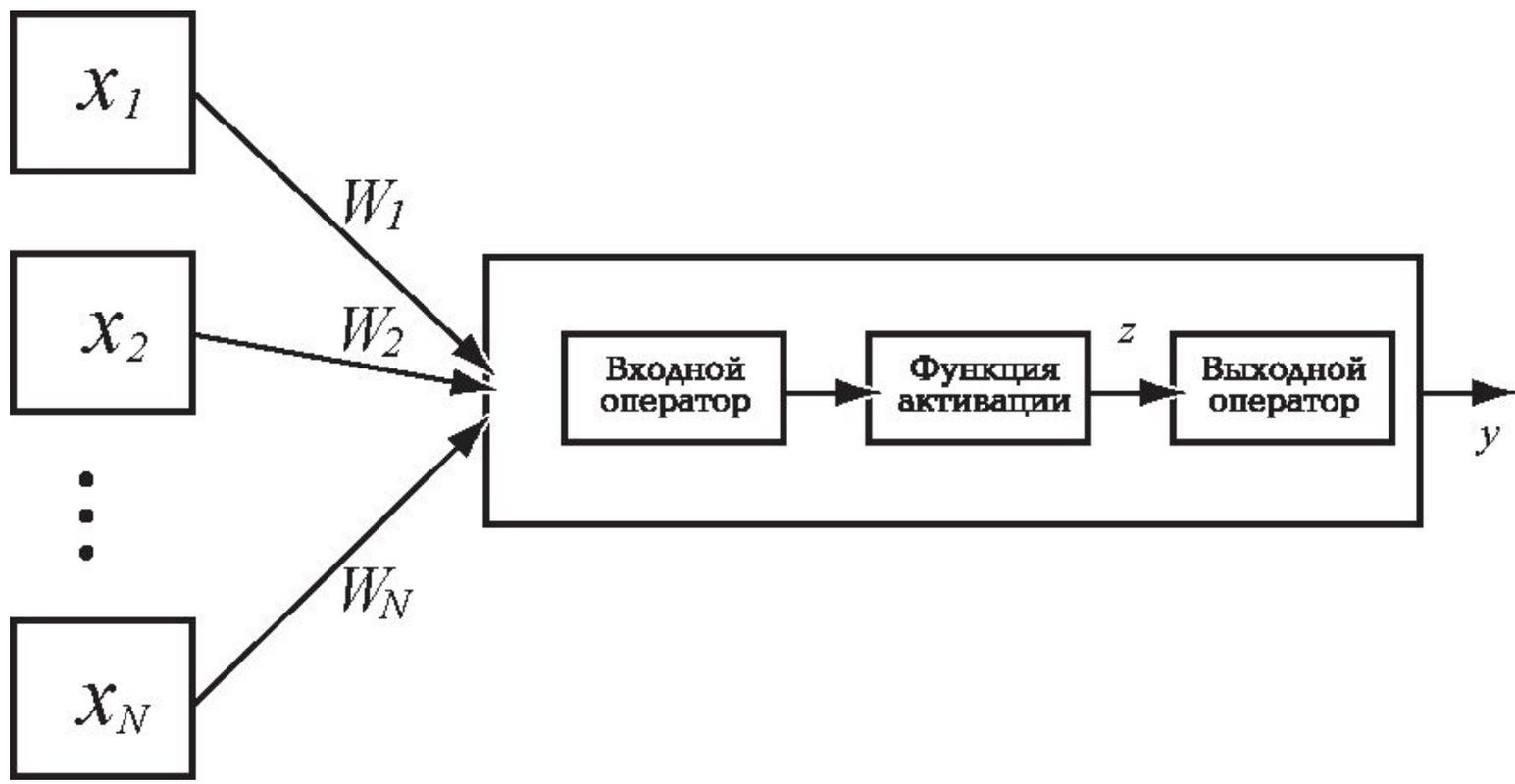


Рис.2.1 – Структура искусственного нейрона

Входной оператор

- сумма взвешенных входов

$$f(x, w) = \sum_{i=1} w_i x_i ;$$

- максимальное значение взвешенных входов

$$f(x, w) = \max_i (w_i x_i) ;$$

- произведение взвешенных входов

$$f(x, w) = \prod_{i=1}^N w_i x_i ;$$

- минимальное значение взвешенных входов

$$f(x, w) = \min_i (w_i x_i) .$$

Функция активации

- линейная

$$f(z) = K z, \quad K = \text{const},$$

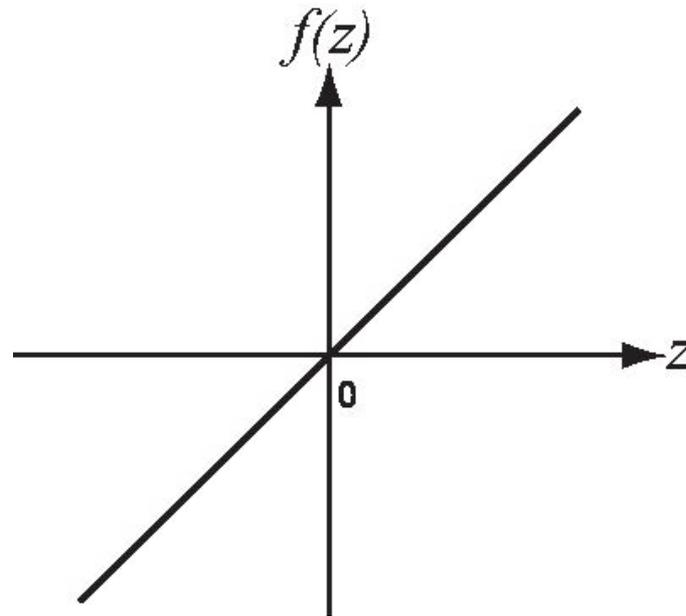


Рис.2.2 – Линейная функция

- линейная биполярная с насыщением

$$f(z) = \begin{cases} 1 & \text{при } z > \alpha_2; \\ kz & \text{при } -\alpha_1 \leq z \leq \alpha_2; \\ -1 & \text{при } z < -\alpha_1. \end{cases}$$

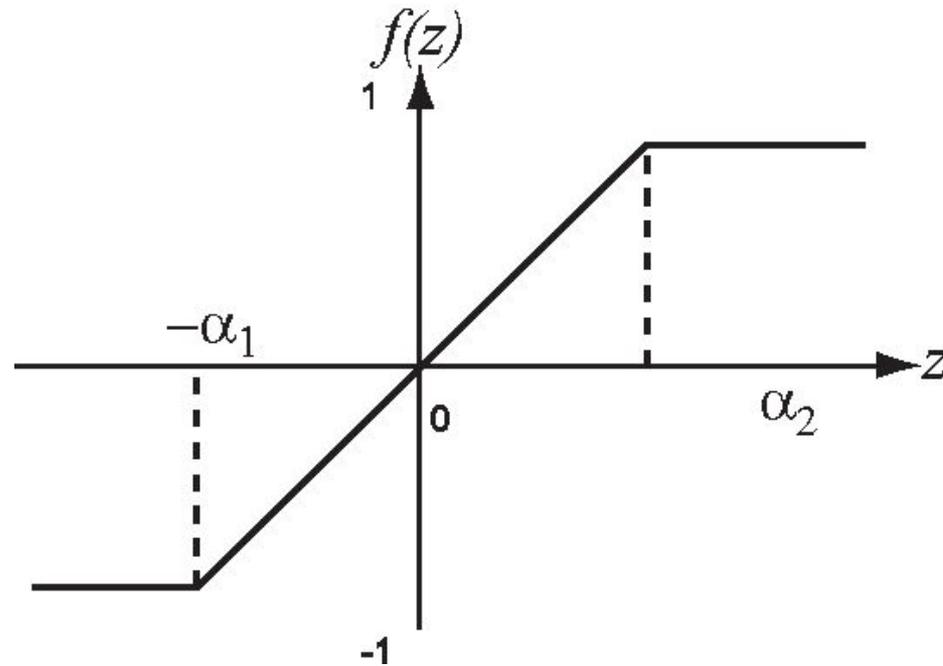


Рис.2.3 – Линейная с насыщением

- линейная униполярная с насыщением

$$f(z) = \begin{cases} 1 & \text{при } z \geq \frac{1}{2\alpha}; \\ \alpha z + 0,5 & \text{при } |z| < \frac{1}{2\alpha}; \\ 0 & \text{при } z \leq -\frac{1}{2\alpha}. \end{cases}$$

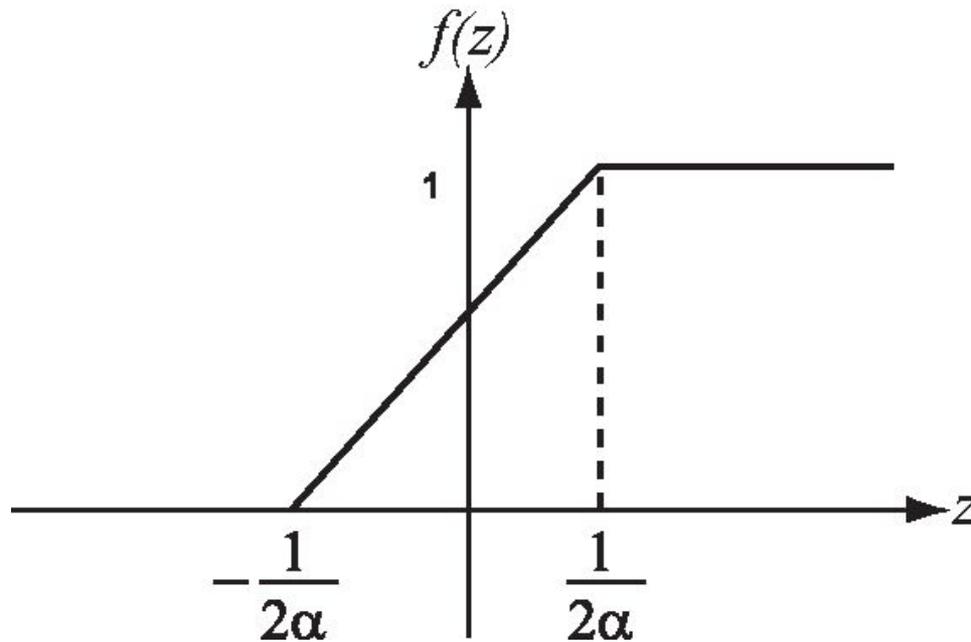


Рис.2.4 – Линейная униполярная функция с насыщением

– функция Хевисайда - униполярная пороговая функция

$$f(z) = \begin{cases} 1 & \text{при } z \geq \alpha; \\ 0 & \text{при } z < \alpha. \end{cases}$$

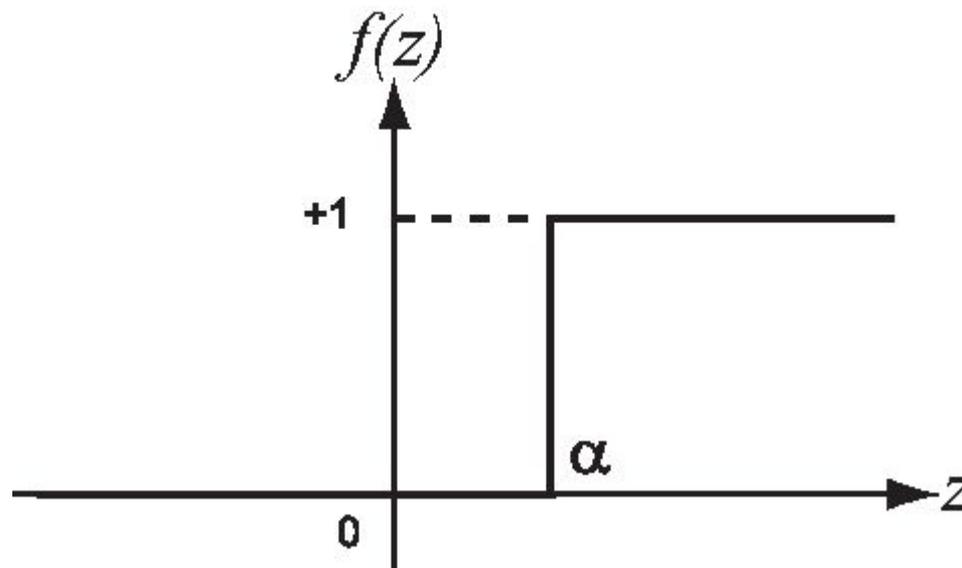


Рис.2.5 – Пороговая функция

– биполярная пороговая функция

$$f(z) = \begin{cases} 1 & \text{при } z \geq \alpha; \\ -1 & \text{при } z < \alpha. \end{cases}$$

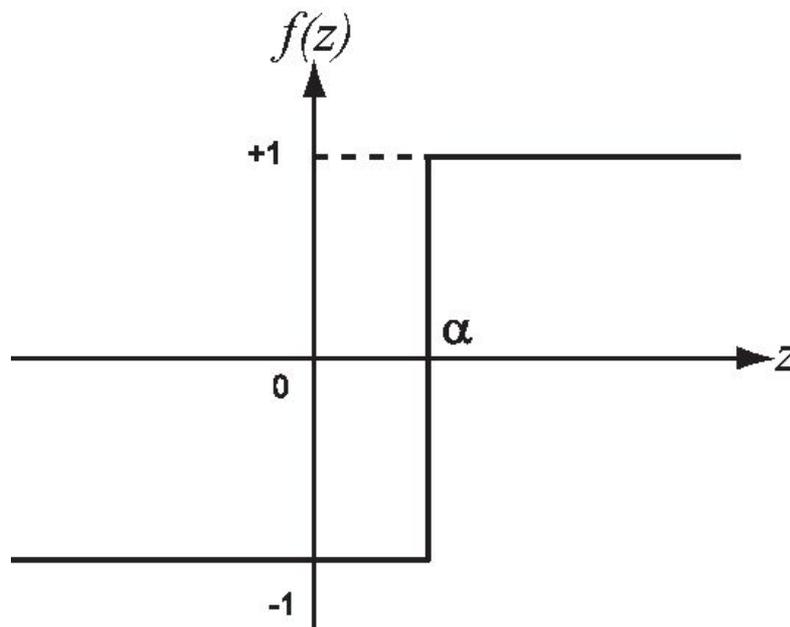


Рис.2.6 – Биполярная пороговая функция

**- логистическая
(униполярная)**

$$f_{\log}(z) = \frac{1}{1 + e^{-\alpha z}}$$

$$\frac{d}{dz} f_{\log}(z) = \alpha f_{\log}(z) (1 - f_{\log}(z))$$

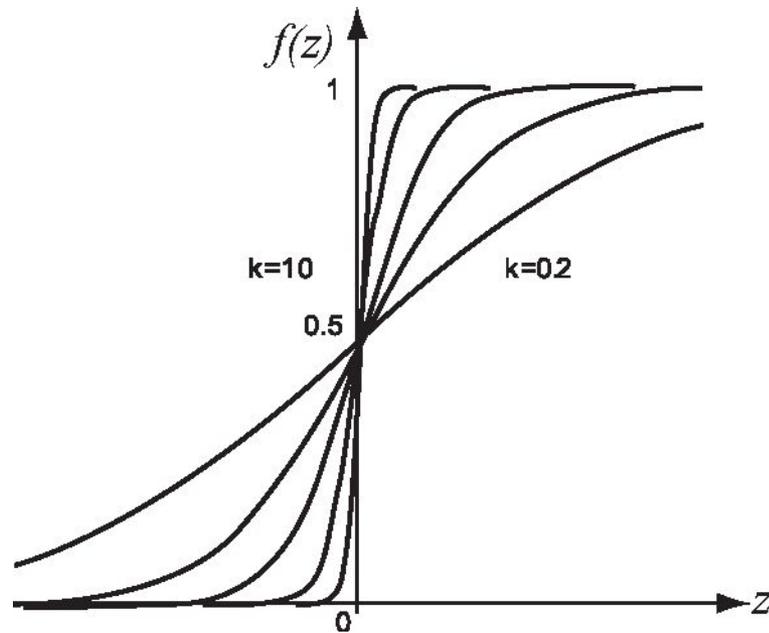


Рис.2.7 – Логистическая функция

-гиперболического тангенса (биполярная)

$$f_{th}(z) = \tanh(\alpha z) = \frac{e^{\alpha z} - e^{-\alpha z}}{e^{\alpha z} + e^{-\alpha z}}$$

$$\frac{d}{dz} f_{th}(z) = 1 - \tanh^2(\alpha z)$$

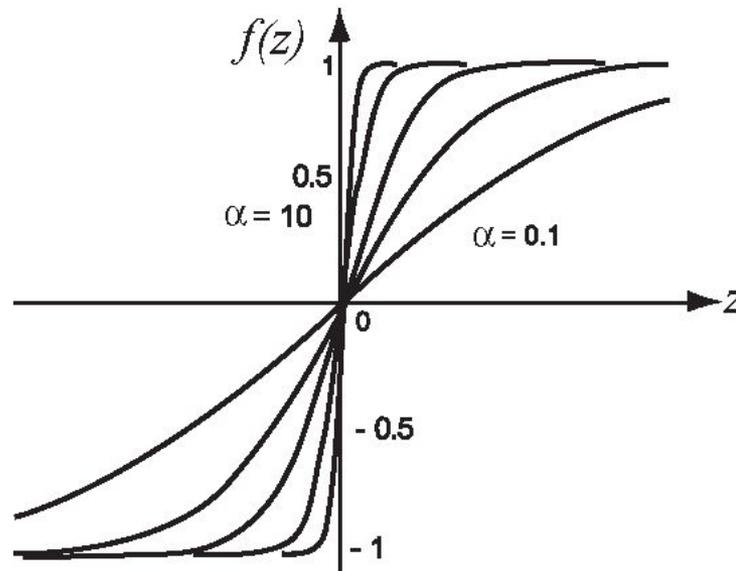


Рис. 2.8 – Функция гиперболического тангенса

Выражение логистической функции и функции гиперболического тангенса друг через друга:

$$f_{tn}(z) = \frac{e^z (1 - e^{-2z})}{e^z (1 + e^{-2z})} = \frac{2 - (1 + e^{-2z})}{1 + e^{-2z}} = 2f_{\log}(2z) - 1$$

$$f_{\log}(z) = \frac{1}{2} \left(\tanh\left(\frac{z}{2}\right) - 1 \right)$$

- модулированная сигмоида

$$f(x, y) = \frac{1}{1 + e^{-x-y-\theta}} - \frac{1}{1 + e^{x-y-\theta}}$$

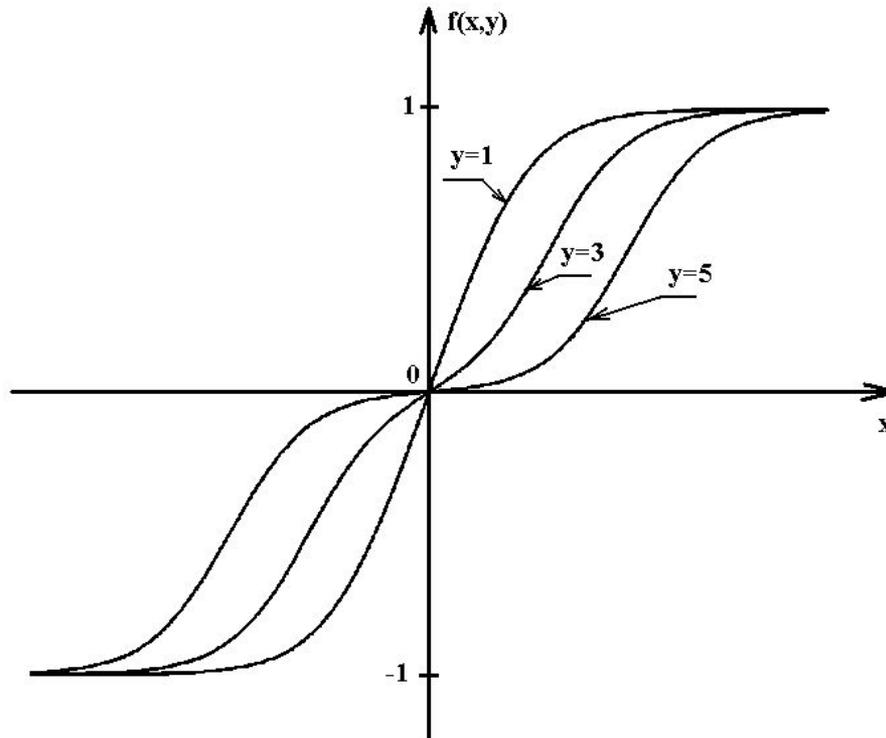


Рис. 2.9 – Функция “модулированная сигмоида”

- синусоидальная с насыщением (биполярная)

$$f_{sn}(z) = \begin{cases} 1 & \text{при } z \geq \alpha; \\ \sin z & \text{при } |z| < \alpha; \\ -1 & \text{при } z \leq -\alpha, \end{cases}$$

$$\frac{d}{dz} f_{sn}(z) = \sqrt{1 - f_{sn}^2(z)}$$

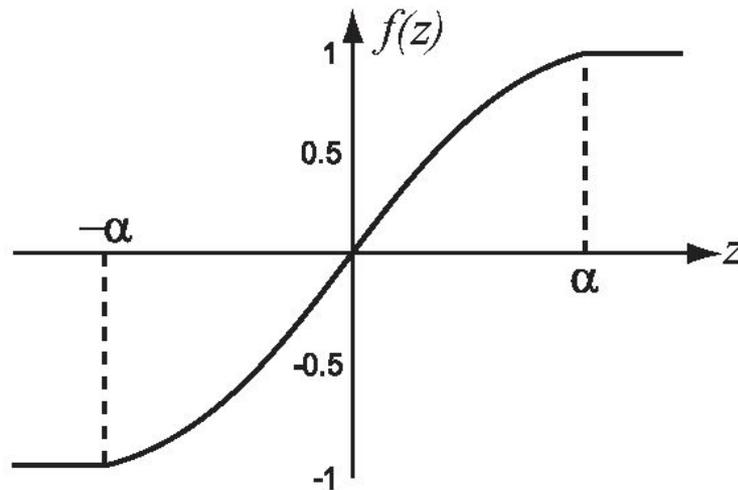


Рис.2.10 – Функция синусоидальная с насыщением

- косинусоидальная с насыщением (униполярная)

$$f_{CS}(z) = \begin{cases} 1 & \text{при } z \geq \frac{\pi}{2}; \\ \frac{1}{2} (1 + \cos(z - \frac{\pi}{2})) & \text{при } |z| < \frac{\pi}{2}; \\ 0 & \text{при } z \leq -\frac{\pi}{2}. \end{cases}$$

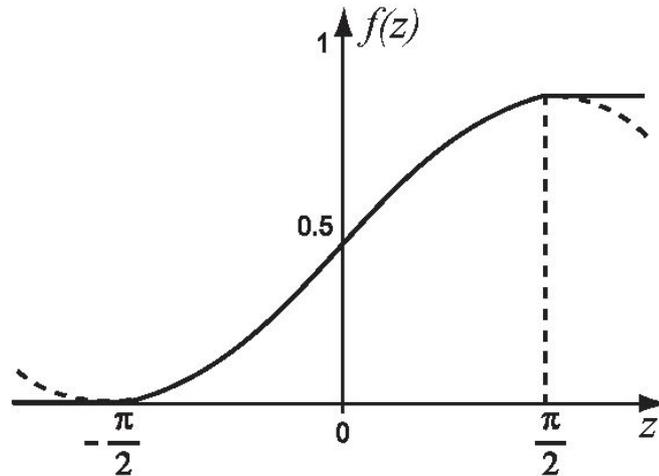


Рис.2.11 – Функция косинусоидальная с насыщением

Топология ИНС

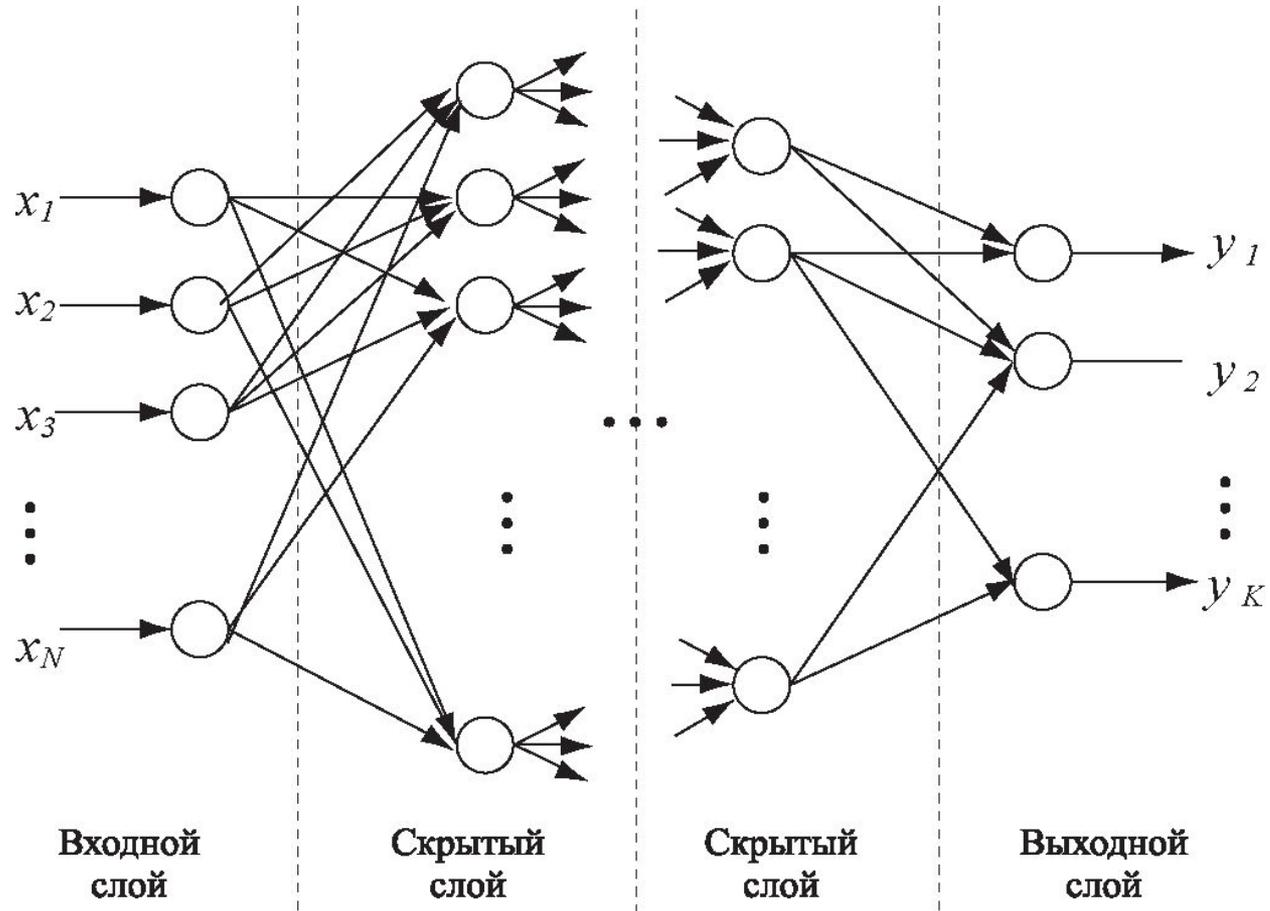


Рис.2.12 – Структура ИНС

Способы представления ИНС

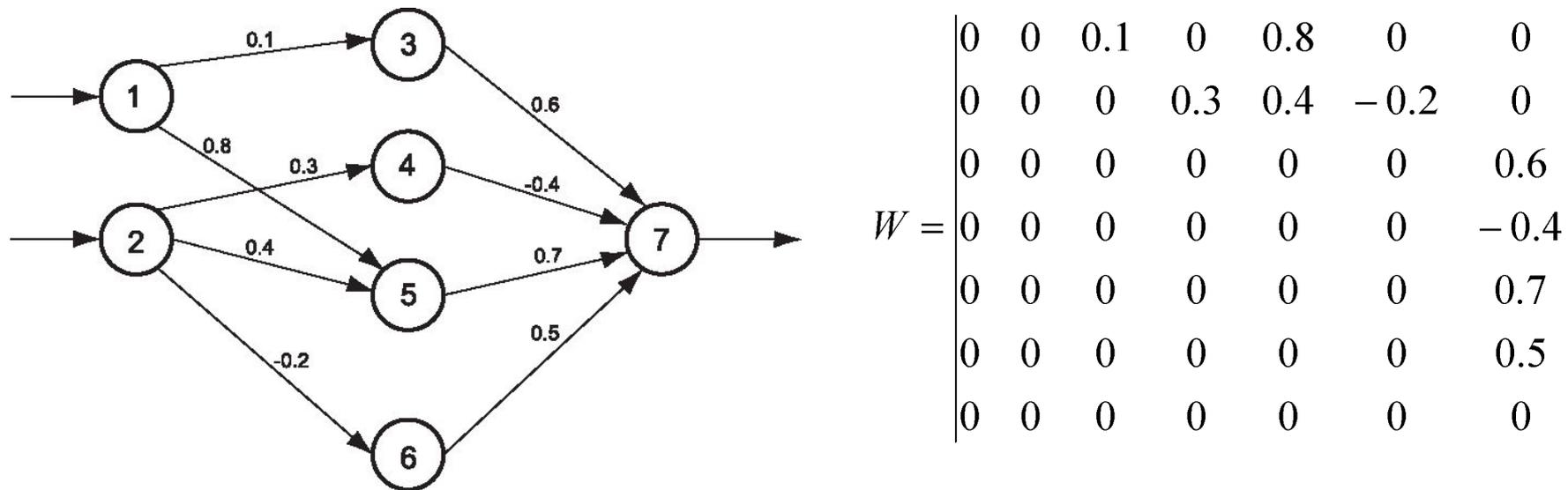


Рис.2.13 – Пример представления ИНС

Топологии ИНС:

1. *ИНС без обратных связей (прямого распространения, feed-forward)*

- *первого порядка;*
- *второго порядка (с “shortcut connections”).*

2. *ИНС с обратными связями (обратного распространения, рекуррентные, Feedback)*

- *с прямыми обратными связями (direct feedback);*
- *с непрямыми обратными связями (indirect feedback);*
- *с латеральными связями (lateral feedback);*
- *полносвязные.*

ИНС прямого распространения

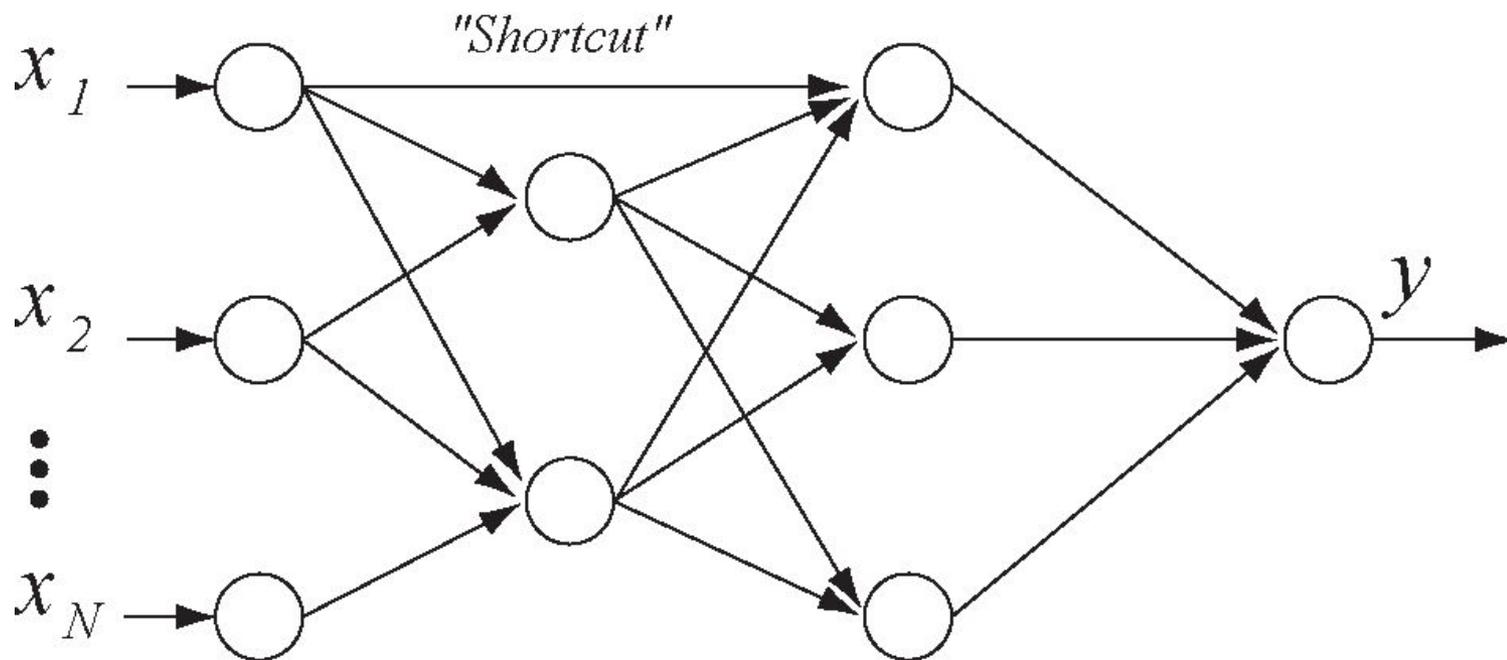


Рис. 2.14 – ИНС прямого распространения второго порядка

ИНС обратного распространения

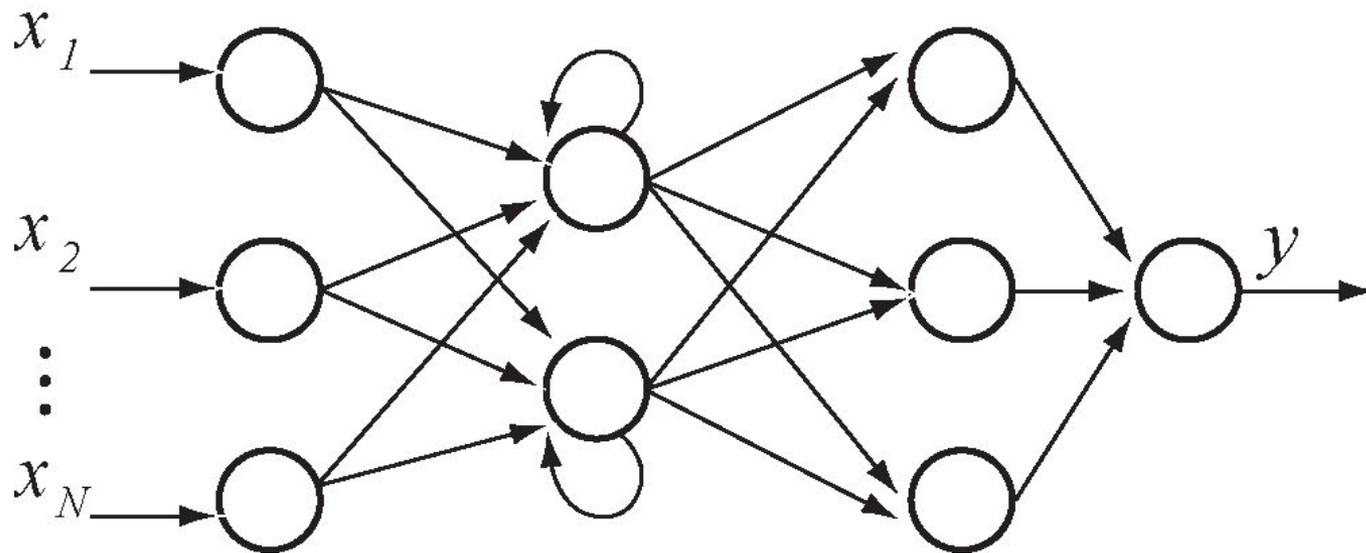


Рис.2.15 – ИНС с прямыми обратными связями

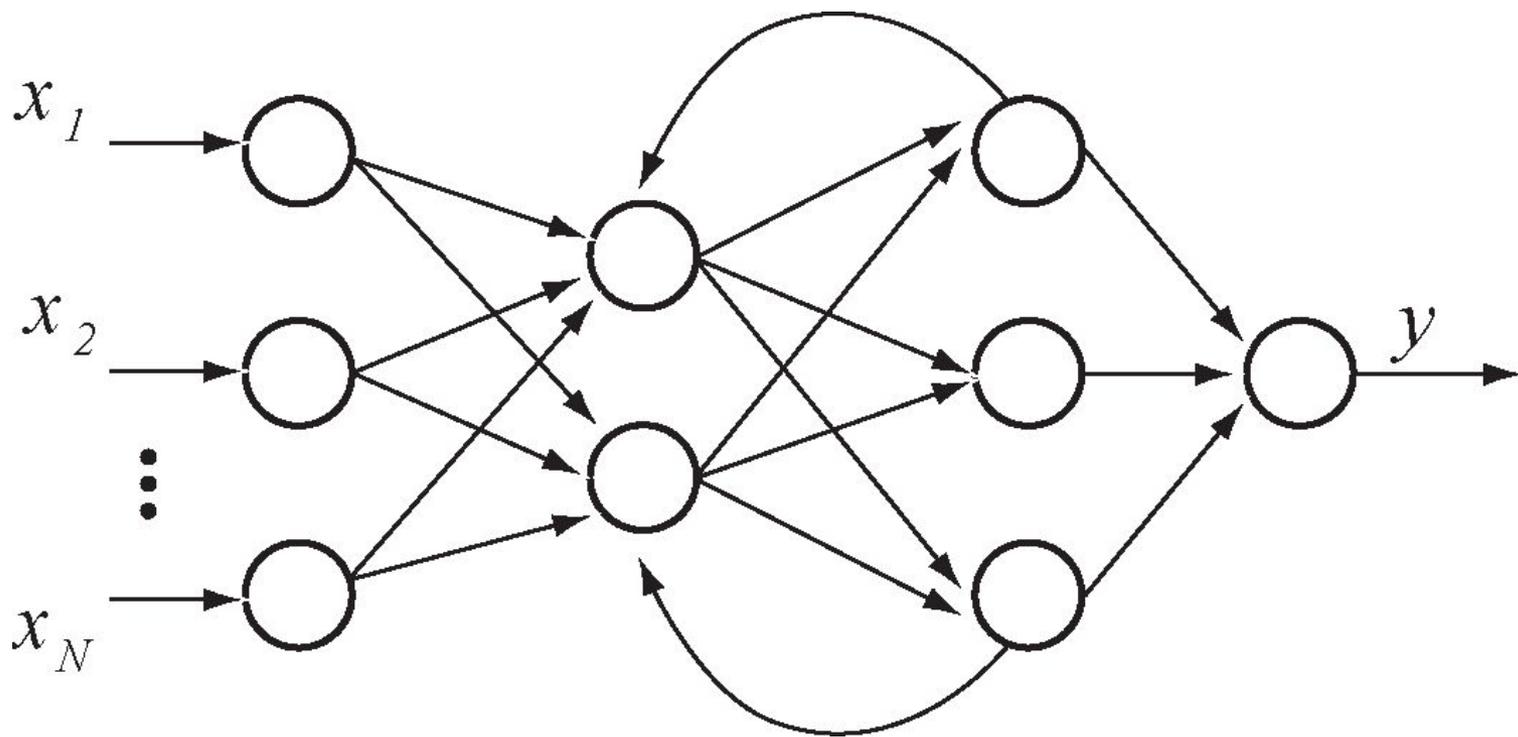


Рис.2.16 – ИНС с непрямыми обратными связями

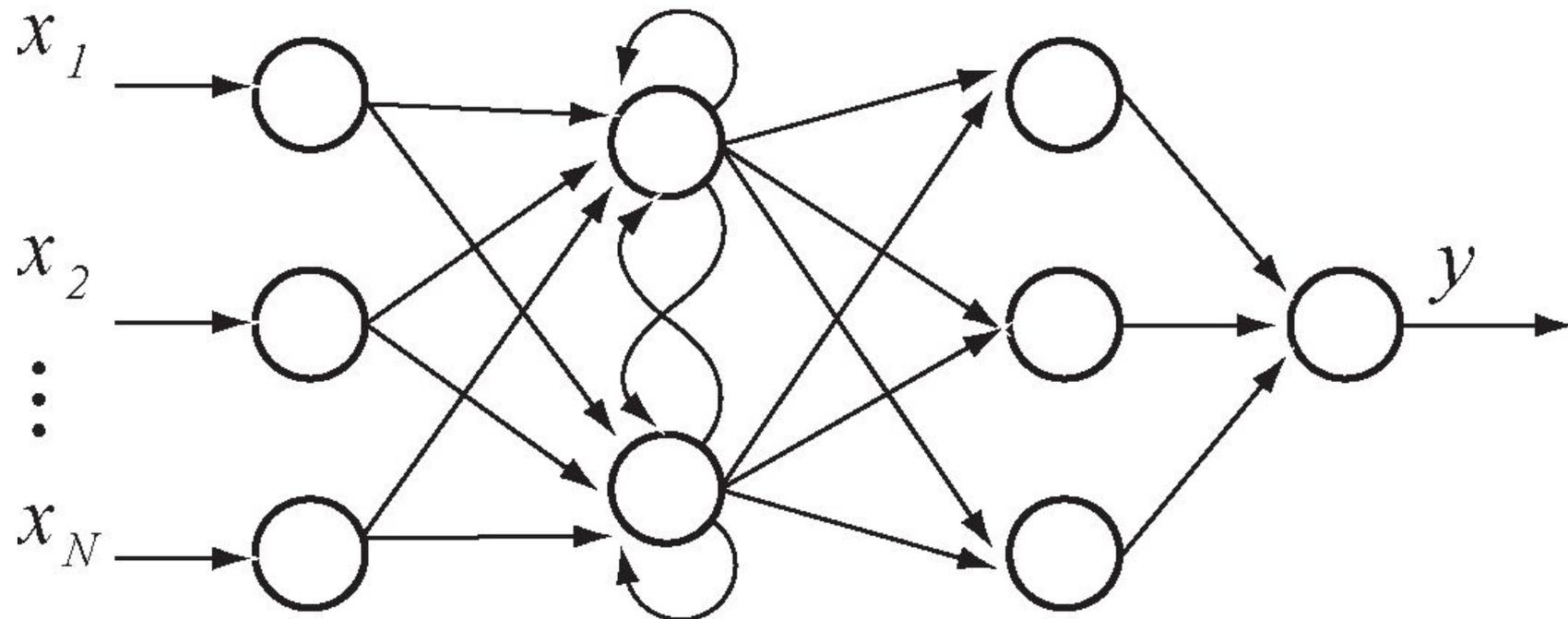


Рис.2.17 – ИНС с латеральными связями

Полносвязные ИНС

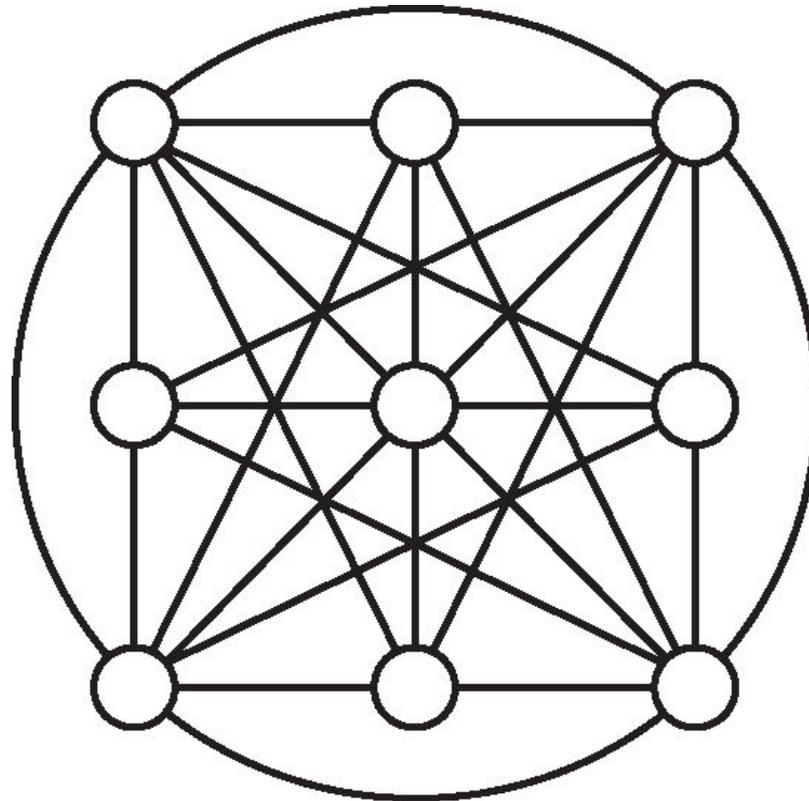


Рис.2.18 – Полносвязная ИНС

Подходы к обучению ИНС:

- изменение конфигурации сети путем образования новых или исключения некоторых существующих связей между нейронами;*
- изменение элементов матрицы связи (весов);*
- изменение характеристик нейронов (вида и параметров активационной функции и т.д.).*

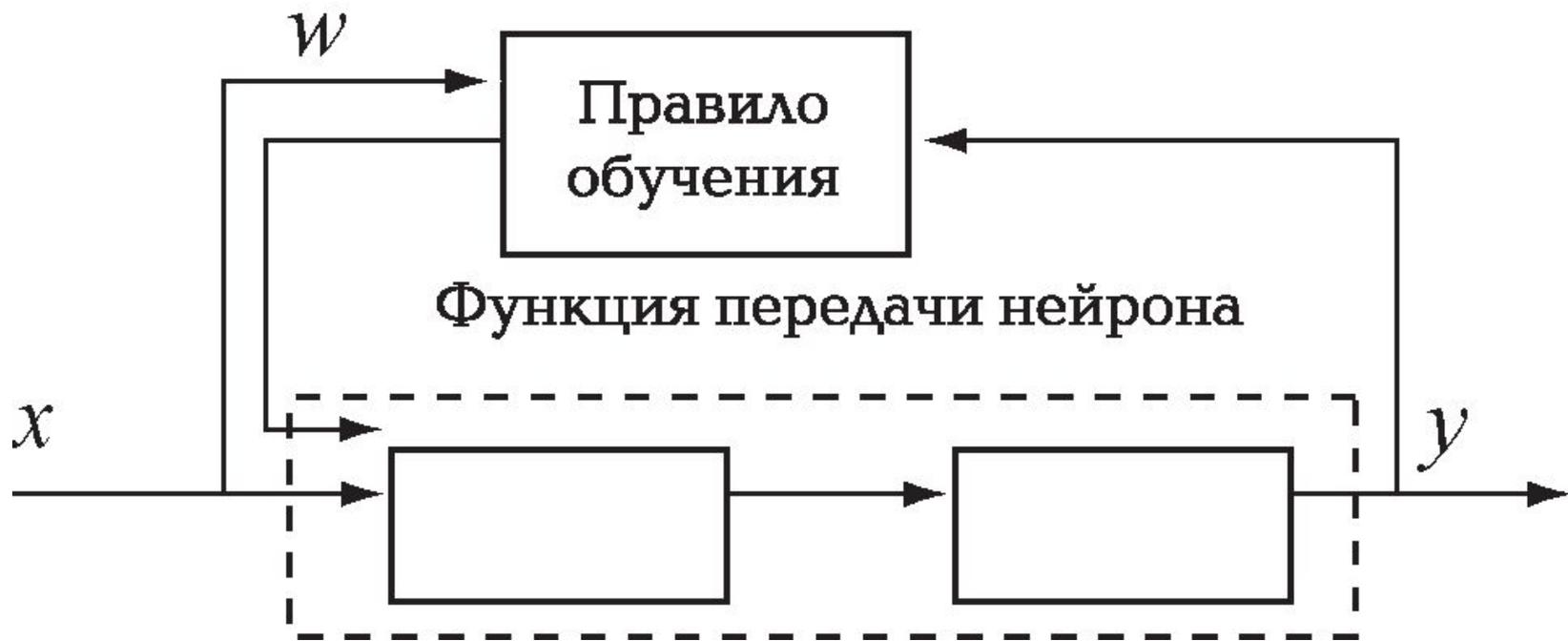


Рис.2.19 – Модель искусственного нейрона

Типы обучения ИНС:

- Обучение с учителем;*
- Обучение без учителя;*
- Подкрепляемое обучение.*

Правило обучения Хэбба (коррелятивное, относительное обучение)

$$\Delta w_{ij} = \gamma x_i y_j \quad ,$$

В векторном виде данное правило может быть записано следующим образом:

для непрерывного времени

$$\frac{\partial w}{\partial t} = \gamma \mathbf{x} y$$

для дискретного времени

$$w(k + 1) = w(k) + \gamma \mathbf{x}(k) y(k) \quad .$$

Дельта-правило

$$\Delta w_{ij} = \gamma x_i (y_j^* - y_j)$$

Расширенное дельта-правило

$$\Delta w_{ij} = \gamma x_i \delta_j$$

$$\delta_j = \begin{cases} f'_j(x, w) (y_j^* - y_j) & \text{если } j\text{-нейрон выходного слоя;} \\ f'_j(x, w) \sum_m \delta_m w_{im} & \text{в противном случае;} \end{cases}$$

-Конкурентное обучение;

-Стохастическое обучение;

-Градиентные методы обучения

$$\Delta w_{ij} = \gamma \nabla_w I(w)$$

где γ - коэффициент, влияющий на скорость обучения;

$$\nabla_w I(w) = \frac{\partial I(w)}{\partial w_{ij}}$$

Модели искусственных нейронов.

$$y_j = f\left(\sum_{i=1}^N w_{ji}x_i + \theta_j\right)$$

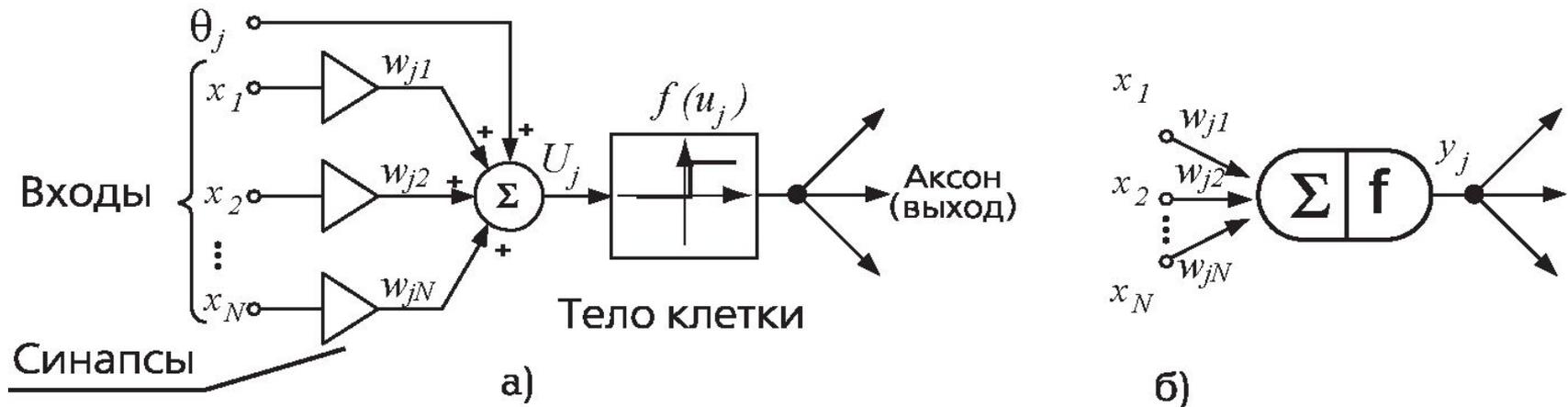


Рис. 2.20 – Модель искусственного нейрона

Маккаллоха-Питтса и её обозначение

$$y_j = f(u_j) = f\left(\frac{\sum_{i=0}^N G_{ji}x_i}{\sum_{i=0}^N G_{ji}}\right),$$

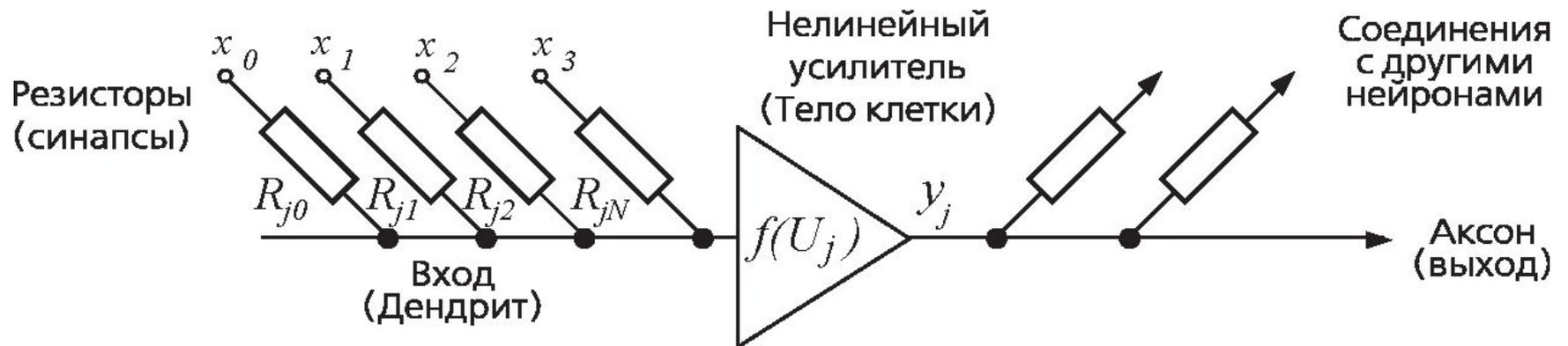


Рис. 2.21 – Электронная аналоговая модель нейронной клетки

$$y_j = f \left[\frac{1 + \sum_{i=1}^n a_{ji} x_i}{1 + \sum_{i=1}^m b_{ji} v_i} - 1 \right], \quad f(u_j) = \begin{cases} u_j, & \text{если } u_j \geq 0; \\ 0 & \text{в противном случае.} \end{cases}$$

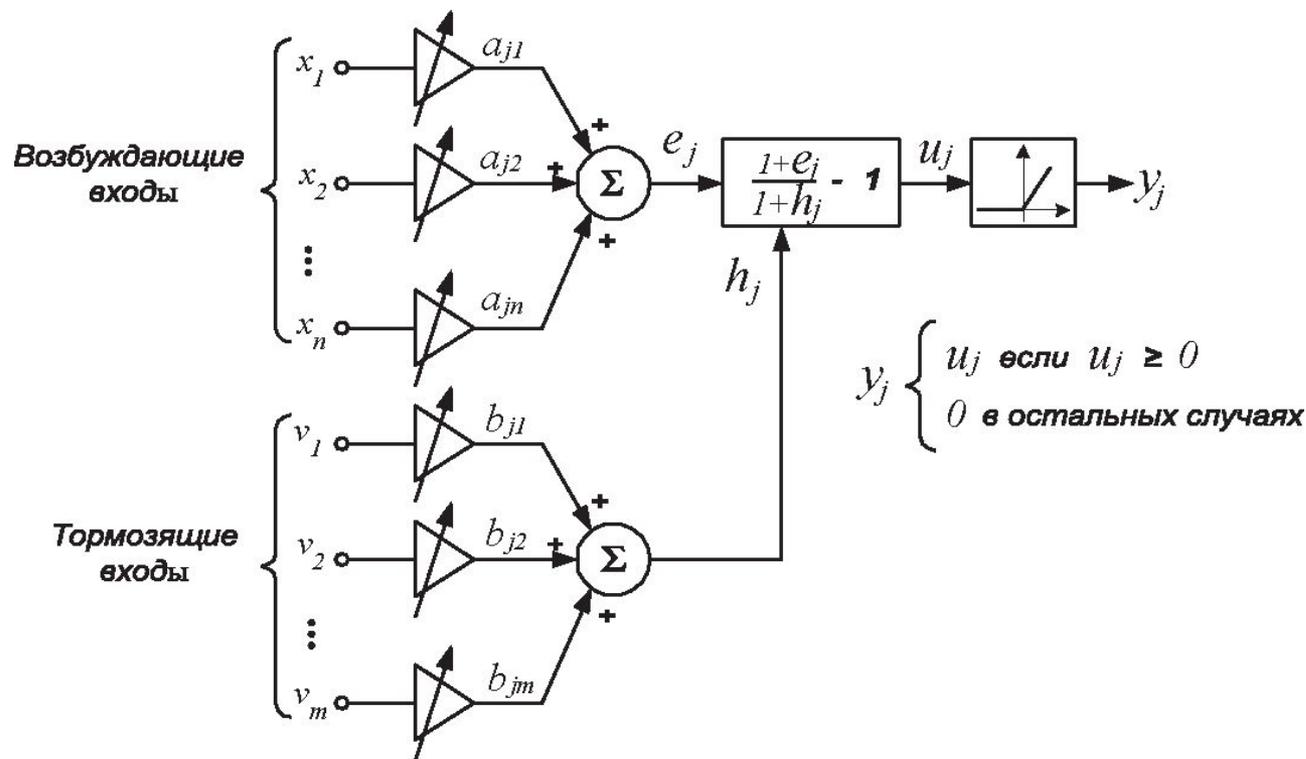


Рис. 2.22 – Упрощенная модель нейрона Фукушимы

Модель искусственного нейрона Хопфилда

$$C_j \frac{du_j}{d\tau} = -\frac{u_j}{R_j} + \sum_{i=1}^N \frac{v_i}{R_{ji}} + I_j,$$

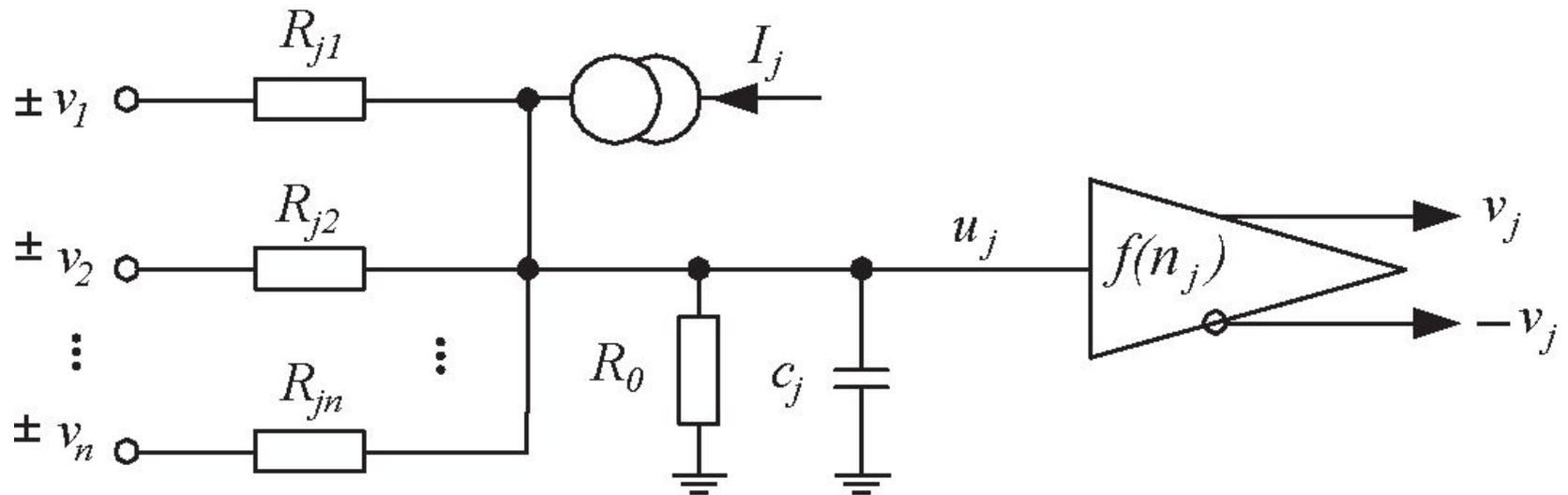


Рис. 2.23 – Модель Хопфилда динамической нейронной клетки

$$T_{0j} \frac{du_j}{d\tau} = -a_j u_j + \left(\sum_{i=0}^N w_{ji} x_i + \theta_j \right),$$

где $T_j = \tau_j C_j$ - постоянная интегрирования;

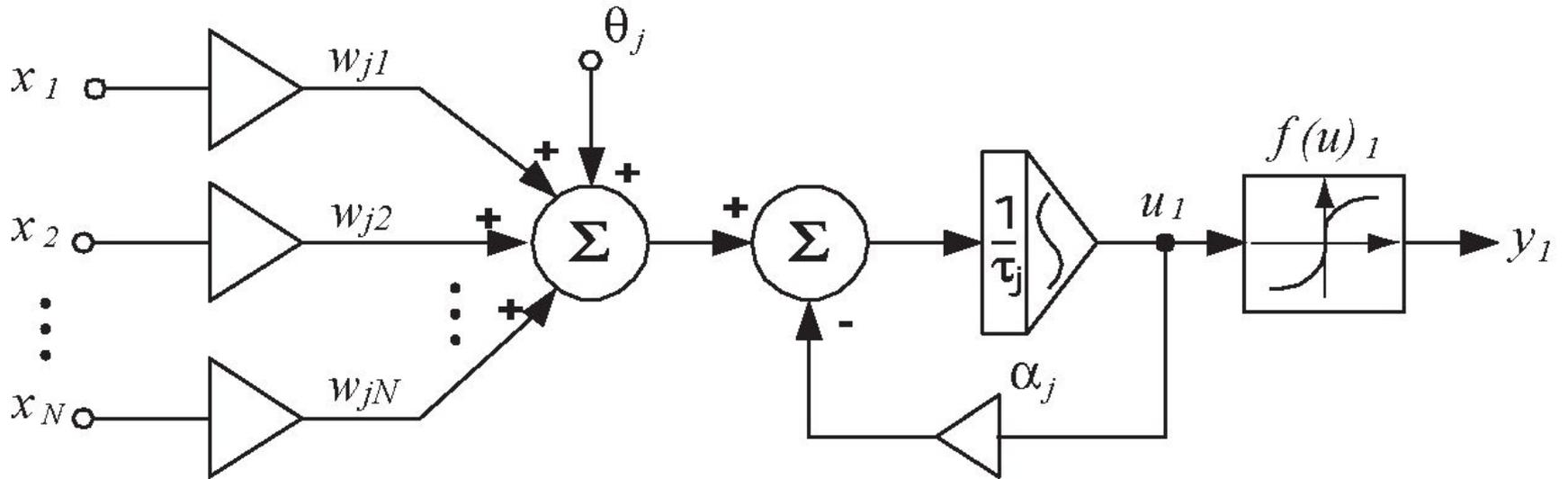


Рис.2.24 – Структура нейрона Хопфилда

Модель Гроссберга

$$T_j \frac{du_j}{d\tau} = -\alpha_j u_j + (\gamma_j - \beta_j u_j) \left(\sum_{i=1}^n w_{ji} f_i(u_i), f_i(u_i, \theta_j) \right),$$

$$\frac{dw_{ji}}{d\tau} = \left[-\theta_{ji} w_{ji} + d_{ji} f_i(u_i) h_i(u_i) \right]$$

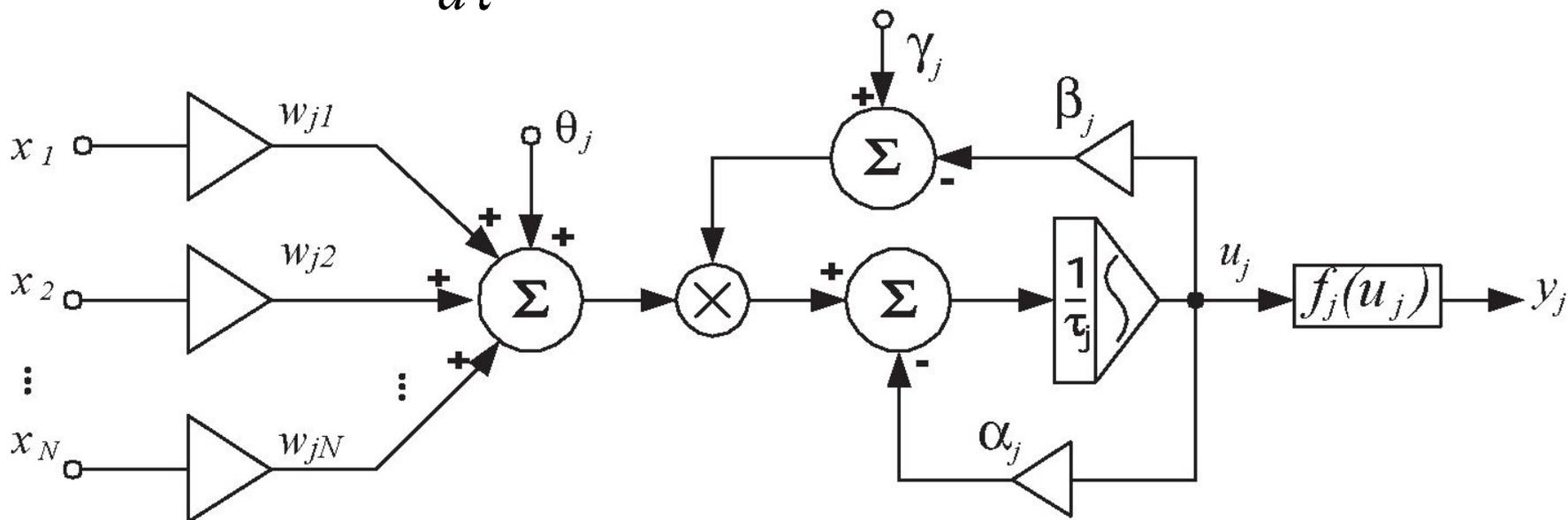


Рис.2.25 – Функциональная модель нейрона Гроссберга

Сигма-ПИ-нейрон

$$z(w, x) = \sum_{i=1}^m w_i \prod_{j=1}^{N_i} x_j, \quad y = f(w, x) = f\left(\sum_{i=1}^m w_i \prod_{j=1}^{N_i} x_j\right)$$

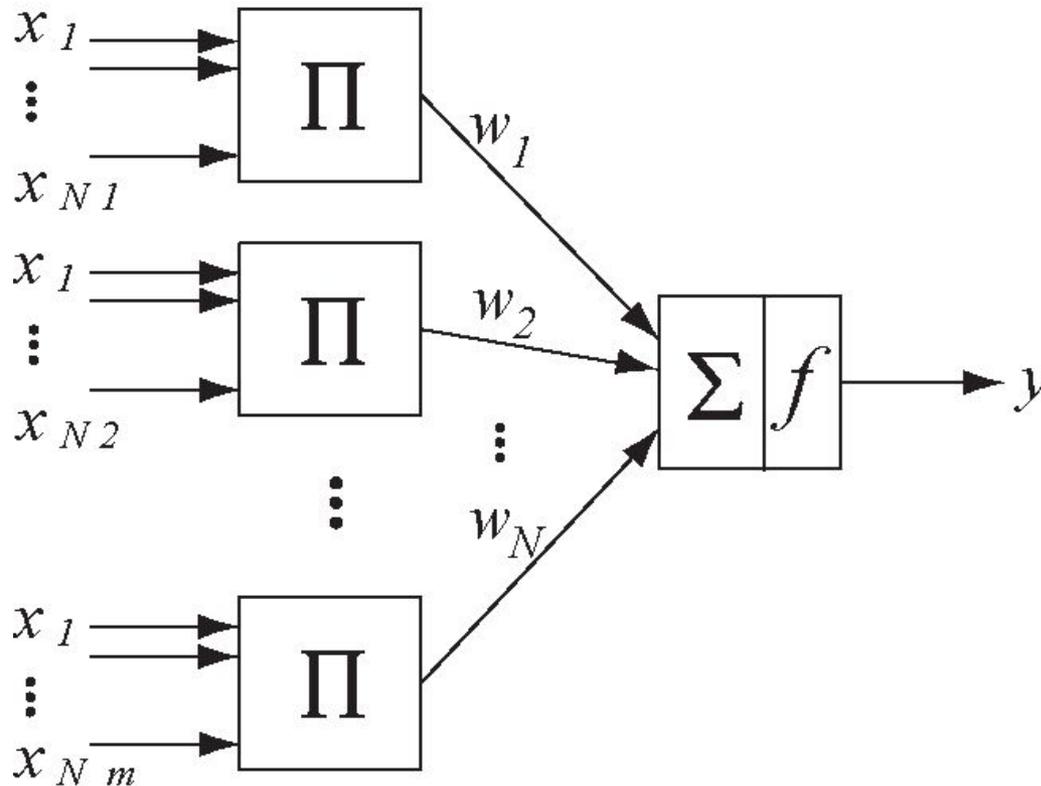


Рис.2.26 – Сигма-Пи нейрон

Стохастический нейрон

$$y = \begin{cases} +1 & \text{с вероятностью } P(z|y=1); \\ -1 & \text{с вероятностью } P(z|y=-1), \end{cases}$$

где $P(z|y=1) + P(z|y=-1) = 1$.

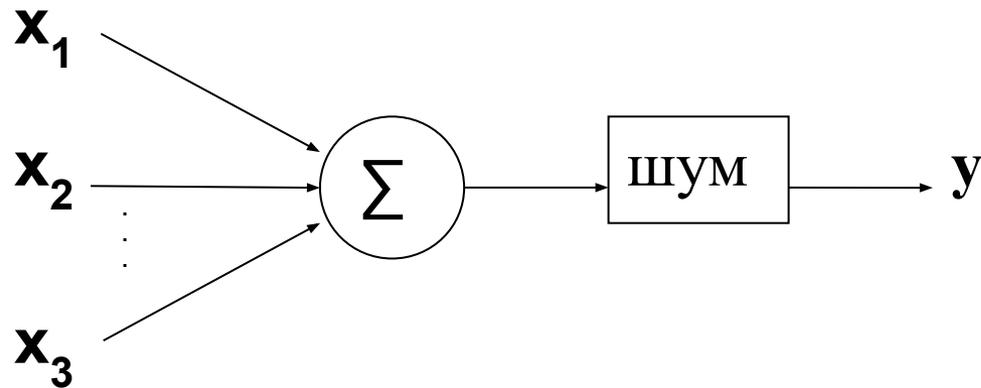


Рис. 2.27 – Модель стохастического нейрона