Лабораторная работа №2

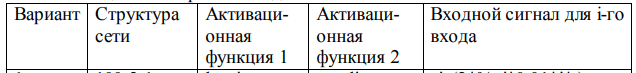
Исследование активационных функций нейронных элементов

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.

Ход работы

Вариант 4





3. (Построить графики активационных функций и их производных в соответствии с вариантом задания, используя соответствующие встроенные функции модуля NeuralNetworks 2.0 пакета Scilab. В случае отсутствия указанных встроенных функции определить их самостоятельно на основе выражений, приведенных в таблице 2.1.)

Описаны функции poslin и logsig (аналог ann\_logsig\_activ).

function **a**=poslin(**n**)

**a** = **n**

for i=1:length(**n**)

if **a**(i) < 0 then

**a**(i) = 0

end

end

endfunction

function **y**=logsig(**x**)

**y** = 1 ./ (1 + exp(-**x**))

endfunction

Описаны производные активационных функций.

function **y**=d\_poslin(**x**)

**y** = zeros(**x**)

for i=1:length(**y**)

if **x**(i) >= 0 then

**y**(i) = 1

end

end

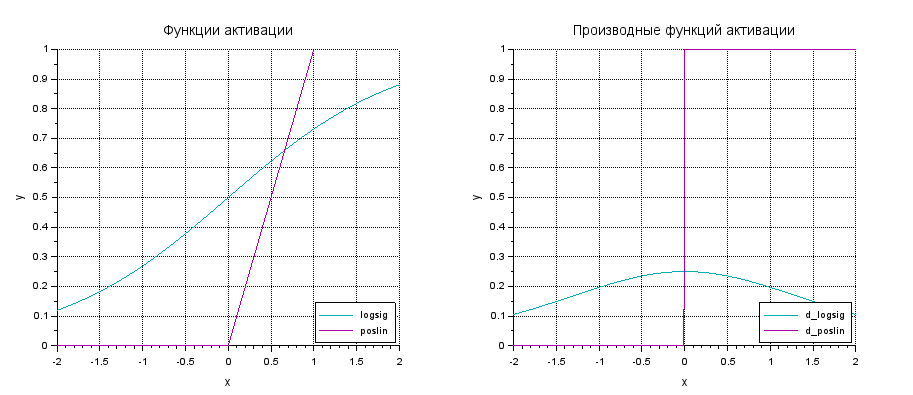
endfunction

function **y**=d\_logsig(**x**)

**y** = logsig(**x**) .\* (1 - logsig(**x**))

endfunction

Построены графики активационных функций и их производных.



4. (Реализовать две сети прямого распространения в соответствии с заданной структурой, запрограммировав вычисления в соответствии с выражением (2.5) для двух видов активационных функций, заданных по варианту.)

Реализованы две сети прямого распространения для двух видов активационных функций.

Для logsig:

*// количество нейронов в слоях*

s1 = 100

s2 = 8

s3 = 1

*// размер входного вектора*

R = 5

*// количество входных векторов*

N = 10

rand("seed", 42)

*// генерация входа*

P = rand(R, N)

function **a**=ff\_logsig(**W**, **P**, **b**)

*// прямое распространение*

**a** = logsig(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=init\_nn\_layer(**P**, **S**)

*// инициализация слоя*

W = rand(**S**, size(**P**, 1))

b = zeros(**S**, 1)

**a** = ff\_logsig(W, **P**, b)

endfunction

a1 = init\_nn\_layer(P, s1)

a2 = init\_nn\_layer(a1, s2)

a3 = init\_nn\_layer(a2, s3)

Для poslin:

*// количество нейронов в слоях*

s1 = 100

s2 = 8

s3 = 1

*// размер входного вектора*

R = 5

*// количество входных векторов*

N = 10

rand("seed", 42)

*// генерация входа*

P = rand(R, N)

function **a**=ff\_poslin(**W**, **P**, **b**)

*// прямое распространение*

**a** = poslin(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=init\_nn\_layer(**P**, **S**)

*// инициализация слоя*

W = rand(**S**, size(**P**, 1))

b = zeros(**S**, 1)

**a** = ff\_poslin(W, **P**, b)

endfunction

a1 = init\_nn\_layer(P, s1)

a2 = init\_nn\_layer(a1, s2)

a3 = init\_nn\_layer(a2, s3)

5. (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 (например, для варианта 1 F=0.01Гц). 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.)

Выполнено моделирование двух нейросетей на сгенерированных сигналах с разной инициализацией весов.

*// количество нейронов в слоях*

s1 = 100

s2 = 8

s3 = 1

*// размер входного вектора*

R = 5

*// минимальная частота входного сигнала*

F = 0.04

*//формирование матрицы p входных сигналов*

t = (0:0.1:2/F)

fi = 2 \* %pi \* F .\* t

P = sin(fi)

for i=2:R

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

end

grand("setsd", 42)

function **a**=ff\_logsig(**W**, **P**, **b**)

*// прямое распространение*

**a** = logsig(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=ff\_poslin(**W**, **P**, **b**)

*// прямое распространение*

**a** = poslin(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=init\_nn\_layer(**P**, **S**)

*// инициализация весов и смещений*

R = size(**P**, 1)

*// W = grand(S,R,'unf',-10,10)*

scale = 1/sqrt(R);

W = grand(**S**,R,'unf',-scale,scale)

b = zeros(**S**, 1)

**a** = ff\_logsig(W, **P**, b)

*// a = ff\_poslin(W, P, b)*

endfunction

a1 = init\_nn\_layer(P, s1)

a2 = init\_nn\_layer(a1, s2)

a3 = init\_nn\_layer(a2, s3)

6. (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.)

Вычислены выходные значения всех нейронов сети.

*// количество нейронов в слоях*

s1 = 100

s2 = 8

s3 = 1

*// размер входного вектора*

R = 5

*// минимальная частота входного сигнала*

F = 0.04

*//формирование матрицы p входных сигналов*

t = (0:0.1:2/F)

fi = 2 \* %pi \* F .\* t

P = sin(fi)

for i=2:R

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

end

grand("setsd", 42)

function **a**=ff\_logsig(**W**, **P**, **b**)

*// прямое распространение*

**a** = logsig(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=ff\_poslin(**W**, **P**, **b**)

*// прямое распространение*

**a** = poslin(**W** \* **P** + repmat(**b**, 1, size(**P**, 2)))

endfunction

function **a**=init\_nn\_layer(**P**, **S**)

*// инициализация весов и смещений*

R = size(**P**, 1)

*// W = grand(S,R,'unf',-10,10)*

scale = 1/sqrt(R);

W = grand(**S**,R,'unf',-scale,scale)

b = zeros(**S**, 1)

**a** = ff\_logsig(W, **P**, b)

*// a = ff\_poslin(W, P, b)*

endfunction

a1 = init\_nn\_layer(P, s1)

a2 = init\_nn\_layer(a1, s2)

a3 = init\_nn\_layer(a2, s3)

Построены указанные графики.

clf(1)

figure(1)

subplot(221)

*// активность 1-ого нейрона скрытого слоя*

plot(t, a2(1,:))

xgrid

xtitle("Активность нейрона скрытого слоя")

subplot(222)

*// активность нейрона выходного слоя*

plot(t, a3(1,:))

xgrid

xtitle("Активность нейрона выходного слоя")

subplot(223)

*// гистограмма общей активности нейронов скрытого слоя*

histplot(20, a2, style=17)

xgrid

xtitle("Общая активность нейронов скрытого слоя")

subplot(224)

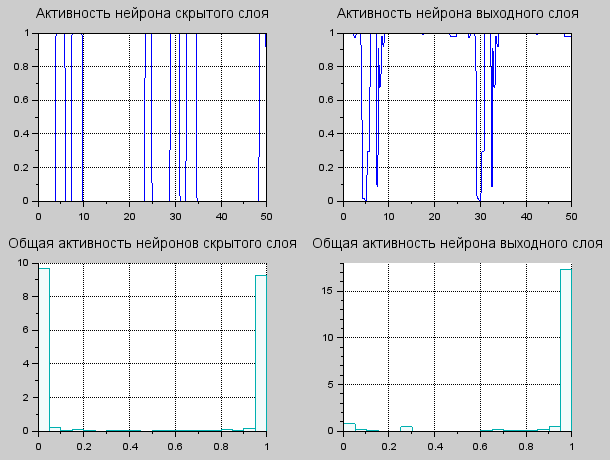
*// гистограмма общей активности нейрона выходного слоя*

histplot(20, a3, style=17)

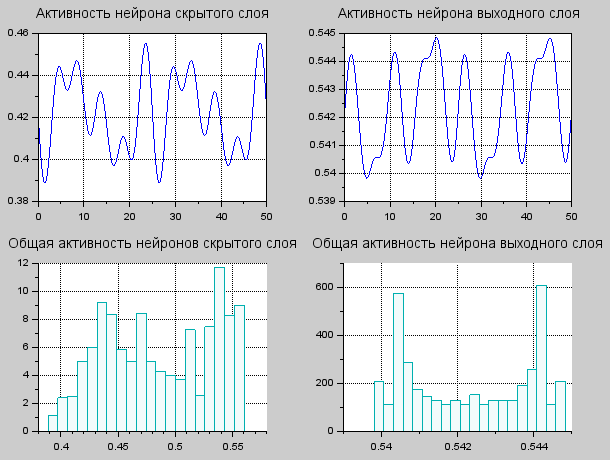
xgrid

xtitle("Общая активность нейрона выходного слоя")

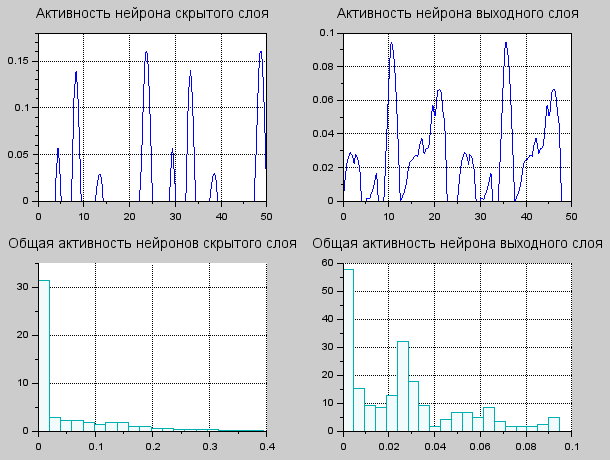
Функция активации logsig, метод инициализации 1.

****

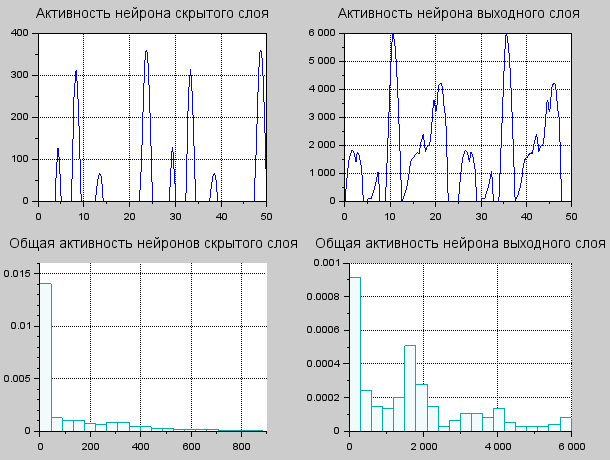
Функция активации logsig, метод инициализации 2.



Функция активации poslin, метод инициализации 1.

****

Функция активации poslin, метод инициализации 2.



Вычислены статистики.

*// статистики скрытого слоя*

hl\_mean = mean(a2)

hl\_sd = stdev(a2)

printf("Средняя активность нейронов скрытого слоя: %f\n", hl\_mean)

printf("Стандартное отклонение нейронов скрытого слоя: %f\n", hl\_sd)

*// статистики выходного слоя*

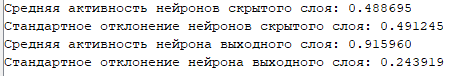
ol\_mean = mean(a3)

ol\_sd = stdev(a3)

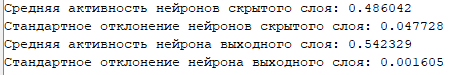
printf("Средняя активность нейрона выходного слоя: %f\n", ol\_mean)

printf("Стандартное отклонение нейрона выходного слоя: %f\n", ol\_sd)

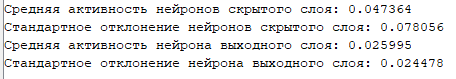
Функция активации logsig, метод инициализации 1.



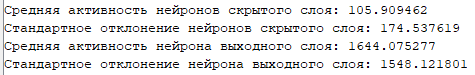
Функция активации logsig, метод инициализации 2.



Функция активации poslin, метод инициализации 1.



Функция активации poslin, метод инициализации 2.



7. (Вычислить значения производных нелинейностей для всех нейронов. Построить графики значений производных нелинейности для одного из нейронов скрытого слоя и нейрона выходного слоя, а также гистограмму производных функции активации скрытых нейронов и выходного (для двух способов инициализации и для каждой из двух реализованных сетей). Вычислить среднее значение и стандартное отклонение производных для нейронов скрытого и выходного слоёв)

Вычислены значения производных нелинейностей для всех нейронов.

d\_a1 = d\_logsig(a1)

d\_a2 = d\_logsig(a2)

d\_a3 = d\_logsig(a3)

*//d\_a1 = d\_poslin(a1)*

*//d\_a2 = d\_poslin(a2)*

*//d\_a3 = d\_poslin(a3)*

Построены графики значений производных нелинейности для одного из нейронов скрытого слоя и нейрона выходного слоя, а также гистограммы производных функции активации скрытых нейронов и выходного (для двух способов инициализации и для каждой из двух реализованных сетей).

subplot(221)

*// активность 1-ого нейрона скрытого слоя*

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

xgrid

xtitle("Производные для нейрона скрытого слоя")

subplot(222)

*// активность нейрона выходного слоя*

plot(t, d\_a3(1,:))

xgrid

xtitle("Производные для нейрона выходного слоя")

subplot(223)

*// гистограмма общей активности нейронов скрытого слоя*

histplot(20, d\_a2, style=17)

xgrid

xtitle("Производные для нейронов скрытого слоя")

subplot(224)

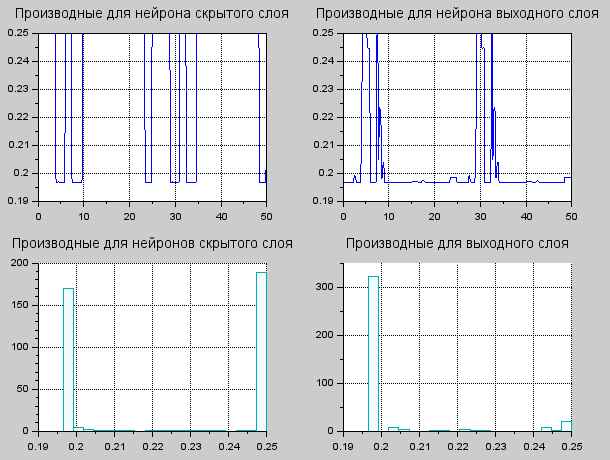
*// гистограмма общей активности нейрона выходного слоя*

histplot(20, d\_a3, style=17)

xgrid

xtitle("Производные для выходного слоя")

Функция активации logsig, 1-ый способ инициализации.

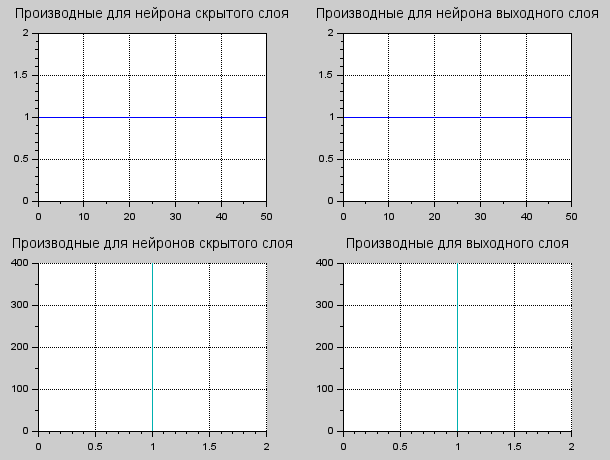


Функция активации logsig, 2-ый способ инициализации.

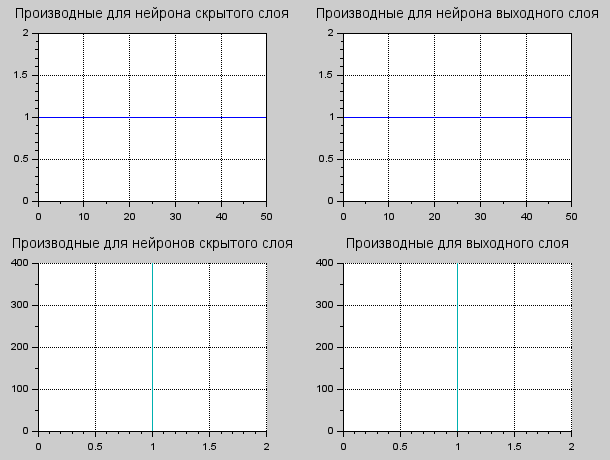


Функция активации poslin, 1-ый способ инициализации.

(Функция poslin возвращает нули и положительные числа, d\_poslin их все превращает в 1)



Функция активации poslin, 2-ый способ инициализации.



Вычислены среднее значение и стандартное отклонение производных для нейронов скрытого и выходного слоёв.

*// статистики скрытого слоя*

hl\_mean = mean(d\_a2)

hl\_sd = stdev(d\_a2)

printf("Среднее значение производных скрытого слоя: %f\n", hl\_mean)

printf("Стандартное отклонение производных скрытого слоя: %f\n", hl\_sd)

*// статистики выходного слоя*

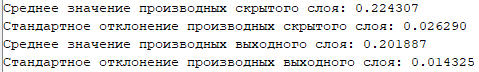
ol\_mean = mean(d\_a3)

ol\_sd = stdev(d\_a3)

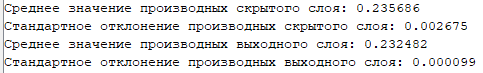
printf("Среднее значение производных выходного слоя: %f\n", ol\_mean)

printf("Стандартное отклонение производных выходного слоя: %f\n", ol\_sd)

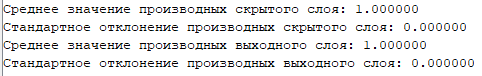
Функция активации logsig, 1-ый способ инициализации.



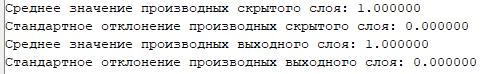
Функция активации logsig, 2-ый способ инициализации.



Функция активации poslin, 1-ый способ инициализации.



Функция активации poslin, 2-ый способ инициализации.



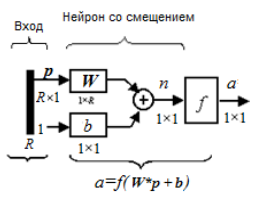
8. (Выполнить анализ полученных результатов, обратив внимание на характер гистограмм. Сделать выводы о характере распределения значений активностей и производных для заданных активационных функций и каждого из способов инициализации. Сравнить результаты для каждой из двух реализованных сетей.)

Вывод

Выполняя эту работу, я углубила теоретические знания в области архитектуры нейронных сетей, исследовала свойства активационных функций нейронных элементов, приобрела практические навыки моделирования простейшей нейронной сети прямого распространения.

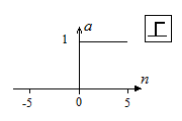
Ответы на контрольные вопросы

1. *Нарисуйте схему нейронного элемента с векторным входом и запишите выражение для вычисления выходного значения.*

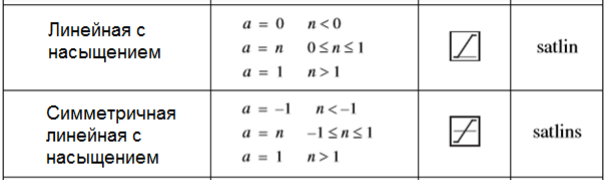


2. *Приведите выражение функции активации с жестким ограничением (функция Хевисайда) и ее графическое представление.*

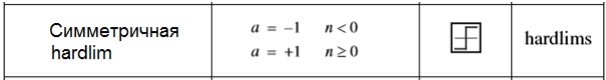
a = 0, если n < 0, и 1, если n > 0.



3. *Приведите выражения для линейных функций активации (satlin, satlins) с насыщением и их графическое представление.*

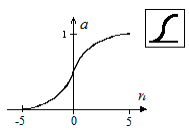


4. *Приведите выражение для знаковой функции активации и ее графическое представление.*

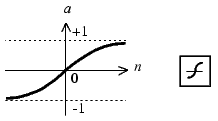


5. *Приведите выражения для униполярной и биполярной сигмовидных функций активации и их графическое представление.*

Униполярная: Логистическая (сигмовидная) функция активации (logsig). Эта функция описывается соотношением a = 1/(1 + exp(–n)).



Биполярная: Тангенциальная (сигмовидная) функция активации (tansig). Эта функция описывается соотношением a = (exp(n)-(exp(-n))/(exp(n) + exp(–n)).



6. *Приведите выражение для выходных сигналов НЭ с квадратической радиальной функцией.*

(чего-чего?)

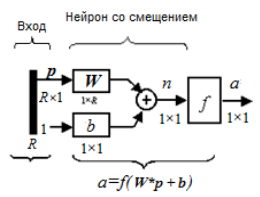
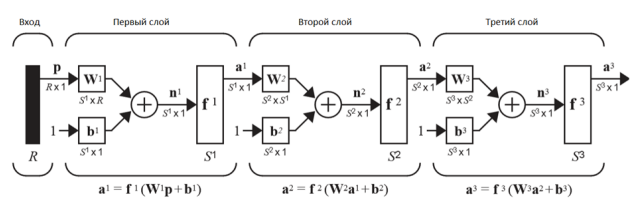
7. *Нарисуйте схему и приведите аналитическое выражение функции преобразования для линейного порогового элемента.*

(опять непонятно)

8. *Продемонстрируйте на компьютере, каким образом проводились исследования функций активации?*

-

9. *Нарисуйте структурные схемы однослойной и многослойной искусственных нейронных сетей с прямыми связями.*

10. *Поясните понятия входного, скрытого и выходного слоев. Как сокращенно обозначают структуру сетей прямого распространения?*

Входной слой – первый слой нейронной сети, на который поступают входные данные.

Скрытый слой – промежуточный слой, данные на который поступают из предыдущего слоя нейронной сети, а выходят в следующий слой.

Выходной слой – последний слой нейронной сети, который выдает окончательные данные.

Сокращенно структура нейронной сети описывается количеством нейронов на каждом слое, записанными через дефис.

11. *Приведите выражения для вычисления производных активационных функций, указанных в таблице 2.1.*

Hardlim: d = 0

Satlin: d = 0, если n < 0; 1, если 0≤n≤1; 0, если 1≤n.

Satlins: d = 0, если n < -1; 1, если -1≤n≤1; 0, если 1≤n.

Poslin: d = 0, если n<0; 1, если 0≤n.

Compet: d = 0.

12. *Используя правила дифференцирования, выведите выражения для вычисления производных логистической и тангенциальной сигмовидных активационных функций.*

13. *Что такое иницализация нейронной сети и как её выполняют? Приведите примеры кода на языке Scilab.*

Под инициализацией нейронной сети подразумевают присвоение значений весов и смещений. Как правило, значения весов выбираются случайные, а смещения нулевые.

W = rand(S, R)

b = zeros(S, 1)

14. *Нарисуйте структурную схему 3-х слойной сети прямого распространения и запишите на языке Scilab фрагмент кода для вычисления её выходного значения, если последовательность входных векторов p представляется в виде матрицы.*

В коде выше есть

15. *Какой желательный характер распределений должны иметь выходные значения нейронов нейросети и почему? Как этого добиваются?*

16. *Почему нежелательно, чтобы производные активационных функций принимали нулевые значения?*

Это приводит к потере эффективности градиентных алгоритмов обучения.