

Chap02- 感知机

I. 模型

$$\mathbb{R}^n \times \rightarrow Y, Y = \{-1, +1\}$$

$$f(x) = \text{sign}(w \cdot x + b)$$

$w \in \mathbb{R}^n$: weight ; $b \in \mathbb{R}$: bias ; $F = \{f \mid f(x) = w \cdot x + b\}$

$$\text{sign}(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

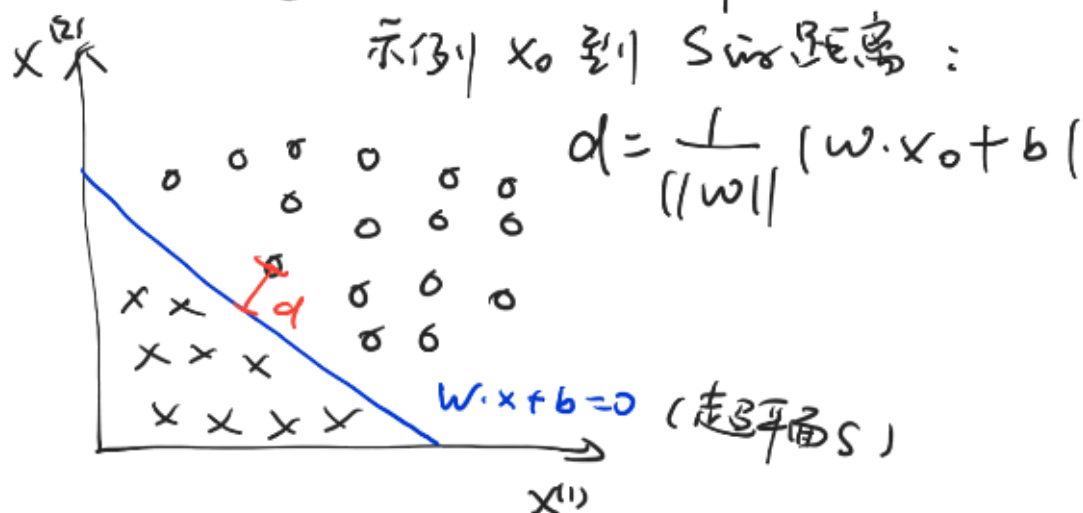
II. 策略

1. Loss function

① 0-1 Loss function : 不连续, 不可导
不易极小化损失 (优化)

② L_2 -Norm Loss function :

示例 x_0 到 S 的距离 :



对于误分类的点 (x_i, y_i) 来说, $-y_i(w \cdot x_i + b) > 0$

$$d = -\frac{1}{\|w\|} y_i (w \cdot x_i + b)$$

假设误分类点集合为 M , 总距离为

$$D = -\frac{1}{\|w\|} \sum_{x_i \in M} y_i (w \cdot x_i + b)$$

($\|w\|$ 为 const, 则得到) 模型 loss function,

$$L(w, b) = - \sum_{x_i \in M} y_i (w \cdot x_i + b)$$

在假设空间 F 中选取 $L(w, b)$ 最小的模型参数 w, b

IV. 算法 = 梯度下降法

$$[\text{梯度} \Rightarrow \begin{cases} \nabla_w L(w, b) = - \sum_{x_i \in M} y_i x_i \\ \nabla_b L(w, b) = - \sum_{x_i \in M} y_i \end{cases}$$

1. 算法 =

① 任选一个超平面 (w_0, b_0)

② 在训练集中选取数据 (x_i, y_i)

③ if $y_i (w \cdot x_i + b) \leq 0$

$$\begin{cases} w \leftarrow w + \eta y_i x_i \\ b \leftarrow b + \eta y_i \end{cases}$$

η ($0 < \eta \leq 1$), 为 learning rate

④ 转至②, 重复迭代直至训练集中没有误分类点

Input = $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$,

$x_i \in \mathbb{R}^n$, $y_i \in \{-1, +1\}$, η ($0 < \eta \leq 1$)

Output = w, b ; $f(x) = \text{sign}(w \cdot x + b)$

注 = 采用不同的初值或选取不同的分类点 (x_i, y_i) (步骤②中), 最终结果 w, b 可能不同.

如果训练集线性可分, 则算法收敛

2. 对偶形式

2.1 原始形式中初值 $w_0 = 0, b_0 = 0$

$$\Rightarrow \begin{cases} w = \sum_{i=1}^n \alpha_i y_i x_i \\ b = \sum_{i=1}^n \alpha_i y_i \end{cases} \quad \alpha_i = n_i \eta$$

算法: Input: $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$,
 $x_i \in \mathbb{R}^n, y_i \in \{-1, +1\}; \eta (0 < \eta \leq 1)$

Output: α, b ;

$$f(x) = \text{Sign} \left(\sum_{i=1}^n \alpha_i y_i x_i \cdot x + b \right)$$

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$$

① $\alpha \leftarrow 0, b \leftarrow 0$

② 在训练集中选取 (x_i, y_i)

③ if $y_i \left(\sum_{j=1}^n \alpha_j y_j x_j \cdot x + b \right) \leq 0$

$$\alpha_i \leftarrow \alpha_i + \eta$$

$$b \leftarrow b + \eta y_i$$

④ 迭代 ②, 直至没有误分数据

上次修改：2018年10月3日