# О СОВМЕСТНОМ РАВНОМЕРНОМ ПРИБЛИЖЕНИИ ПОЛИНОМИАЛЬНЫХ ФУНКЦИЙ И ИХ ПРОИЗВОДНЫХ НЕЙРОННЫМИ СЕТЯМИ ПРЯМОГО РАСПРОСТРАНЕНИЯ СИГНАЛА

**Н. С. Узенцова, С. П. Сидоров**

Саратовский государственный университет
механико-математический факультет

Поддержано РФФИ

2013

# Содержание



1. Нейронные сети;
2. Нейронная сеть прямого распространения с одним входом и одним выходом;
3. Одновременное приближение алгебраических многочленов и их производных нейронными сетями прямого распространения.

# 1. Нейронные сети

# Архитектура многослойной сети прямого распространения



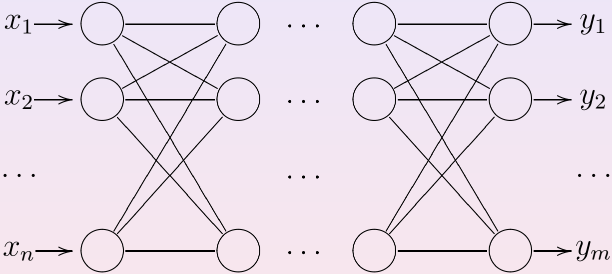


Рис. : Архитектура многослойной сети прямого распространения

# Нейрон — составная часть ИНС

Нейрон состоит из элементов трех типов:

- синапсы
- сумматор
- нелинейный преобразователь

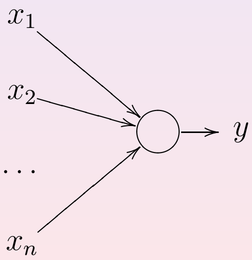


Рис. : Нейрон

# Математическая модель нейрона



$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right),$$

где

- $n$ — число входов
- $x_i$ — $i$-ая компонента входного вектора (входного сигнала)
- $w_i$ — вес синапса
- $b$ — значение смещения
- $\sigma$ — нелинейное преобразование (функция активации)
- $y$ — выходной сигнал нейрона

# Виды функций активаций



В качестве функции активации $\sigma$ нейронов берут обычно одну из следующих:

- пороговая функция активации, $\sigma(s) = \begin{cases} 1, & s \geqslant 0 \\ 0, & s < 0 \end{cases}$ ;
- экспоненциальная сигмоида, $\sigma(s) = \frac{1}{1+e^{-\alpha s}}$;
- рациональная сигмоида, $\sigma(s) = \frac{s}{|s|+\alpha}$.

# 2. Нейронная сеть прямого распространения с одним входом и одним выходом

# Обозначения



Пусть $\sigma : \mathbb{R} \to \mathbb{R}$ есть некоторая функция.
Пусть $r \in \mathbb{N}$.

Рассмотрим нейронную сеть $F : \mathbb{R} \to \mathbb{R}$ прямого распространения с одним входом и одним выходом:

$$F(x) = \sum_{j=1}^{N} c_j \sigma(w_j x + b_j),$$

где $c_j, w_j, b_j \in \mathbb{R}$, $j = 1, \ldots, N$, есть параметры сети.
Обозначим множество всех таких отображений $\mathcal{F}_N(\sigma)$.
Полиномиальная функция $f(x) : \mathbb{R} \to \mathbb{R}$ степени $r$ имеет вид:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_r x^r,$$

где $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, r$.

## Лемма 1

Производная функции $F$ порядка $s$ существует в каждой точке $x \in \mathbb{R}$, при этом справедливо равенство

$$\frac{d^s F(x)}{dx^s} = \sum_{j=1}^{N} c_j w_j^s Q_s(\sigma(\theta_j)), \qquad (1)$$

где

$$Q_s(\sigma(\theta_j)) = (-1)^{s+1} s! \theta_j^{-s} \sigma^s(\theta_j)(1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j),$$

$$\theta_j = \theta_j(x) = w_j x + b_j.$$

Предположим, что

$$\frac{d^s F(0)}{dx^s} = \frac{d^s f(0)}{dx^s} = (s!) a_s, s = 0, 1, 2, \ldots r,$$

Преобразуем в матричное равенство и получаем $\Omega c = a$, где

**Лемма 2**

Пусть $\delta > 0$ и $w_i = (i-1)\delta$, $i = 1, \ldots, N$. Если $c_1, \ldots, c_N$ есть решение системы (1), то справедливы неравенства

$$|c_j| \leq \psi_{\max} q_{\max} \sum_{i=0}^{r} |i! a_i| \delta^{-i}, \; j = 1, \ldots, N,$$

где $\psi_{\max}$ есть положительное число, зависящее от $N$ и $r$, $q_{\max}$ есть положительное число, зависящее от порогового значения в скрытом слое.

# Основной результат



**Теорема**

Пусть $x_{max} > 0$ – произвольное число. Сеть с одним скрытым слоем, содержащем $N$ узлов, может аппроксимировать любой алгебраический многочлен степени $N-1$ совместно с его производными с любой степенью точности на отрезке $[-x_{max}, x_{max}]$, т.е. $\forall f \in \mathcal{P}_{N-1}$, $\forall \varepsilon > 0$, $\exists F(x) \in \mathcal{F}_N(\sigma)$ такая, что:

$$\sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| < \varepsilon, \forall x \in [-x_{max}, x_{max}],$$

где $\varepsilon$ является ошибкой равномерного приближения на $[-x_{max}, x_{max}]$.

# Доказательство теоремы



Пусть $f \in \mathcal{P}_{N-1}$ есть произвольный многочлен. Возьмем $\varepsilon > 0$ и покажем, что для некоторой $F(x) \in \mathcal{F}_N(\sigma)$ для всех $x \in [-x_{max}, x_{max}]$ будет

$$\sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| < \varepsilon.$$

Потребуем

$$\frac{d^s F(0)}{dx^s} = \frac{d^s f(0)}{dx^s}, s = 0, 1, 2, \dots N-1,$$

Пусть $r = N - 1$. Так как $F^{(s)}(x)$ имеет производную любого порядка в каждой точке вещественной оси, то функцию $F^{(s)}(x)$ можно разложить в ряд Тейлора в окрестности точки $x = 0$. Так, для $s = 0$ имеем

$$F(x) = F(0) + F'(0)x + \frac{1}{2!}F''(0)x^2 + \dots + \frac{1}{r!}F^{(r)}(0)x^r + R_{r,0}(x),$$

где $R_{r,0}(x)$ есть остаточный член форме Лагранжа:

$$R_{r,0}(x) = \frac{F^{(r+1)}(\zeta_0)}{(r+1)!}x^{r+1}.$$

# Доказательство теоремы



Имеем:

$$F^{(s)}(x) = \sum_{l=s}^{r} \frac{F^{(l)}(0)x^{l-s}}{(l-s)!} + R_{r,s}(x),$$

где $R_{r,s}(x)$ есть остаточный член в форме Лагранжа,

$$R_{r,s}(x) = \frac{F^{(r+1)}(\zeta_s)}{(r+1-s)!}x^{r+1-s}.$$

Имеем:

$$\sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| = \sum_{s=0}^{N-1} |R_{r,s}(x)|.$$

Согласно лемме 1

$$R_{r-s}(x) = \frac{x^{r+1-s}}{(r+1-s)!} \sum_{j=1}^{N} c_j w_j^{r+1} Q_{r+1}(\sigma(\theta_j(\zeta_s))),$$

где $-x_{max} \leq \zeta \leq x_{max}$.

Возьмем $0 < \delta < 1$. Положим $w_j = (j-1)\delta$, $j = 1, 2, \dots, N$. Согласно лемме 1 для $s \geq 1$

$$Q_s(\sigma(\theta_j)) = (-1)^{s+1} s! \theta_j^{-s} \sigma^s(\theta_j)(1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j)$$

и $Q_0(\sigma_j) = \sigma_j$, где $\sigma_j(x) = \sigma(\theta_j(x))$. Так как для всех $x$ имеет место $-1 < \sigma_j < 1$, то $|Q_0(\sigma_j)| \leq 1$ и существует $M^0$ такое, что $|\sigma^s(\theta_j)(1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j)| < M^0$. Имеем, учитывая лемму 2,

$$\sum_{s=0}^{N-1} |R_{r,s}(x)| \leq \sum_{s=0}^{N-1} \left( \frac{x_{max}^{r+1-s}}{(r+1-s)!} \sum_{j=1}^{N} |c_j| |w_j^{r+1}| |Q_{r+1}(\sigma_j(\zeta_s))| \right) \leq$$

$$\leq \sum_{s=0}^{N-1} \left( \frac{x_{max}^{r+1-s}}{(r+1-s)!} \psi_{\max} q_{\max} \sum_{j=1}^{N} \sum_{p=0}^{r} |p! a_p| \delta^{-p} (r+1)! j^{r+1} \delta^{r+1} M^0 \right) \leq$$

$$\leq \delta \psi_{max} q_{max} M^0 \sum_{p=0}^{r} |p! a_p| \sum_{s=0}^{N-1} \left( \frac{(r+1)!}{(r+1-s)!} x_{max}^{r+1-s} \sum_{j=1}^{N} j^{r+1} \right).$$

Обозначим

$$M = \psi_{max} q_{max} M^0 \sum_{p=0}^{r} |p! a_p| \sum_{s=0}^{N-1} \left( \frac{(r+1)!}{(r+1-s)!} x_{max}^{r+1-s} \sum_{j=1}^{N} j^{r+1} \right),$$

т.е. $M$ есть конечное число, не зависящее от $x$. Имеем

$$\sum_{s=0}^{N-1} |R_{r,s}(x)| \leq \delta M.$$

Взяв $\delta < \frac{\varepsilon}{(M+1)}$, получаем $\sum_{s=0}^{N} |R_r(x)| < \varepsilon$, где $\varepsilon$ есть ошибка аппроксимации. Теорема доказана.

БЛАГОДАРЮ ЗА ВНИМАНИЕ!