**实验一：感知器学习到线性分类**

**15051207 王诗丽**

**2018/05/03**

**本实验所有代码链接在电子版文件里**

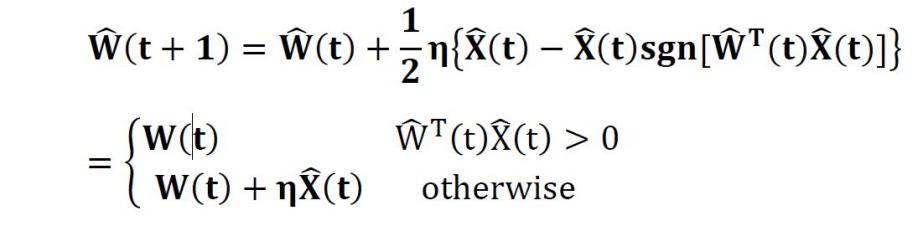
**一．简介**

线性感知器是两类分类器最简单的学习算法之一。给定一组数据点在d维空间,属于两类,ω1ω2,试图找到一个线性的算法分离两类样本之间的超平面。如果样本是在一个、两个或三个维度上，分离超平面将分别是点、线或平面。我们研究的具体算法是一种特殊的算法，它使用梯度下降法来确定一个经过仔细定义的目标函数，从而得出一个解。

**二．原理和理论**

假设两个类的样本在特征空间中是线性可分的。即，有一个平面G(X) = WTX + wn+1 = 0其中，所有属于第一类的样本都在平面的一边，而第二类的所有样本都在另一边。如果存在这样的平面，感知器算法的目标是学习任何一个这样的平面，给定数据点。学习完成后超平面确定。这将使未来分类新的样本点变得容易，因为样本在某一侧会使G(X) = WTX + wn+1 值为正，当点在另一侧的时候，值为负。

通过感知机学习的准则，权重可以拓展为 W^ = (w1 w2 … wn wn+1)T Rn+1，特征向量可写为 X^ = (x1 x2 … xn 1)T Rn+1。同时，分类平面能写为G(X) = W^ TX^ = 0。更新权重的学习规则可写为：



其中是可以适当调整的学习速率，以提高学习过程的收敛效率。

**三．实验目标**

(1)了解线性感知器学习算法的工作原理。

(2)了解各参数对算法学习速率和收敛性的影响。

(3)了解数据分布对算法学习能力的影响。

**四．实验流程**

1.4内容和程序

阶段1:

(1)根据上文第1.2节的原理和理论，设计并编写了感知器学习的程序代码，对两个类进行线性分类。

(2)创建一个线性可分模式数据集，每个类有50多个样本。

(3)初始化权向量W(0)，并为学习选择合适的值。速率 (0,1)

(4)用数据集运行程序并记录最终结果，注意收敛的次数。

阶段2:

重复上述步骤(2)~(4)在第1阶段的不同数据集，在两类模式之间有不同的间隔。记下你的观察。

阶段3:

研究不同学习速率对不同的可分性的影响。

阶段4:

在前一节中探究这些问题，并设计实验来回答这些问题。完成并提交一份关于所有实验结果的实验报告，并对实验结果进行对比分析和总结。

**五.实验数据**

**1.相同数据集，相同初始权向量w=[1,1],b=0，不同学习率对收敛次数的影响**

**1.1 实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 学习率 | 最终权重 | 最终截距 | 收敛次数 | 实验图片 |
| 0.1 | [-0.18252259,  1.62558615] | -35 | 235 | 1-1 |
| 0.2 | [0.399401,0.41894692] | -17.8 | 238 | 1-2 |
| 0.3 | [-0.03453636,2.46920406] | -55.8 | 252 | 1-3 |
| 0.4 | [0.7882074,250731699] | -68.8 | 230 | 1-4 |
| 0.5 | [1.62256688,0.9925345] | -81 | 216 | 1-5 |
| 0.6 | 1.28041612 | -102 | 228 | 1-6 |

(横坐标为学习率)

**1.2 实验讨论**

如图可看出学习率对收敛次数影响不大

**2.相同数据集，相同学习率，不同相同初始权向量w,b，不同学习率对收敛次数的影响**

**2.1实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 初始权重 | 初始截率 | 最终权重 | 最终截距 | 收敛次数 | 实验图片 |
| [ 52.68 -23.157] | -538.846 | [ 52.68048939 ,-23.15773146] | -538.846 | 11641 | 3-1 |
| [0, 0] | 0 | [ 52.36628847 ，-23.02866532] | -535.4 | 11543 | 3-2 |
| [100,100] | -100 | [ 52.72681869 ,-23.06641482] | -542.8 | 9754 | 3-3 |
| [1000,1000] | -200 | [ 50.87454382 ,-21.94649674] | -540.4 | 6954 | 3-4 |
| [10000,10000] | -200 | [ 130.7325812 , -57.39781666] | -1339.0 | 17225 | 3-5 |

(横坐标为不同初始权向量)

**2.2 结果讨论**

如图可看出不同初始权向量对收敛次数有影响，其具体关系不明显。

**3.不同数据集，相同初始权向量w=[0,0],b=0，相同学习率对收敛次数的影响**

**3.1实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | 最终权重 | 最终截距 | 收敛次数 | 实验图片 |
| 第一个样本为中心点（40，40），极差为10；第二个样本中心点（10，30），极差为10 | [ 18.75608895 -7.43341372] | -257.8 | 2927 | 2-1 |
| 第一个样本为中心点（40，40），极差为5；第二个样本中心点（10，30），极差为5 | [ 8.93962152 -5.18912634] | -0.2 | 5 | 2-2 |
| 第一个样本为中心点（40，40），极差为5；第二个样本中心点（30，30），极差为5 | [ 5.51995043 1.31546099] | -233.6 | 11036 | 2-3 |
| 第四个样本为中心点（40，40），极差为10；第二个样本中心点（30，30），极差为10 | \ | \ | 不收敛 | 2-4 |

**3.2结果讨论**

由实验结果可看出，不同数据样本对收敛次数影响很大。不同样本之间差异越大，即整体离得越远；收敛次数越小，收敛性能越好；若样本差异小，甚至不可分，则无法确定超平面将两样本分离

**4.总结**

本实验采用控制变量法；自变量分别为初始权重、学习率、数据样本对因变量感知器算法学习的收敛次数的影响，得到三个结论：

4.1.不同初始权向量对收敛次数有影响，其具体关系不明显

4.2.学习率对收敛次数影响不大

4.3.不同数据样本对收敛次数影响很大。不同样本之间差异越大，即整体离得越远；收敛次数越小，收敛性能越好；若样本差异小，甚至不可分，则无法确定超平面将两样本分离