Homework1

付宇辉 无05 2020010768

1. 证明：

首先记,如此

优化目的是要得到,即找到相应的，使得准确分类。

由于数据线性可分，则遇到误分类点的次数有上界，即存在使得所有数据点被准确分类。

使用随机梯度下降，会获得一系列,当变大，会越来越接近

用两个向量的夹角去衡量两个向量的相似程度（向量的长度并不重要）。

下面证明，有限次迭代内，会收敛

已知，存在使得,即存在比划分超平面更近的距离。

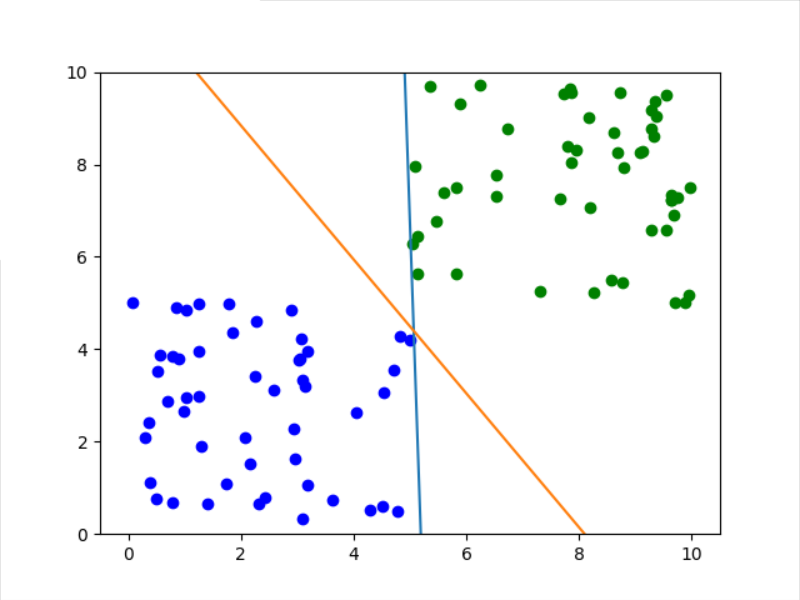
令

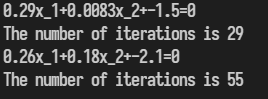
所以

得到

即证明了有限次可得到收敛的

2.使用numpy+matplotlib





附源代码：

import numpy as np

import matplotlib.pyplot as plt

def update(X, y ,w, b, lr):

    cnt = 0

    flag = 1

    while flag:

        flag = 0

        for i in range(100):

            if y[i] \* (w.dot(X[i]) + b) <= 0 :

                w += lr \* y[i] \* X[i]

                b += lr \* y[i]

                flag = 1

                cnt += 1

    print(f"{w[0]:4.2}x\_1+{w[1]:4.2}x\_2+{b:4.2}=0")

    print(f"The number of iterations is {cnt}")

    res\_x = np.linspace(0, 10, 500)

    res\_y = - (w[0] \* res\_x + b) / w[1]

    plt.plot(res\_x,res\_y)

if \_\_name\_\_ == '\_\_main\_\_':

    x\_n = np.random.uniform(0, 5, [50, 2])

    x\_p = np.random.uniform(5, 10, [50, 2])

    X = np.append(x\_n, x\_p).reshape((100,2))

    y = np.array([-1 if i < 50 else 1 for i in range(100)])

    plt.scatter(x\_n[:,0], x\_n[:,1], marker = 'o', color = 'blue')

    plt.scatter(x\_p[:,0], x\_p[:,1], marker = 'o', color = 'green')

    lr = 0.1

    w = np.zeros(2)

    b = 0

    update(X, y, w, b, lr)

    w = np.ones(2)

    b = 1

    update(X, y, w, b, lr)

    plt.ylim(0, 10)

    plt.show()