Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное

учреждение высшего образования  
 «Севастопольский государственный университет»

**Отчёт**

по лабораторной работе

**Исследование активационных функций нейронных элементов**

Дисциплина «Основы нейронных сетей»

Выполнил студент гр. ГНКЗ-8

Мовенко К.М.

Вариант − 5

**Севастополь**

**2023**

# ЦЕЛЬ РАБОТЫ

Углубление теоретических знаний в области архитектуры нейронных сетей, исследование свойств активационных функций нейронных элементов, приобретение практических навыков моделирования простейшей нейронной сети прямого распространения.

# ЗАДАНИЕ

Таблица 1 – Вариант задания №5

|  |  |  |  |
| --- | --- | --- | --- |
| Структура сети | Активационная функция 1 | Активационная функция 2 | Входной сигнал для i-го входа сети |
| 100-10-1 | tansig | purelin | sin(2\*%pi\*0.01\*i\*t) |

1. Построить графики активационных функций и их производных в соответствии с вариантом задания (Таблица 1), используя соответствующие встроенные функции модуля NeuralNetworks 3.0 пакета Scilab. В случае отсутствия указанных встроенных функций определить их самостоятельно на основе выражений, приведенных в методических указаниях;
2. Реализовать две сети прямого распространения в соответствии с заданной структурой, запрограммировав вычисления в соответствии с выражением для двух видов активационных функций, заданных по варианту;
3. Выполнить моделирование двух нейросетей, сгенерировав входные сигналы в соответствии с вариантом задания на интервале времени от 0 до 2/F, где F – минимальная частота входного гармонического сигнала. При этом инициализацию сетей выполнить двумя способами:
   1. случайными значениями из диапазона от -10 до +10;
   2. случайными значениями из диапазона от -1/sqrt(R) до +1/sqrt(R), где R число входов нейрона;
4. Вычислить выходные значения всех нейронов сети. Построить графики активностей одного из нейронов скрытого слоя и нейрона выходного слоя, а также гистограмму общей активности всех нейронов скрытого слоя и отдельно гистограмму активности выходного нейрона для двух способов инициализации для каждой из двух реализованных сетей. Вычислить значение средней активности и стандартного отклонения на выходах нейронов скрытого и выходного слоёв;
5. Вычислить значения производных нелинейностей для всех нейронов в точках, соответствующих значениям входных сигналов этих нейронов. Построить графики вычисленных значений производных для одного из нейронов скрытого слоя и нейрона выходного слоя, а также гистограмму значений производных функции активации скрытых нейронов и выходного нейрона (для двух способов инициализации и для каждой из двух реализованных сетей). Вычислить среднее значение и стандартное отклонение производных для нейронов скрытого и выходного слоёв.
6. Выполнить анализ полученных результатов, обратив внимание на характер гистограмм. Сделать выводы о характере распределения значений активностей и значений производных для заданных активационных функций и каждого из способов инициализации. Сравнить результаты для каждой из двух реализованных сетей.

# ЛИСТИНГИ ПРОГРАММ

Листинг 1 – Построение графиков активационных функций и их производных

*// Диапазон значений x*

x = [-10:0.05:10];

*// Значения функции tansig и её производной*

y1 = ann\_tansig\_activ(x);

y2 = ann\_d\_tansig\_activ(y1);

*// График функции tansig и её производной*

subplot(1, 2, 1);

plot(x, y1, 'r', x, y2, 'b');

title('График функции tansig и её производной');

xlabel('x');

ylabel('y');

legend('tansig', 'tansig''');

*// Значения функции purelin и её производной*

y1 = ann\_purelin\_activ(x);

y2 = ann\_d\_purelin\_activ(y1);

*// График функции purelin и её производной*

subplot(1, 2, 2);

plot(x, y1, 'r', x, y2, 'b');

title('График функции purelin и её производной');

xlabel('x');

ylabel('y');

legend('purelin', 'purelin''');

Листинг 2 – Исследование сети с активационной функцией tansig

*// Задание исходных данных*

F = 0.01; *// минимальная частота входного сигнала*

R = 100; *// число входов слоя*

S1 = 10; *// число нейронов слоя 1*

S2 = 1; *// число нейронов слоя 2*

*// Формирование матрицы входных сигналов*

t = 0:0.1:2/F *// отсчёты времени от 0 до 2/F*

fi = (2\*%pi\*F).\*t *// аргументы синуса*

p = [sin(fi)]; *// входные сигналы для 1-го входа*

for i = 2:R *// формирование входных сигналов для i-х входов*

p = [p; sin(fi\*i)];

end

*// Инициализация весов и смещений для слоя 1*

scale = 1/sqrt(R); *// масштабирование*

*//scale = 10;*

w1 = grand(S1, R, "unf", -scale, scale); *// вес*

b1 = zeros(S1, 1); *// смещение*

*// Моделирование слоя 1*

n1 = w1\*p + repmat(b1, 1, size(p, 2)); *// сетевые значения*

a1 = ann\_tansig\_activ(n1); *// активность на выходе 1-го слоя*

d\_a1 = ann\_d\_tansig\_activ(a1); *// производная активационной функции*

*// Вычисление статистик слоя 1*

mean\_a1 = mean(a1); *// среднее значение функции*

stdev\_a1 = stdev(a1); *// стандартное отклонение*

mean\_d\_a1 = mean(d\_a1); *// среднее значение производной*

stdev\_d\_a1 = stdev(d\_a1); *// стандартное отклонение*

*// Инициализация весов и смещений для слоя 2*

scale = 1/sqrt(S1) *// масштабирование*

*//scale = 10;*

w2 = grand(S2, S1, "unf", -scale, scale); *// вес*

b2 = zeros(S2, 1) *// смещение*

*// Моделирование слоя 2*

n2 = w2\*a1 + repmat(b2, 1, size(a1, 2)) *// сетевые значения*

a2 = ann\_tansig\_activ(n2); *// активность на выходе 2-го слоя*

d\_a2 = ann\_d\_tansig\_activ(a2); *// производная активационной функции*

*// Вычисление статистик слоя 2*

mean\_a2 = mean(a2) *// среднее значение функции*

stdev\_a2 = stdev(a2) *// стандартное отклонение*

mean\_d\_a2 = mean(d\_a2) *// среднее значение производной*

stdev\_d\_a2 = stdev(d\_a2) *// стандартное отклонение*

*// Графики первого слоя*

clf(1);

figure(1);

subplot(2, 2, 1)

plot(t, a1(1,:), t, a1(2,:));

title('Активность нейронов слоя 1')

subplot(2, 2, 2)

plot(t, d\_a1(1,:), t, d\_a1(2,:));

title('Производные функции активации слоя 1')

subplot(2, 2, 3)

histplot(20, a1);

title('Гистограмма активности нейронов слоя 1')

subplot(2, 2, 4)

histplot(20, d\_a1);

title('Гистограмма производных функции tansig')

*// Графики второго слоя*

clf(2);

figure(2);

subplot(2, 2, 1)

plot(t, a2(1,:));

title('Активность нейронов слоя 2')

subplot(2, 2, 2)

plot(t, d\_a2(1,:));

title('Производные функции активации слоя 2')

subplot(2, 2, 3)

histplot(20, a2);

title('Гистограмма активности нейронов слоя 2')

subplot(2, 2, 4)

histplot(20, d\_a2);

title('Гистограмма производных функции tansig')

Листинг 3 – Исследование сети с активационной функцией purelin

*// Задание исходных данных*

F = 0.01; *// минимальная частота входного сигнала*

R = 100; *// число входов слоя*

S1 = 10; *// число нейронов слоя 1*

S2 = 1; *// число нейронов слоя 2*

*// Формирование матрицы входных сигналов*

t = 0:0.1:2/F *// отсчёты времени от 0 до 2/F*

fi = (2\*%pi\*F).\*t *// аргументы синуса*

p = [sin(fi)]; *// входные сигналы для 1-го входа*

for i = 2:R *// формирование входных сигналов для i-х входов*

p = [p; sin(fi\*i)];

end

*// Инициализация весов и смещений для слоя 1*

scale = 1/sqrt(R); *// масштабирование*

*//scale = 10;*

w1 = grand(S1, R, "unf", -scale, scale); *// вес*

b1 = zeros(S1, 1); *// смещение*

*// Моделирование слоя 1*

n1 = w1\*p + repmat(b1, 1, size(p, 2)); *// сетевые значения*

a1 = ann\_purelin\_activ(n1); *// активность на выходе 1-го слоя*

d\_a1 = ann\_d\_purelin\_activ(a1); *// производная активационной функции*

*// Вычисление статистик слоя 1*

mean\_a1 = mean(a1); *// среднее значение функции*

stdev\_a1 = stdev(a1); *// стандартное отклонение*

mean\_d\_a1 = mean(d\_a1); *// среднее значение производной*

stdev\_d\_a1 = stdev(d\_a1); *// стандартное отклонение*

*// Инициализация весов и смещений для слоя 2*

scale = 1/sqrt(S1) *// масштабирование*

*//scale = 10;*

w2 = grand(S2, S1, "unf", -scale, scale); *// вес*

b2 = zeros(S2, 1) *// смещение*

*// Моделирование слоя 2*

n2 = w2\*a1 + repmat(b2, 1, size(a1, 2)) *// сетевые значения*

a2 = ann\_purelin\_activ(n2); *// активность на выходе 2-го слоя*

d\_a2 = ann\_d\_purelin\_activ(a2); *// производная активационной функции*

*// Вычисление статистик слоя 2*

mean\_a2 = mean(a2) *// среднее значение функции*

stdev\_a2 = stdev(a2) *// стандартное отклонение*

mean\_d\_a2 = mean(d\_a2) *// среднее значение производной*

stdev\_d\_a2 = stdev(d\_a2) *// стандартное отклонение*

*// Графики первого слоя*

clf(1);

figure(1);

subplot(2, 2, 1)

plot(t, a1(1,:), t, a1(2,:));

title('Активность нейронов слоя 1')

subplot(2, 2, 2)

plot(t, d\_a1(1,:), t, d\_a1(2,:));

title('Производные функции активации слоя 1')

subplot(2, 2, 3)

histplot(20, a1);

title('Гистограмма активности нейронов слоя 1')

subplot(2, 2, 4)

histplot(20, d\_a1);

title('Гистограмма производных функции purelin')

*// Графики второго слоя*

clf(2);

figure(2);

subplot(2, 2, 1)

plot(t, a2(1,:));

title('Активность нейронов слоя 2')

subplot(2, 2, 2)

plot(t, d\_a2(1,:));

title('Производные функции активации слоя 2')

subplot(2, 2, 3)

histplot(20, a2);

title('Гистограмма активности нейронов слоя 2')

subplot(2, 2, 4)

histplot(20, d\_a2);

title('Гистограмма производных функции purelin')

# ХОД работы

Был построен график функции tansig и её производной (Рисунок 1). Тангенциальная сигмовидная функция видоизменяется от -1 до 1, производная локализуется в области [-4; 4] и достигает максимума при .

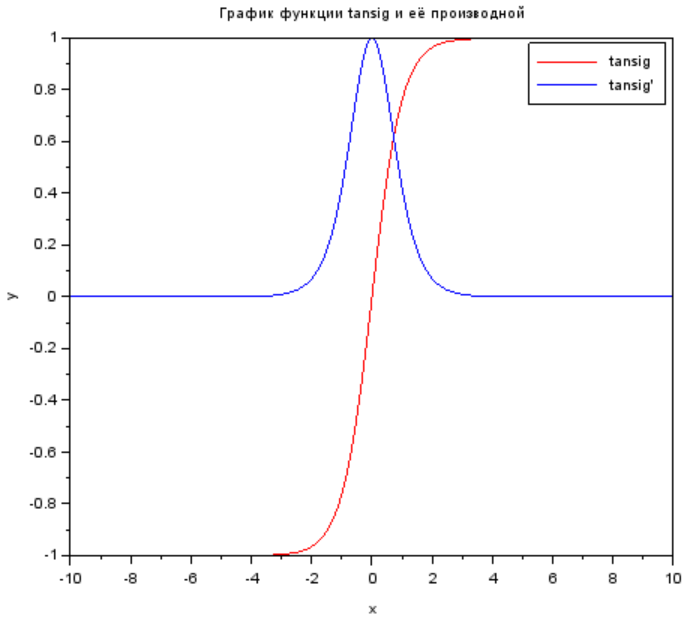


Рисунок 1 – График функции tansig и её производной

Был построен график линейной функции purelin и её производной (Рисунок 2). Производная функции всюду равна 1.

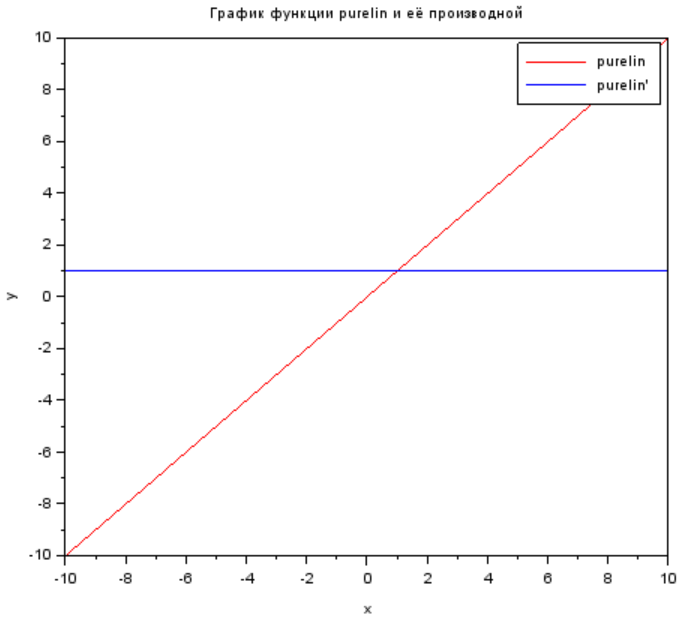


Рисунок 2 – График функции purelin и её производной

Было написано два программных блока, моделирующих сети прямого распространения с заданными по варианту параметрами. Сначала была смоделирована сеть с активационной функцией tansig.

Были выведены графики активности нейронов сети при инициализации её случайными значениями из диапазона [-10; 10] (scale = 10).

Для первого слоя: практически все активности нейронов равны -1 или 1 – двум крайним значениям, полученная производная чаще всего равна 0 (Рисунок 3).

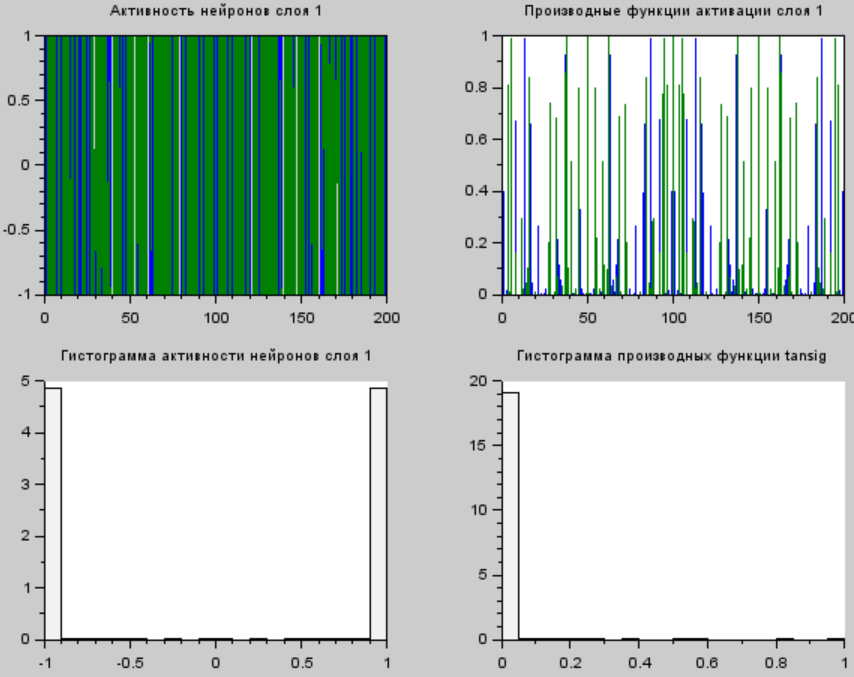


Рисунок 3 − Графики активности нейронов первого слоя (tansig, scale =10)

Для второго слоя ситуация аналогичная, активность равна 1 или -1, производная равна 0 (Рисунок 4).

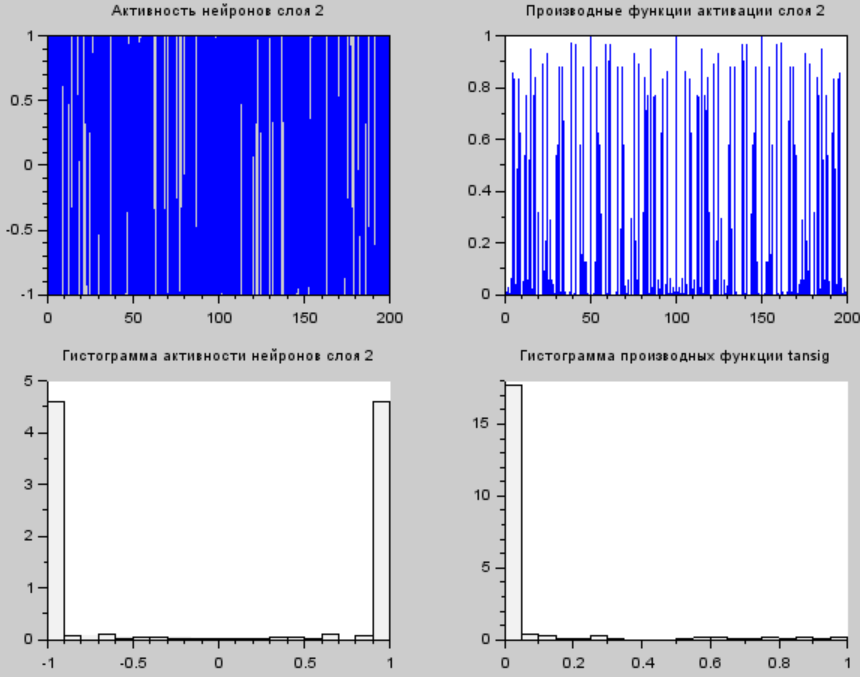


Рисунок 4 − Графики активности нейронов второго слоя (tansig, scale =10)

Затем были выведены графики активности нейронов той же сети при масштабировании обратно пропорциональном числу входов (scale = 1/sqrt(R)).

Для первого слоя: активности нейронов распределены вокруг 0 в диапазоне [-1; 1], полученная производная приближена к 1 (Рисунок 5).

Для второго слоя ситуация схожа: активность лежит в диапазоне [-0,5; 0,5], производная всегда больше нуля (Рисунок 6).

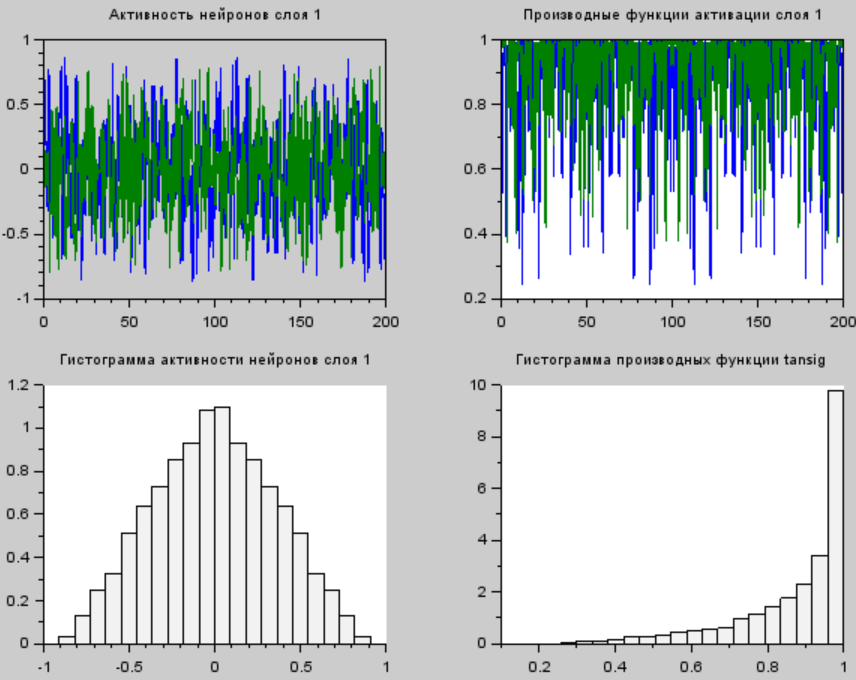


Рисунок 5 − Графики активности нейронов первого слоя (tansig, scale =1/sqrt(R))

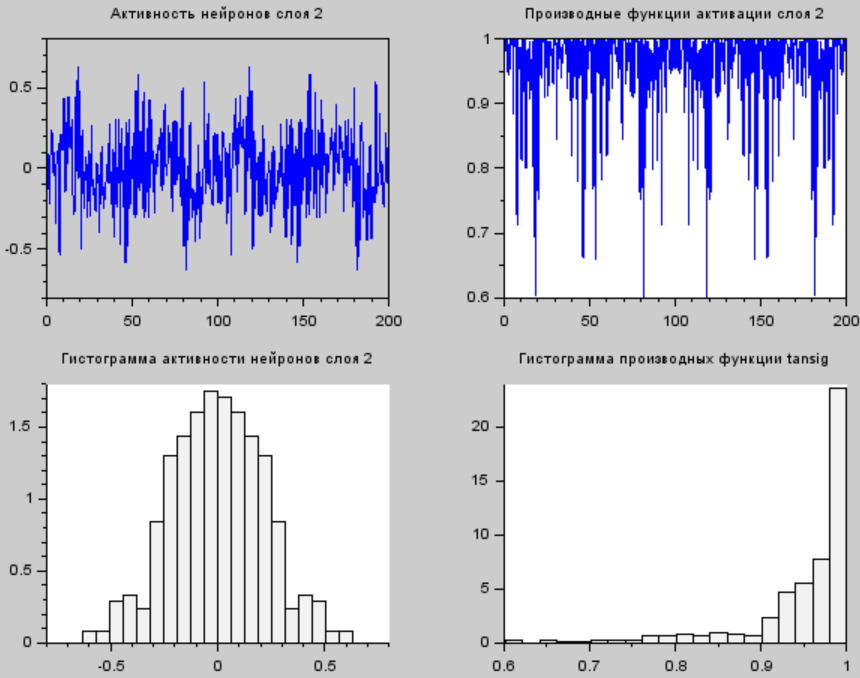


Рисунок 6 − Графики активности нейронов второго слоя (tansig, scale =1/sqrt(R))

Была рассмотрена сеть с активационной функцией purelin при инициализации в промежутке [-10; 10].

Для первого слоя: активности распределены в диапазоне [-150; 150], производная строго равна 1 (Рисунок 7).

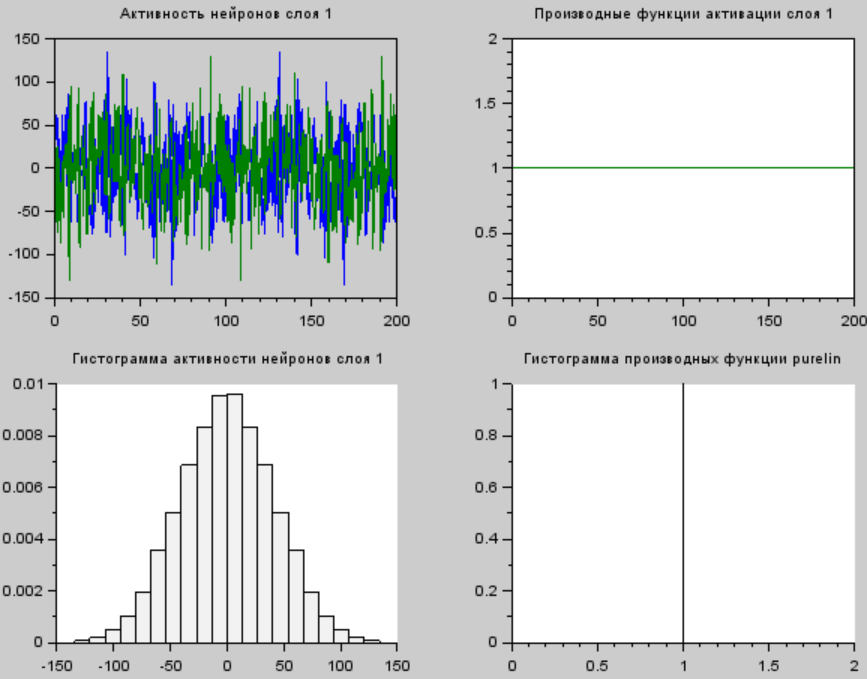


Рисунок 7 − Графики активности нейронов первого слоя (purelin, scale =10)

Для второго слоя диапазон активности увеличился до [-2000; 2000], производная всё так же равна 1 (Рисунок 8).

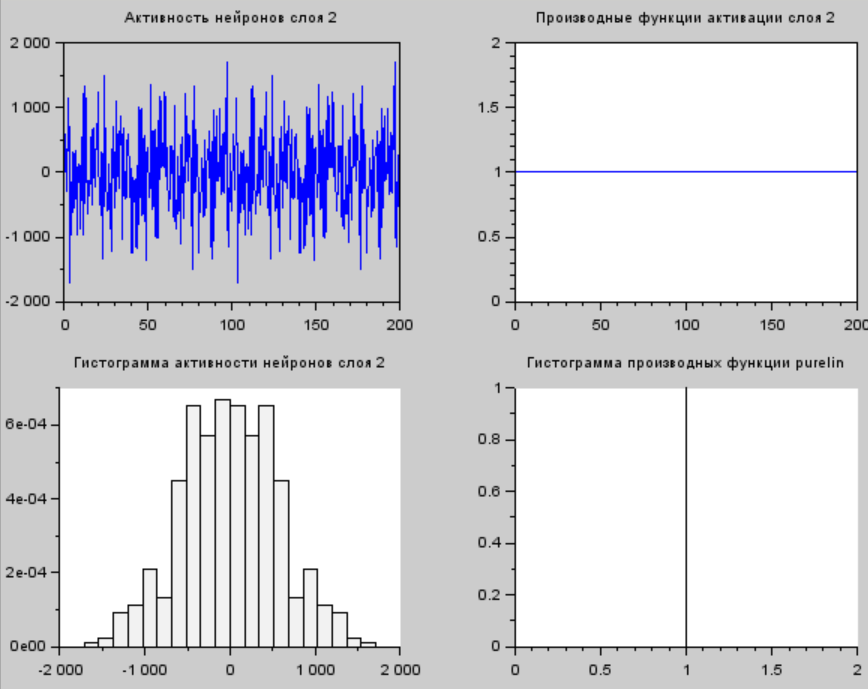


Рисунок 8 − Графики активности нейронов второго слоя (purelin, scale =10)

При настройке масштабирования (scale = 1/sqrt(R)) значения на графиках изменились. Так для первого слоя диапазон изменения активности нейронов снизился до [-1,5; 1,5], производная всё так же осталась равна 1 (Рисунок 9).

Для второго слоя: активность лежит в диапазоне [-0,5; 0,5], производная строго равна 1 (Рисунок 1).

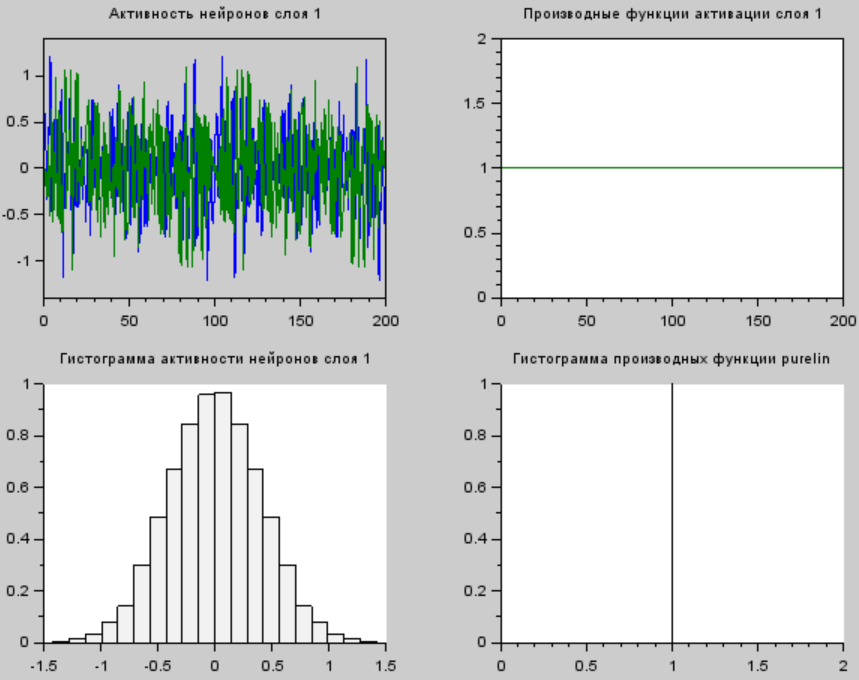


Рисунок 9 − Графики активности нейронов первого слоя (purelin, scale =1/sqrt(R))

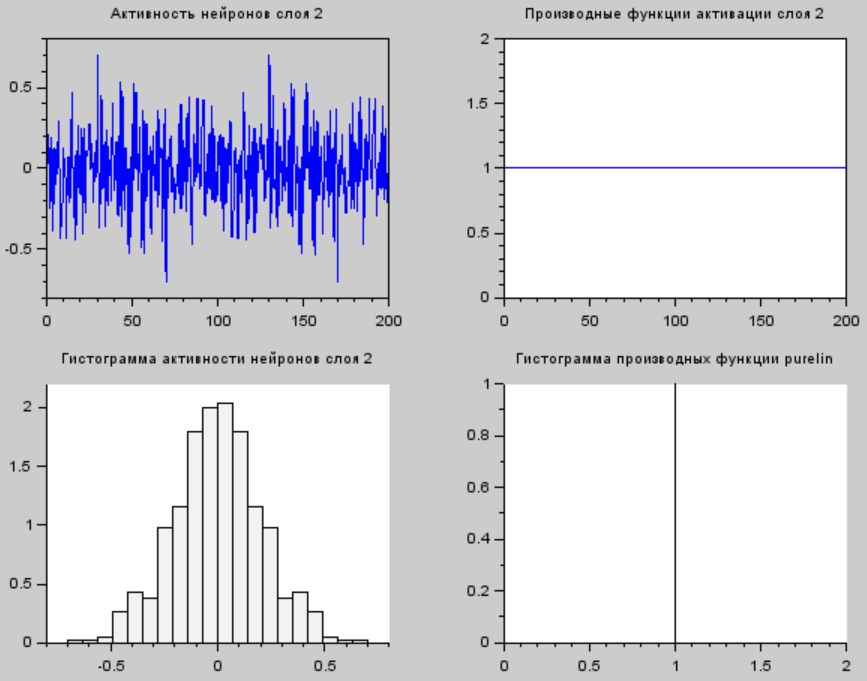


Рисунок 10 − Графики активности нейронов второго слоя (purelin, scale =1/sqrt(R))

# ВЫВОД

В ходе работы было выполнено моделирование двухслойных нейронных сетей прямого распределения, произведён анализ активационных функций и их производных для каждого слоя сетей. Найденные значения активности нейронов на выходах каждого слоя и производных были визуализированы с помощью графиков и гистограмм, что позволило провести анализ активационных функций.

Функция tansig реализует нелинейную зависимость между входами и выходами нейронной сети и подходит для ограничения активности нейронов в диапазоне [-1; 1]. При её использовании необходимо учитывать области насыщения – ниже -4 и выше 4.

Функция purelin реализует линейную активацию нейронов. Она не изменяет форму входных данных, принимает на вход любые значения и локализирована на всей числовой прямой.

Также был изучен способ настройки нейронной сети путём инициализации её матрицы W с масштабированием относительно числа входов нейрона. Он позволяет уменьшать сетевые значения нейронов так, чтобы все соответствующие им значения активационной функции лежали в области, где функция локализована. Это может быть полезным при решении задач, где необходимо подавать на выходы нейронов значения, ограниченный некоторой областью.