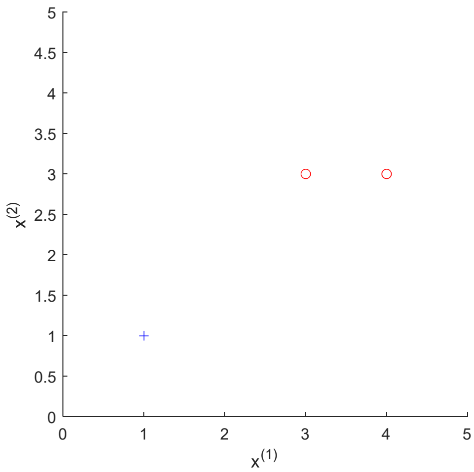
感知机例题

如下图所示的训练数据集，其正实例点是 $x_1=(3,3)^T$， $x_2=(4,3)^T$，负实例点是 $x_3=(1,1)^T$，试用基本的感知机学习算法求感知机模型 $f(x)=\mathrm{sign}(w\cdot x+b)$。这里，$w=\left(w^{(1)},w^{(2)}\right)^T$， $x=\left(x^{(1)},x^{(2)}\right)^T$，初始值可取 $w_0=0$， $b_0=0$，损失函数采用 $\min\limits_{w,b}L(w,b)=-\sum\limits_{x_k\in M}y_k\left(w^T\cdot x_k+b\right)$，参数更新公式采用 $w_{t+1}=w_t+\eta y_k x_k$，$b_{t+1}=b_t+\eta y_k$，$\eta=1$，$k=1,2,3$， $y_1=y_2=+1$， $y_3=-1$。



解：构建最优化问题：

$$\min_{w,b}L(w,b)=-\sum_{x_k\in M}y_k\left(w\cdot x_k+b\right)$$

下面求解向量参数 w 和标量 b。

（1）取初值 $w_0=0$， $b_0=0$

（2）下面，对于不能正确分类的点，就更新参数 w 和 b，若可以正确分类该点，就不更新。

对 $x_1=(3,3)^T$， $y_1\left(w_0\cdot x_1+b_0\right)=0$，未能被正确分类，更新 w 和 b，

$$w_1=w_0+y_1x_1=(3,3)^T，\quad b_1=b_0+y_1=1$$

得到线性模型

$$w_1\cdot x+b_1=3x^{(1)}+3x^{(2)}+1$$

（3）对 x_1， x_2，显然 $y_k\left(w_1 \cdot x_k + b_1\right) > 0$，被正确分类，不修改 w 和 b；

对 $x_3 = (1,1)^T$， $y_3\left(w_1 \cdot x_3 + b_1\right) < 0$，被误分类，更新 w 和 b。

$$w_2 = w_1 + y_3 x_3 = (2,2)^{T+}，\quad b_2 = b_1 + y_3 = 0$$

得到线性模型

$$w_2 \cdot x + b_2 = 2x^{(1)} + 2x^{(2)}$$

如此继续下去，直到

$$w_7 = (1,1)^T，\quad b_7 = -3$$

$$w_7 \cdot x + b_7 = x^{(1)} + x^{(2)} - 3$$

对所有数据点 $y_k\left(w_7 \cdot x_k + b_7\right) > 0$，没有误分类点，损失函数达到极小。

分离超平面： $x^{(1)} + x^{(2)} - 3 = 0$

感知机模型为： $f(x) = \mathrm{sign}\left(x^{(1)} + x^{(2)} - 3\right)$

迭代过程见下表

迭代次数	误分类点	w	b	$w \cdot x + b$
0		0	0	0
1	x_1	$(3,3)^T$	1	$3x^{(1)} + 3x^{(2)} + 1$
2	x_3	$(2,2)^T$	0	$2x^{(1)} + 2x^{(2)}$
3	x_3	$(1,1)^T$	-1	$x^{(1)} + x^{(2)} - 1$
4	x_3	$(0,0)^T$	-2	-2
5	x_1	$(3,3)^T$	-1	$3x^{(1)} + 3x^{(2)} - 1$
6	x_3	$(2,2)^T$	-2	$2x^{(1)} + 2x^{(2)} - 2$
7	x_3	$(1,1)^T$	-3	$x^{(1)} + x^{(2)} - 3$
8	0	$(1,1)^T$	-3	$x^{(1)} + x^{(2)} - 3$

这是在计算中误分类点先后取 $x_1, x_3, x_3, x_3, x_1, x_3, x_3$ 得到的分离超平面和感知机模型。

如果在计算中误分类点依次取$x_1, x_3, x_3, x_3, x_2, x_3, x_3, x_3, x_1, x_3, x_3$，那么得到的分离超平面是$2x^{(1)} + x^{(2)}$-5=0。

感知机学习算法采用不同的初始值或选取不同的误分类点，解可以不同。