

两类感知器

$$\hat{y} = \text{sgn}(w^T x + b)$$

算法 3.1 两类感知器的参数学习算法

```

输入: 训练集  $\mathcal{D} = \{(x^{(n)}, y^{(n)})\}_{n=1}^N$ , 最大迭代次数  $T$ 
1 初始化  $w_0 \leftarrow 0, k \leftarrow 0, t \leftarrow 0$ ;
2 repeat
3   对训练集  $\mathcal{D}$  中的样本随机排序;
4   for  $n = 1 \dots N$  do
5     选取一个样本  $(x^{(n)}, y^{(n)})$ ;
6     if  $w_k^T (y^{(n)} x^{(n)}) \leq 0$  then
7        $w_{k+1} \leftarrow w_k + y^{(n)} x^{(n)}$ ;
8        $k \leftarrow k + 1$ ;
9     end
10     $t \leftarrow t + 1$ ;
11    if  $t = T$  then break;
12  end
13 until  $t = T$ ;
输出:  $w_k$ 

```

$$\text{sgn}(z) = \begin{cases} 1 & z > 0 \\ -1 & z < 0 \\ 0 & z = 0 \end{cases}$$

$$\text{sgn}(z) = \begin{cases} +1 \\ -1 \end{cases}$$

0 或任意值

预测错误:

$$w_1 + \eta(y - \hat{y})x \rightarrow w_2$$

$$b_1 + \eta(y - \hat{y}) \rightarrow b_2$$

例

1. 数据集

| 特征1 (x1) | 特征2 (x2) | 类别 (y) |
|----------|----------|--------|
| 2 | 3 | +1 |
| 1 | 1 | +1 |
| 4 | 2 | -1 |
| 3 | 3 | -1 |

初始化 $W=[0,0]$ 偏置 $b=0$ 学习率 $\eta=0.1$

训练过程:

1 轮迭代:

① 样本 1.

计算预测值

$$\hat{y} = \text{sgn}[\underbrace{0 \times 2 + 0 \times 3}_{w^T x} + \underbrace{0}_b] = \text{sgn}(0) = 0$$

0 ≠ 类别 $y_1 = +1$.

⇒ 预测错误.

更新权重和偏置.

$$W_2 = W_1 + \underbrace{0.1 \times (y - \hat{y})}_{\eta \times (y - \hat{y})} \cdot [2, 3] = [0.2, 0.3]$$

$$b_2 = \underbrace{0}_{b_1} + \underbrace{0.1 \times (y - \hat{y})}_{\eta (y - \hat{y})} = 0.1$$

② 样本 2 (1, 1, +1):

$$\hat{y}_2 = \text{sgn}(0.2 \times 1 + 0.3 \times 1 + 0.1) = \text{sgn}(0.6) = +1$$

$y_2 = y_2 = +1$ 预测正确, 不更新.

③ 样本 3. (4, 2, -1)

$$\hat{y}_3 = \text{sgn}(0.2 \times 4 + 0.3 \times 2 + 0.1) = \text{sgn}(1.5) = +1$$

$+1 = \hat{y}_3 \neq y_3 = -1 \quad \therefore \text{预测错误.}$

更新

$$W_2 + \eta \times (y_3 - \hat{y}_3) \times X_3$$

$$W_3 = [0.2, 0.3] + 0.1 \times (-1 - 1) \times [4, 2] = [-0.2, 0.1]$$

$$b_3 = 0.1 + \eta \times (\hat{y}_3 - y_3)$$
$$b_3 = 0.1 + 0.1 \times (-1 - 1) = -0.1$$

④ 样本 4 $(3, -3, -1)$

$$\hat{y}_4 = \text{sgn}[-0.2 \times 3 + 0.1 \times (-3) - 0.1] = \text{sgn}(-1) = -1$$

$\hat{y}_4 = y_4 = -1$ \therefore 预测正确不更新.

继续进行第二次迭代直到所有样本均正确或达到最大迭代次数. w^* 及 b^* :

如 $w^* = [0.5, -0.5]$ $b^* = -1$

则 $0.5x_1 - 0.5x_2 - 1 = 0$.

\Rightarrow 得出 x_1 与 x_2 关系式 $x_1 - x_2 = 2$:

可见这是一个决策边界.

当 $x_1 - x_2 > 2$ 时为类别 A
当 $x_1 - x_2 < 2$ 时为类别 B.