ФГАОУ ВО «Севастопольский государственный университет»

Институт информационных технологий   
и управления в технических системах

ОТЧЁТ  
ПО ЛАБОРАТОРНОЙ РАБОТЕ №5

«Исследование многослойного персептрона: алгоритмы обратного распространения с адаптивной скоростью обучения и моментом»

по дисциплине «Нейрокомпьютерные технологии»

Выполнила:  
студентка группы ИС/м-21-1-з  
Ускова Екатерина Дмитриевна

Севастополь

2022

# ЦЕЛЬ РАБОТЫ

Углубление теоретических знаний в области архитектуры многослойных нейронных сетей прямого распространения, исследование свойств алгоритмов обучения многослойных нейронных сетей, приобретение практических навыков обучения многослойного персептрона при решении задач классификации и аппроксимации функций.

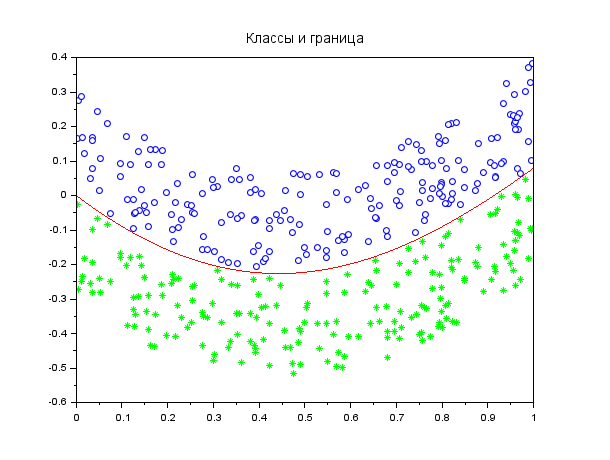
# Ход работы

Задана одномерная функция у=f(x)



Сформированы два множества (класса) точек данных (по 200 точек в каждом множестве), которые располагаются выше и ниже кривой у=f(x), и отстоят от неё на расстояние d=0,03|(ymax - ymin)|.

Классы и кривая у=f(x) отображены на двумерной плоскости.



Разработана программа обучения многослойного персептрона с архитектурой R-S-1 для классификации этих классов, в качестве активационных функций скрытого слоя использована функция ann\_logsig\_active, а в качестве активационной функции выходного слоя ─ ann\_purelin\_active, обучение персептрона выполненяется с использованием функции ann\_FFBP\_gd.

Выполнено предварительное обучение MLP с архитектурой [R-10-1] на небольшом числе эпох itermax =200 при разных значениях параметра скорости обучения lr с целью определения её квазиоптимального устойчивого значения.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| lr | 0.05 | 0.025 | 0.01 | 0.005 |
| MSE | 0.175164 | 0.185217 | 0.190841 | 0.192654 |

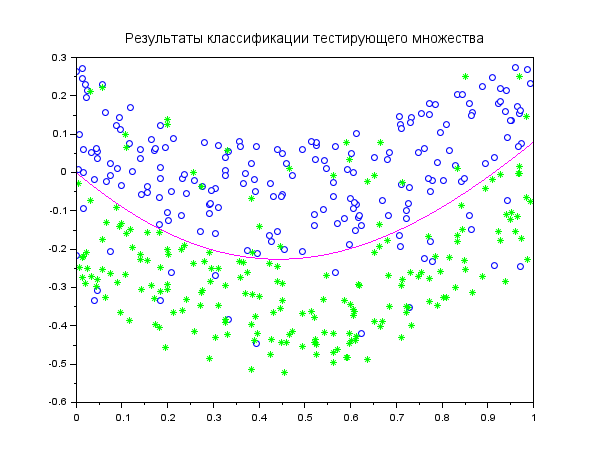
Квазиоптимальное устойчивое значение параметра скорости обучения равно 0.05.

Используя полученное значение скорости обучения lr, выполнено обучение MLP с архитектурой [R-S-1] при разных S (10,30,40,50) на большом числе эпох itermax =1000.

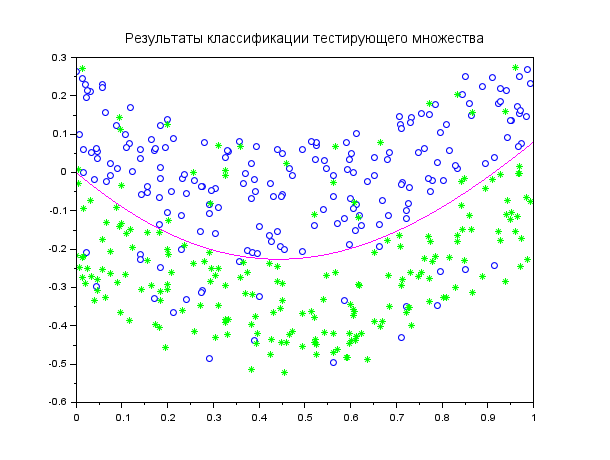
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S | 10 | 30 | 40 | 50 |
| MSE | 0.097353 | 0.093927 | 0.091752 | 0.092312 |

Наименьшая ошибка при S=40.

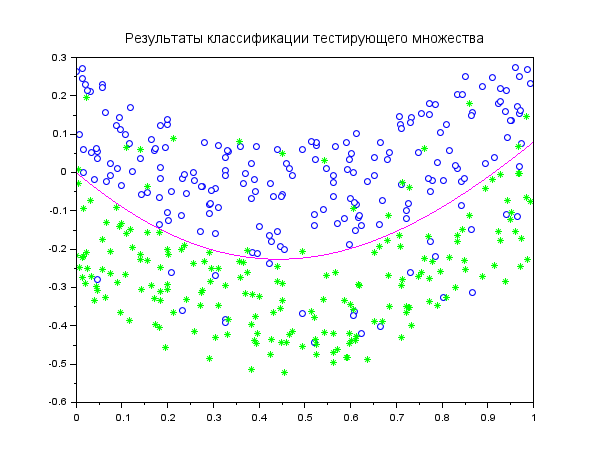
Для различных MLP, обученных в соответствии с предыдущим пунктом, выполнено моделирование MLP и проверено качество классификации на тестовом множестве данных. Для этого сформировано тестовое множество данных аналогично обучающему набору, проведена классификация данных с помощью обученного MLP, отображены точки полученных выходных классов и кривая у=f(x) на двумерной плоскости, вычислены вероятности правильной классификации при разных S.



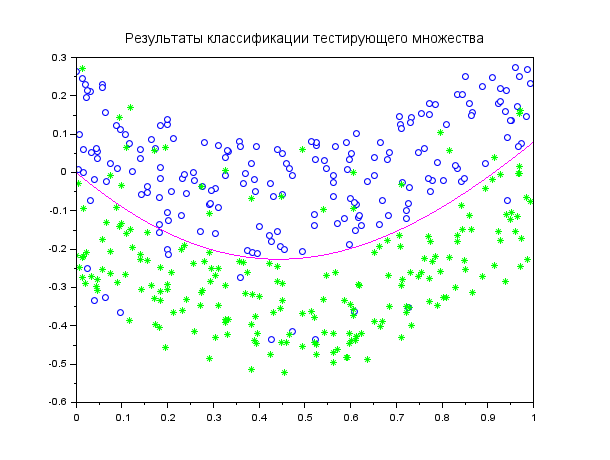
Результаты классификации тестирующего множества, вероятность правильной классификации 0.8775 (S=10)



Результаты классификации тестирующего множества, вероятность правильной классификации 0.8725 (S=30)



Результаты классификации тестирующего множества, вероятность правильной классификации 0.895 (S=40)



Результаты классификации тестирующего множества, вероятность правильной классификации 0.905 (S=50)

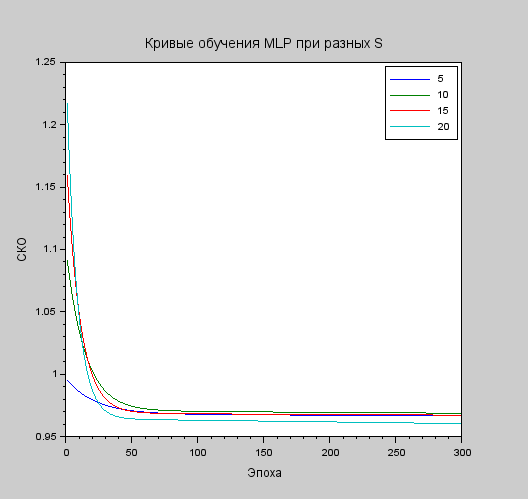
Задана двумерная функция z=f(x,y):



Сформировано подмножество обучающих и тестовых данных. Для этого на плоскости (x,y) выбраны 300 случайных точек и в этих точках определено значение функции z=f(x,y). В качестве входного вектора использован вектор p=[x;y], в качестве целевого значения – t=z. Полученные данные разделены на 2 подмножества: обучающее (80%) и тестовое (20%).

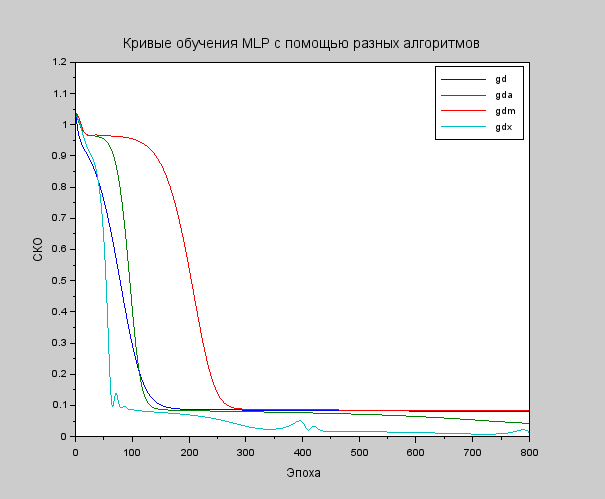
Разработана программа обучения многослойного персептрона с архитектурой R-S-1 для аппроксимации этой функции, в качестве активационной функции скрытого слоя использована функция ann\_tansig\_active, а в качестве активационной функции выходного слоя -- ann\_purelin\_active, обучение персептрона выполняется с использованием модифицированной функции ann\_FFBP\_gd.

Построены кривые обучения MLP при разных значениях S (5 10 15 20) и фиксированной скорости обучения lr (0.005); выбрано квазиоптимальное значение S для дальнейшего использования. Построение кривых выполнено при значении параметра itermax=300.



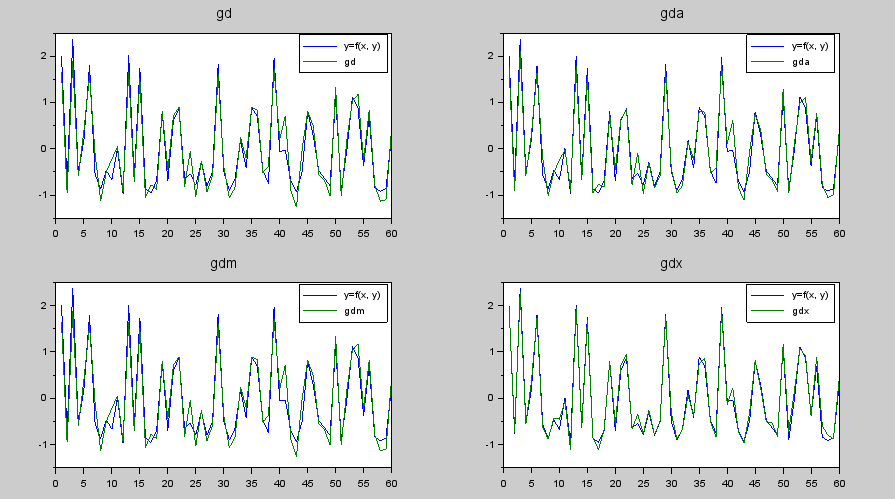
Квазиоптимальным будет значение S=20, так как оно обеспечивает наиболее быстрое снижение СКО и наименьшее конечное значение СКО в ходе обучения.

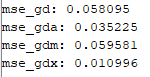
Выполнено обучение MLP (lr = 0.05, S = 20, itermax = 800) c помощью функций ann\_FFBP\_gdm, ann\_FFBP\_gdа, ann\_FFBP\_gdх.



Обучение с помощью функций ann\_FFBP\_gda и ann\_FFBP\_gdx даёт лучший результат. Можно полагать, что это связано с использованием адаптивной скорости обучения.

Используя тестовое подмножество данных, выполнено моделирование 4-х вариантов MLP, обученных с помощью 4-х разных функций, указанных выше. Построены сопоставительные графики значений функции z=f(x,y) и соответствующих значений на выходе MLP, вычислены СКО аппроксимации функции на тестовом подмножестве.





Из рисунка следует, что все алгоритмы дают примерно одинаковые результаты на выбранном числе эпох обучения. Возможно, эти алгоритмы требуют лучшего подбора параметров. Наименьшее значение СКО получается при обучении MLP с помощью алгоритма BP\_gdx.

СКО на обучающем наборе был значительно меньше, чем на тестовом.

# Вывод

Выполняя эту работу, я углубила теоретические знания в области архитектуры многослойных нейронных сетей прямого распространения, исследовала свойства алгоритмов обучения многослойных нейронных сетей, приобрела практические навыки обучения многослойного персептрона при решении задач классификации и аппроксимации функций.