

Lecture2编程作业

PLA算法见: [PLA.py](#)

Pocket算法见: [pocket.py](#)

1. PLA和Pocket的核心算法:

PLA:

```
def train(self, data, label):  
    #训练样本集，寻找分类面  
    length=len(data)  
    y=np.zeros(length)  
    i=0  
    while i<length:  
        y[i]=np.sign(np.dot(self.w, data[i].T))  
        if y[i]!=label[i]:  
            self.w = self.w+label[i]*data[i].T  
            #若模型训练标签与真实标签不一致，则更正w  
            i=0  
        else:  
            i+=1
```

Pocket:

```
def train(self, data, label, num):  
    #训练模型，算法最大迭代次数为num  
    epoch=0  
    length=len(data)  
    y=np.zeros(length)  
    wrong=math.inf  
    while epoch<=num:  
        for i in range(length):  
            epoch+=1  
            y[i]=np.sign(np.dot(self.w, data[i].T))  
            if y[i]!=label[i]:  
                w_mid=self.w+label[i]*data[i].T  
                new_wrong=self.check(w_mid, data, label) #新的错分数量与之前最优的错分数  
                #比  
                if wrong>=new_wrong:  
                    #如果减少，则更换w  
                    wrong = new_wrong  
                    self.w=w_mid  
                    break
```

2.

(1) 数据集

产生两个都具有200个二维向量的数据集 X_1 和 X_2 。数据集 X_1 的样本来自均值向量 $m_1 = [-5, 0]^T$ ，协方差矩阵 $s_1 = I$ 的正态分布，属于“+1”类，数据集 X_2 的样本来自均值向量 $m_2 = [0, 5]^T$ ，协方差矩阵 $s_2 = I$ 的正态分布，属于“-1”类，其中 I 是一个2*2的单位矩阵。其中的数据中80%用于训练，20%用于测试。

(2) 训练集和测试集上，两种算法的分类正确率

PLA: $Accuracy_{(in)} = 1.0, Accuracy_{(out)} = 1.0$

Pocket: $Accuracy_{(in)} = 1.0, Accuracy_{(out)} = 1.0$

(3) 两种算法的运行时间

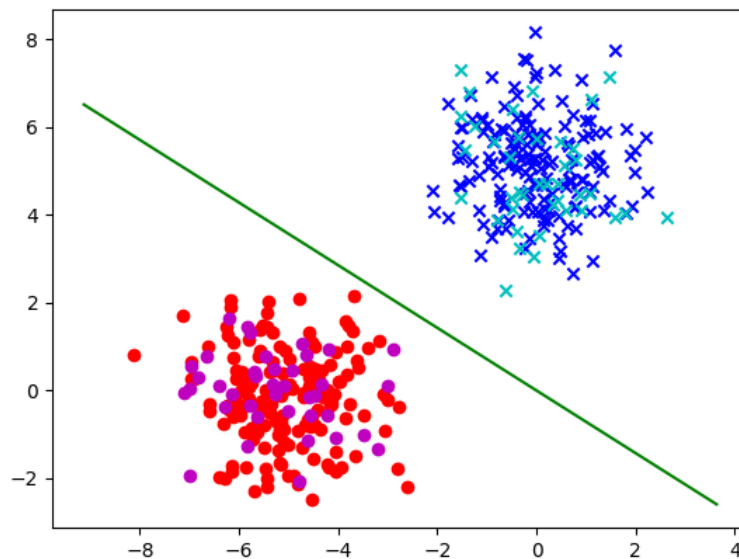
PLA: Running time: 0.00317 Seconds

Pocket: Running time: 0.32208 Seconds (迭代次数为100次)

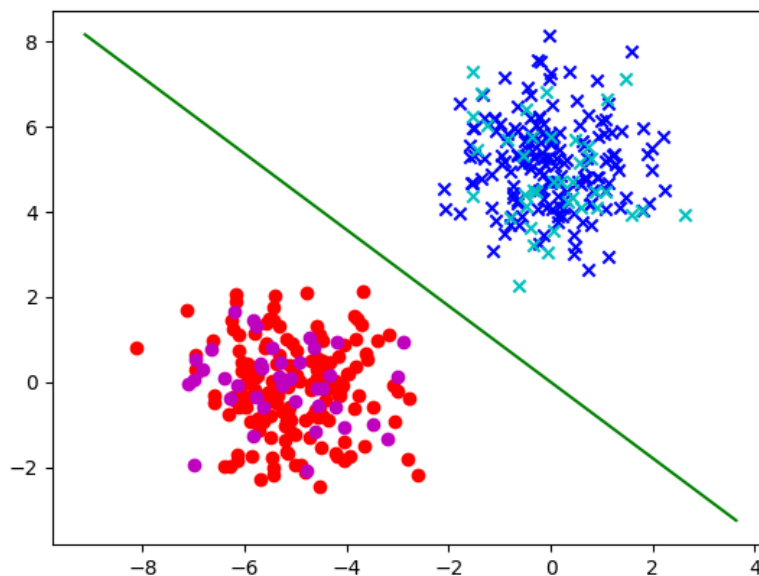
(4) 画出数据集和分类面

(关于作图，训练集的数据标签为“1”和“-1”类分别对应红色和蓝色的o点；测试集的数据标签为“1”和“-1”分别对应紫红色和青色的x点)

PLA:



Pocket: (迭代次数为100次)



3.

(1) 数据集

重复第2题的内容，但数据集 X_1 和数据集 X_2 的均值向量分别改为 $m_1 = [1, 0]^T$ 和 $m_2 = [0, 1]^T$ ，其他不变。

(2) 训练集和测试集上，两种算法的分类正确率

PLA: $Accuracy_{(in)} = 0.703125, Accuracy_{(out)} = 0.7125$ （由于数据不可分割，设迭代次数为100）

Pocket: $Accuracy_{(in)} = 0.79375, Accuracy_{(out)} = 0.725$ （迭代次数为100次）

(3) 两种算法的运行时间

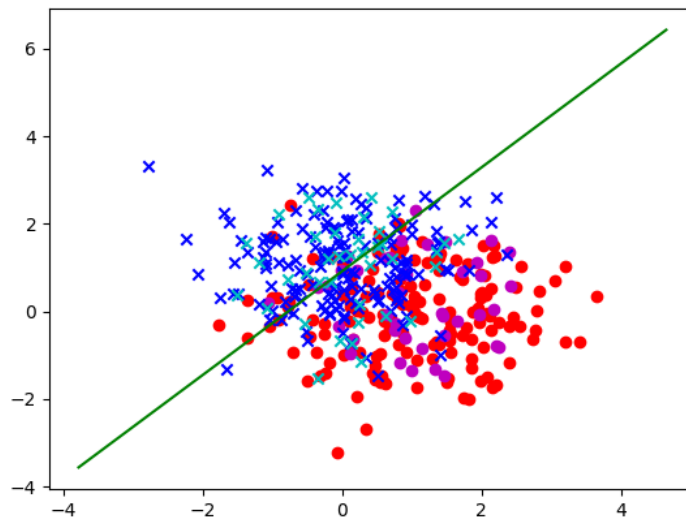
PLA: Running time: 3.85166 Seconds

Pocket: Running time: 9.12105 Seconds

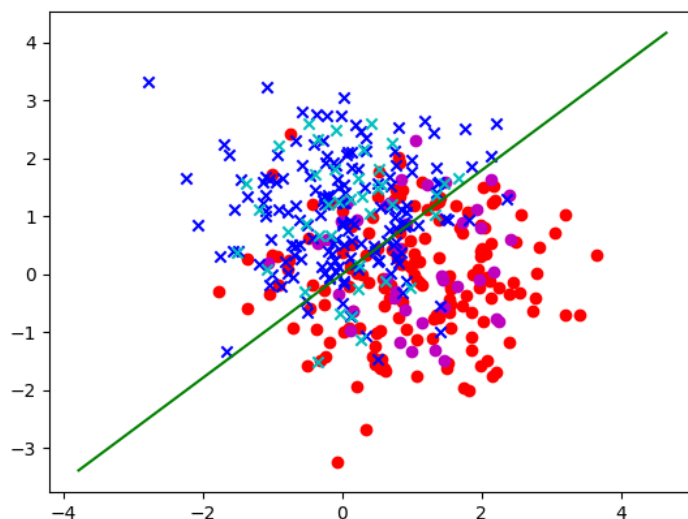
(4) 画出数据集和分类面

（关于作图，训练集的数据标签为“1”和“-1”类分别对应红色和蓝色的o点；测试集的数据标签为“1”和“-1”分别对应紫红色和青色的x点）

PLA:(由于数据不可分割，设迭代次数为100)



Pocket:(迭代次数为100)



4.改变算法中的各类超参数、样本数量、样本分布等，讨论实验结果。

该算法中的超参数为权重更新时的更新率 α , 更新式为: $w_{t+1} = w_t + \alpha y_n x_n$, 本算法中设置 $\alpha = 1$ 进行权重更新。当设置不同更新率时, 迭代收敛速度不同, 若更新率 α 设置过小时, 迭代速度可能过慢; 当设置过大时, 迭代速度可能会很快, 但是不容易收敛, 容易波动。

样本的数量对实验结果影响不大, 样本分布对实验结果影响较大。

感知器在样本集分布较离散可分时, 分类效果会更好。当样本重叠部分较多时, 感知器无法很好的对样本进行分类。若不设置迭代次数, PLA算法可能陷入循环。