

感知机中的几何解释

Haishou Ma

January 24, 2019

1 几何解释

线性方程 $w \cdot x + b = 0$ 对应于特征空间 R^n 中的一个超平面（在平面内就是一条直线，三维空间中就是一个平面），其中 w 是超平面的法向量。超平面将特征空间划分为两个部分，其中法向量指向的空间 $w \cdot x + b > 0$ ，反方向 $w \cdot x + b < 0$ 。超平面也被称为分离超平面。如图1.1所示

2 学习算法几何解释

感知机的学习算法主要是通过不断地纠正错误分类的点，从而达到学习的目的。直观的图如2.1所示，图中一开始的分离超平面为 $w^1 \cdot x + b = 0$ （黑线），法向量为 w^1 ，此时有一个节点 x_i 的负例被错误划分。因此需要更新参数 $w^2 = w^1 + \eta x_i y_i$ （注意：因为是负例，

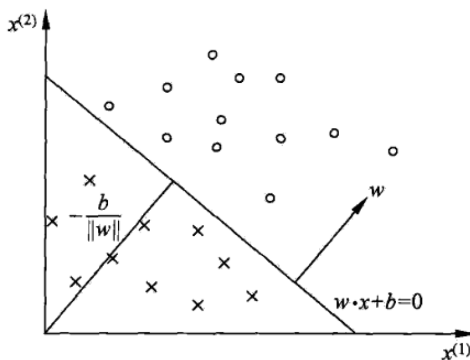


Figure 1.1: 超平面

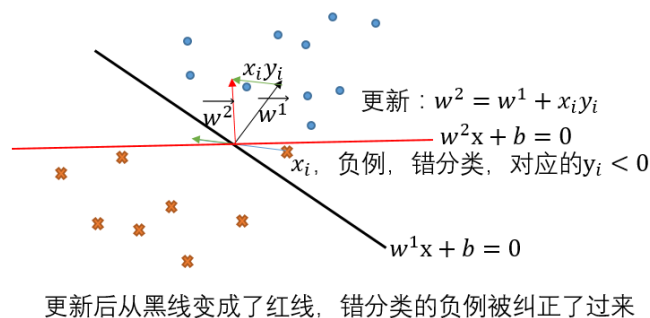


Figure 2.1: 纠正一个负例

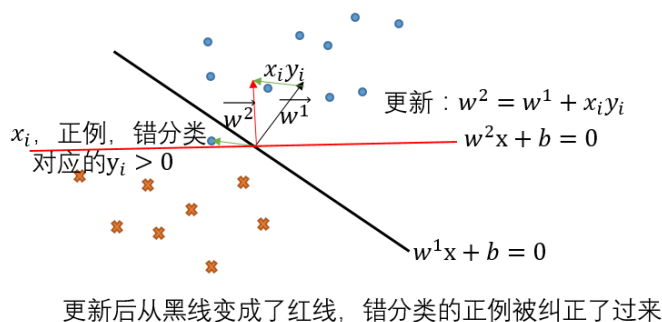


Figure 2.2: 纠正一个正例

所以 $y_i = -1$), 根据向量的加法, 可以得到更新后的超平面法向量为 w^2 , 相应的超平面为 $w^1 x + b = 0$ (红线), 因此原本错误分类的负例就被纠正过来了。

反之, 当错误分类的点为正例时, 也有类似的结果, 如图2.2所示。

总结: 当一个实例点被错误分类时, 即位于分离超平面的错误一侧时, 调整 w, b 的值, 使分离超平面向该误分类点的一侧移动, 以减少该误分类点与超平面之间的距离, 直到超平面越过该误分类点使其被正确分类。(李航: 统计学习方法)