厦门大学信息学院 本科选修课
2021-2022 第二学期

模 式 识 别
Pattern Recognition

主 讲：王 程

第三章　判别函数

3.4 广义线性判别函数

3.5 感知器算法

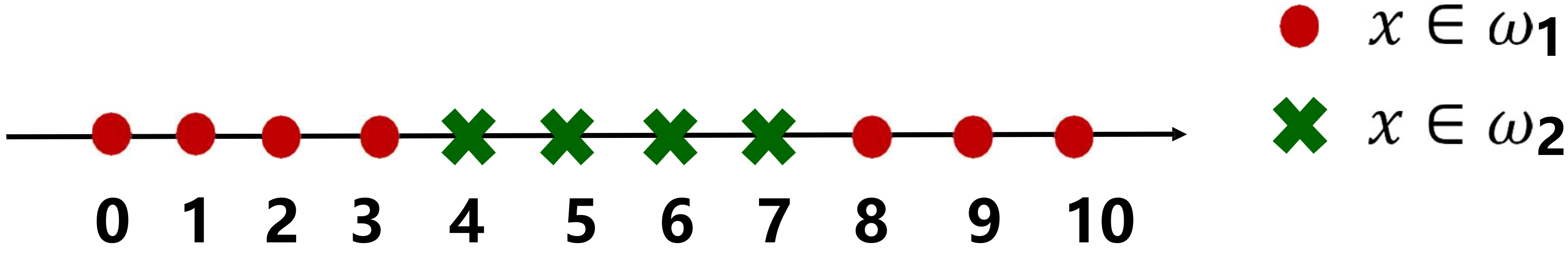
3.6 可训练的确定性分类器的迭代算法

3.7 Fisher线性判别

3.4 广义线性判别函数

问题：

已知10个人的血糖数据如下图：



ω_1 - 不正常血糖人群（高/低血糖）
ω_2 - 正常血糖人群

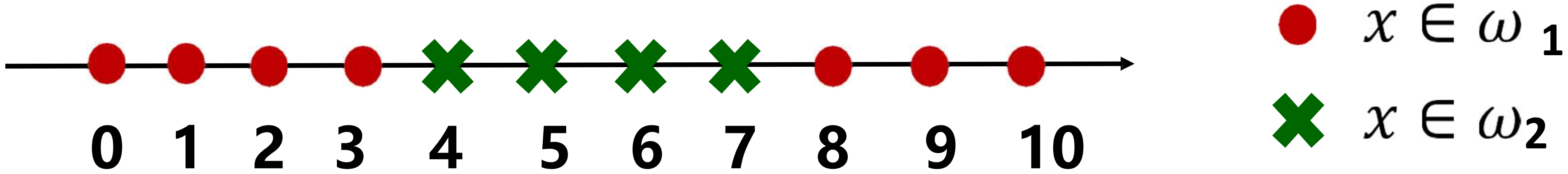
设计一个一维线性判别函数，能将两类人群区分开。

3.4 广义线性判别函数

一维线性判别函数对应的判别界面为:

$$g(x) = wx + w_0 = 0$$

此时判别面为一个点，显然没有任何点（判别面）能 解决这个划分问题。



需要的判别函数满足:

$$\text{如果}\begin{cases} x < b \text{ 或 } x > a & \text{则决策 } x \in \omega_1 \\ b < x < a & \text{则决策 } x \in \omega_2 \end{cases}$$

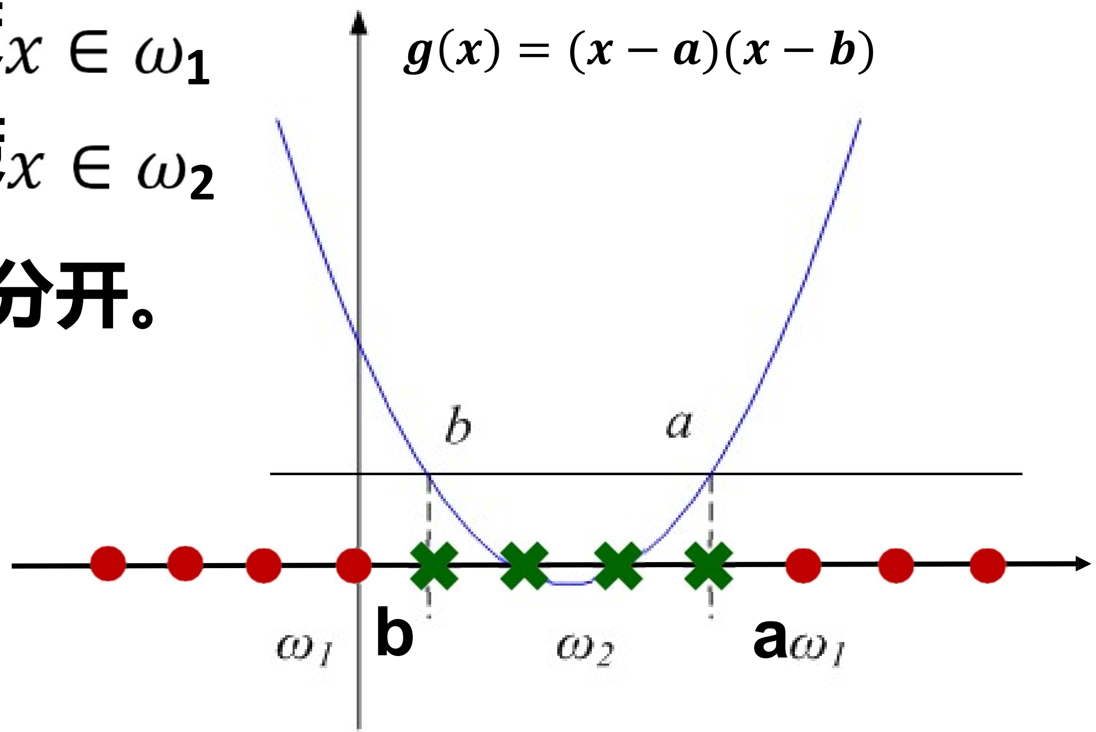
3.4 广义线性判别函数

如果建立一个二次判别函数:

$$g(x) = (x-a)(x-b)$$

如果 $\begin{cases} g(x) > 0 & \textbf{则决策} x \in \omega_1 \\ g(x) < 0 & \textbf{则决策} x \in \omega_2 \end{cases}$

那么可以将两个类别区分开。



3.4 广义线性判别函数

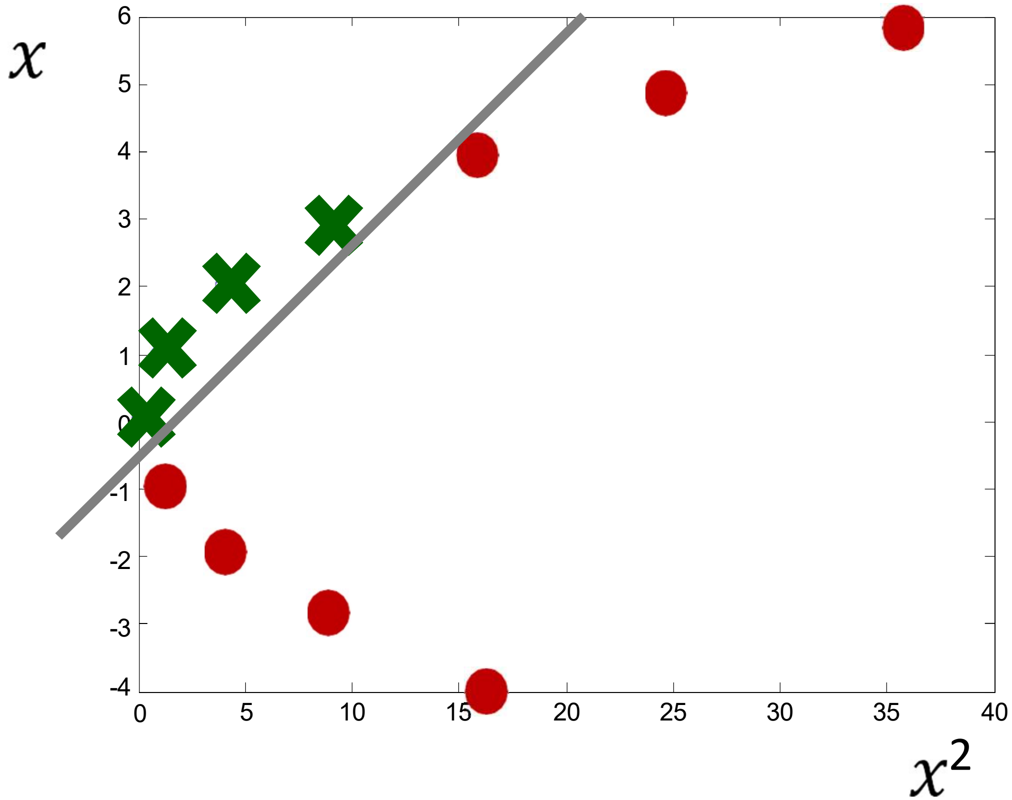
$$g(x) = (x-a)(x-b) = x^2 - (a+b)x + ab$$
$$= (1, -a-b)\begin{pmatrix} x^2 \\ x \end{pmatrix} + ab$$
$$= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{f}(x) + w_0$$

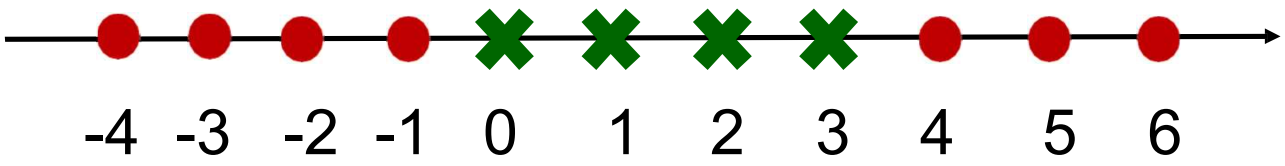
广义线性判别函数形式

$$\boldsymbol{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} x^2 \\ x \end{pmatrix}; \quad \boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{bmatrix} 1 \\ -(a+b) \end{bmatrix};$$

$$w_0 = ab$$

3.4 广义线性判别函数





3.4 广义线性判别函数

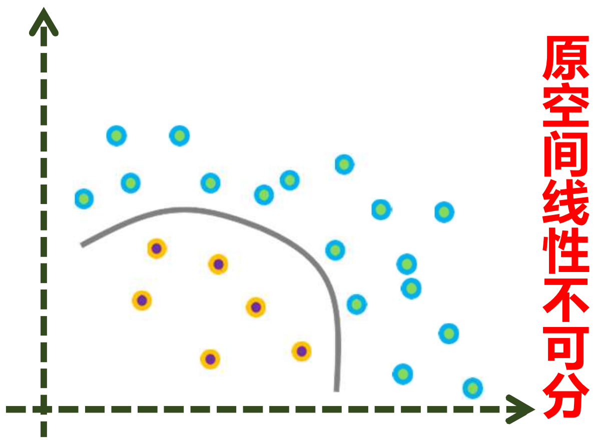
广义线性判别函数

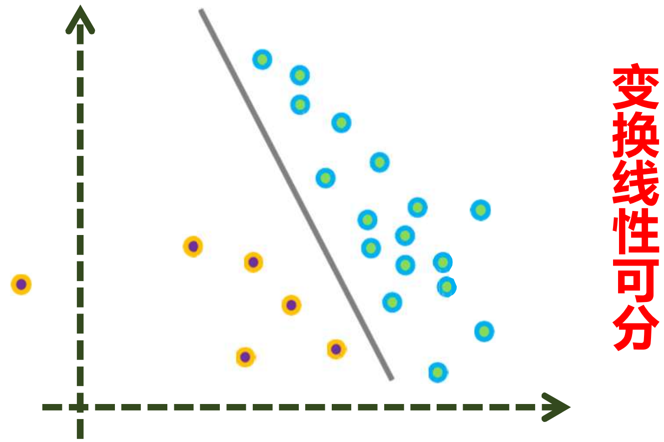
将输入模式进行**非线性变换** $x \in \mathbb{R}^D \to f(x) \in \mathbb{R}^K$，使变换后的模式空间线性可分，在变 换后空间定义的线性判别函数可以表示为：

$$g(x) = w^T f(x) + w_0$$

其中：$f(x) = [f_1(x), f_2(x), \cdots, f_K(x)]^T \in \mathbb{R}^K$ 映射函数将D维的模式映射到H维的变换空间，

w= $(w_1, w_2, \ldots, w_K)^T$ 为广义权向量。





3.4 广义线性判别函数

设原模式空间中的模式为$x \in \mathbb{R}^D$**，变换后空间中的模式为**$f(x)$**,**

则$f(x) = [f_1(x), f_2(x), \cdots, f_K(x)]^T \in \mathbb{R}^H$

下面讨论广义线性判别函数的意义。

3.4 广义线性判别函数

（1）线性判别函数

如$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}$，则线性化后的判别式为：

$$\mathrm{g}(\boldsymbol{x}) = \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x} + w_{\mathbf{0}}$$

其中： $f_{\mathrm{i}}(\boldsymbol{x}) = x_{\mathrm{i}}$

3.4 广义线性判别函数

(2) $f_i(x)$选用二次多项式函数

$$\boldsymbol{f_i(x) = x_{p1}{}^{s}\ x_{p2}{}^{t}}$$

其中： $s,t \in \{0,1\}, p_1, p_2 = 1,2,\dots,D$

$$g(x) = \sum_{j=1}^{D} w_{jj} x_j^2 + \sum_{D}^{D-1} \sum_{l=j+1}^{D} w_{jl} x_j x_l + \sum_{j=1}^{D} w_j x_j + w_0$$

3.4 广义线性判别函数

若$\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2$，则：

$$f(x) = (x_1^2, x_1 x_2, x_2^2, x_1, x_2)^T$$

$$\mathrm{w} = (w_{11}, w_{12}, w_{22}, w_1, w_2)^T$$

$$\begin{aligned} g(x) &= w_{11}x_1^2 + w_{12}x_1x_2 + w_{22}x_2^2 + w_1x_1 + w_2x_2 + w_0 \\ &= \boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{f}(x) + w_0 \end{aligned}$$

3.4 广义线性判别函数

（3）$f_i(x)$选用r次多项式函数

这时的判别函数$g(x)$可用以下递推关系写出：

常数 $g_0(x) = w_0$

一次 $g_1(x) = \sum_{p_1=1}^{D} w_{p_1} x_{p_1} + g_0(x)$

二次 $g_2(x) = \sum_{p_1=1}^{D} \sum_{p_2=p_1}^{D} w_{p_1 p_2} x_{p_1} x_{p_2} + g_1(x)$

r次 $g_r(x) = \sum_{p_1=1}^{D} \sum_{p_2=p_1}^{D} \cdots \sum_{p_r=p_{r-1}}^{D} w_{p_1 p_2 \cdots p_r} x_{p_1} x_{p_2} \cdots x_{p_r} + g_{r-1}(x)$

3.4 广义线性判别函数

缺点：

$g(x)$的项数随着r和D的增加而迅速增大，原来模式 x的维数即使不高，但如果采用次数r较高的多项式 来变换，也会使得变换后的模式$f(x)$的模式很高维，给分类带来较大的困难。

第三章　判别函数

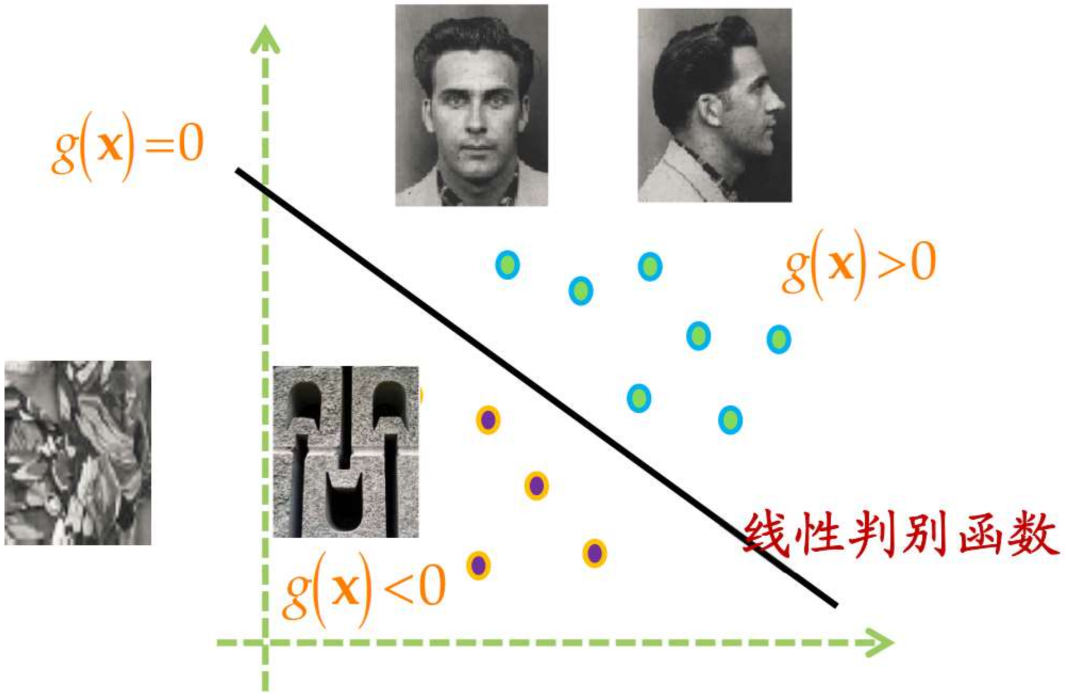
3.4 广义线性判别函数

3.5 感知器算法

3.6 可训练的确定性分类器的迭代算法

3.7 Fisher线性判别

3.5 感知器算法



$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0$$

参数学习阶段（训练）：

$$g(\boldsymbol{x}) = w_1x_1 + w_2x_2 + \cdots + w_0$$

判别检测阶段（测试）：

$$g(\boldsymbol{x}) = \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x} + w_0 \begin{cases} > 0 & \boldsymbol{x} \in \omega_1 \\ < 0 & \boldsymbol{x} \notin \omega_2 \end{cases}$$

3.5 感知器算法

设计线性分类器的主要步骤：

- **采集一系列具有标定的样本集$\{x_1, x_2 \dots, x_N\}$，样本类别标定已知；**
- **确定准则函数进行判别界面参数的学习，包括$\boldsymbol{w}$和 w_0;**
- **用优化方法求出准则函数对应的最优分界面参数$\boldsymbol{w}$和w_0,**

  由此得到判别函数：

$$g(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}) + w_0$$

3.5 感知器算法

常用的准则函数：

1. **感知准则**
2. **Fisher准则**
3. **最小错分准则**
4. **最小平方误差准则**

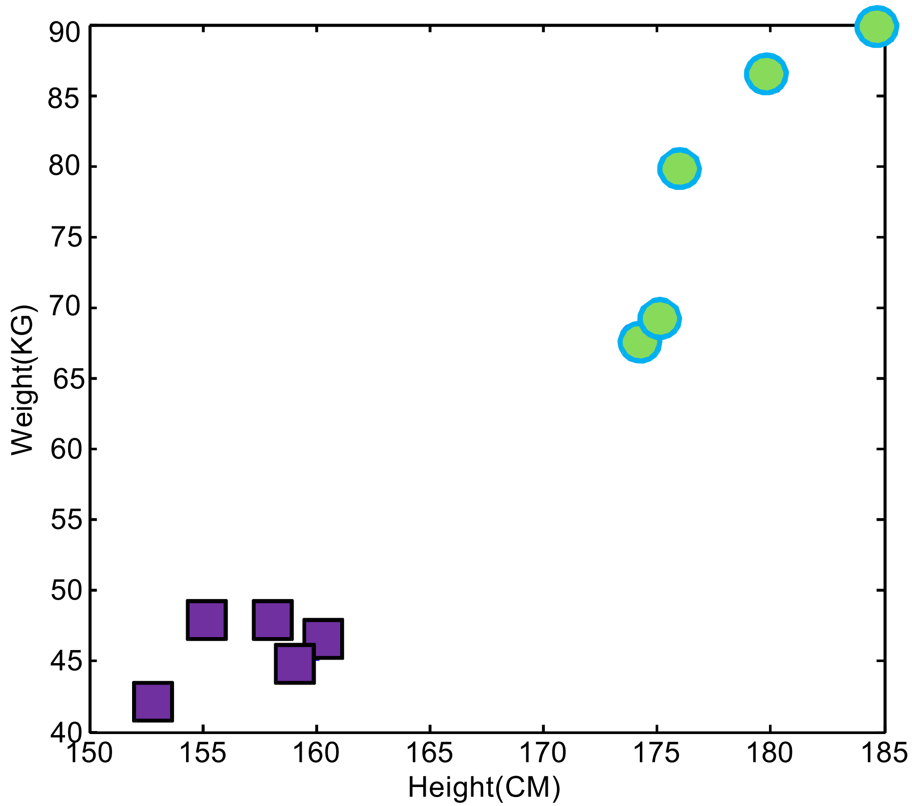
3.5 感知器算法

“感知器”（Perceptron）一词出自于20世纪50年代中期到60年代 中期人们对一种分类学习机模型的称呼，它是属于有关动物和机器学习的仿生学领域中的问题。

当时的一些研究者认为感知器是一种学习机的强有力模型，后来发现估计过高了，但发展感知器的一些相关概念仍然沿用下来。

3.5 感知器算法

已知一个班级所有同学的身高和体重数据，学号为{0,2,3,4,7}的同学为男生，学号为{1,5,6,8,9}的同学为女生，如何分开两类数据



学号	0	1	2	3	4	5	6	7	8	9
身高（cm）	174	155	175	180	185	153	158	176	160	159
体重（kg）	68	48	69	86	90	42	48	80	46	45
性别	男	女	男	男	男	女	女	男	女	女

3.5 感知器算法

如图虚线为初始判别函数为0时对应的超平面，判别函数由右图可知：

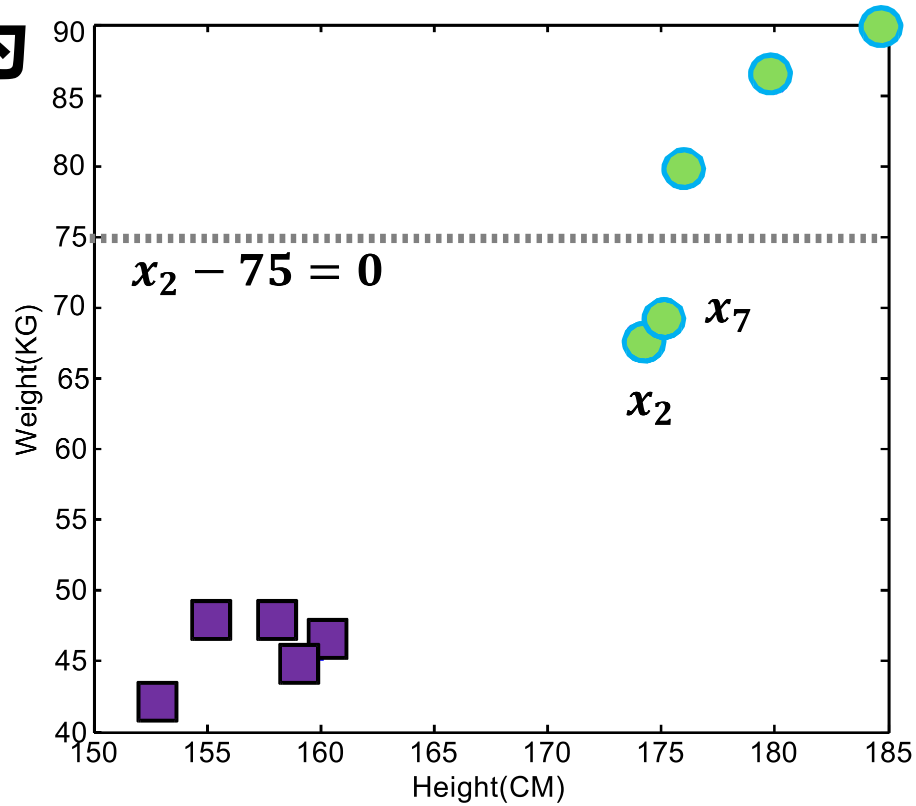
$$g^{(0)}(x) = x_2 - 75$$

该判别函数对应权向量

$$\boldsymbol{w} = (0,1)^T$$

和偏差

$$w_0 = -75$$



因此，根据增广权向量的定义可知该判别函数对应的增广权向量为：

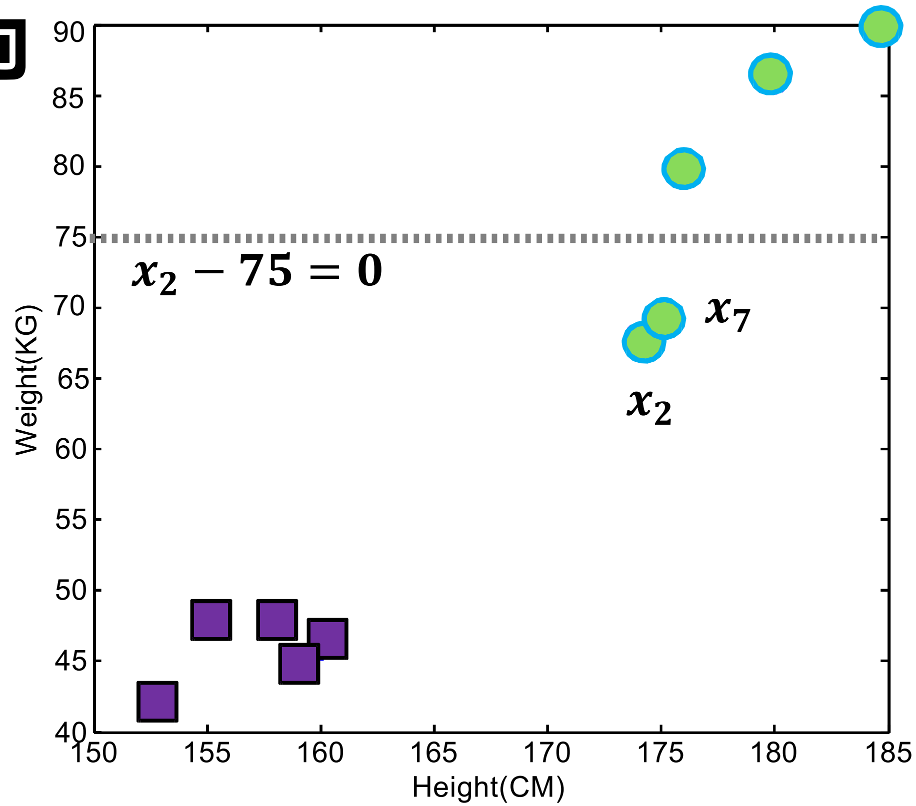
$$\boldsymbol{w} = (0,1,-75)^T$$

3.5 感知器算法

记该权向量为初始增广权向量

$$w^{(0)} = (0,1,-75)^T$$

错分样本为x_2和x_7。

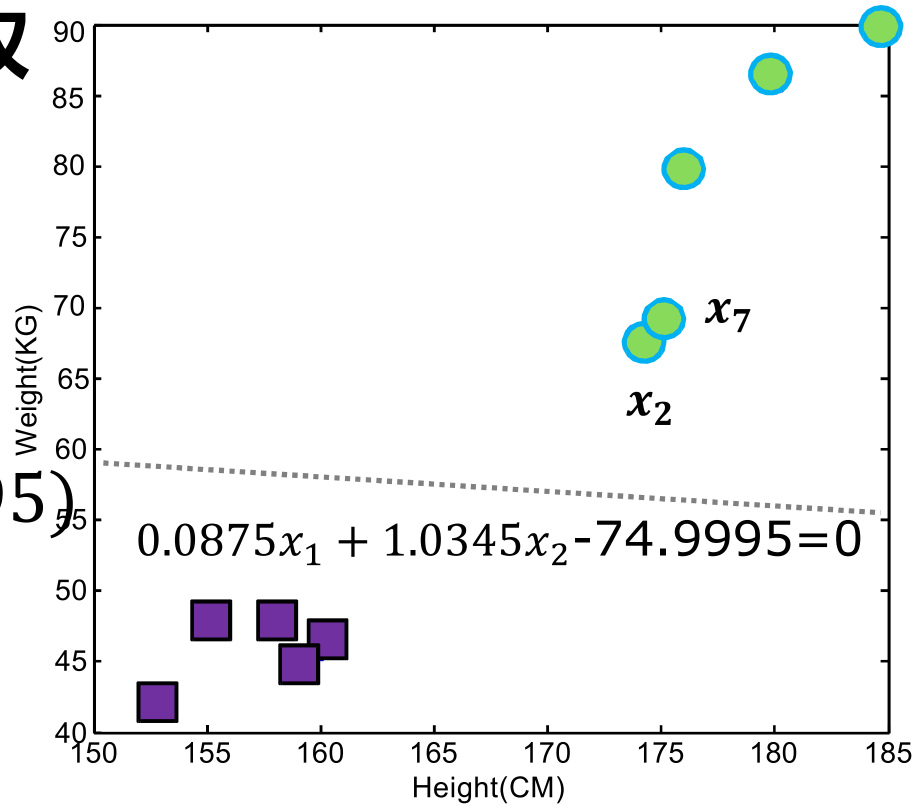


学号	0	1	2	3	4	5	6	7	8	9
身高（cm）	174	155	175	180	185	153	158	176	160	159
体重（kg）	68	48	69	86	90	42	48	80	46	45
性别	男	女	男	男	男	女	女	男	女	女

3.5 感知器算法

我们利用错分样本来调整权向量的取值：

$$w^{(1)} = w^{(0)} + Cx_2$$
$$= (0,1,-75)^T$$
$$+0.0005(175,69,1)^T$$
$$= (0.0875,1.0345,-74.9995)$$



判别函数

$$g^{(1)}(x) = 0.0875x_1 + 1.0345x_2 - 74.9995$$

该判别函数可以正确区分两类样本，停止迭代。

$$g(x) = 0.0875x_1 + 1.0345x_2 - 74.9995$$

3.5 感知器算法

感知器准则：

感知器算法是一种赏罚过程，对正确分类的模式则“赏”，这里用“不罚”，即判别函数参数不变，对错误分类的模式则“罚”。

具体做法是用全部样本训练一轮以后，只要有一个样本是错误的则需要进行第二轮迭代，即用全部模式再训练一次，这样不断反复，直到全部样本训练都能获得正确分类结果时，迭代完成。

3.5 感知器算法

感知器的训练算法如下：

1. **已知训练样本集**$\{x_1, x_2, \dots, x_N\}$以及**它们所属类别**ω_1和ω_2，**同时已知权向量初始值为**$w^{(0)}$
2. **第k次训练中，若**$x_n \in \omega_1$**,但**$w^{(k)^T} x_n \le 0$**,则分类器对第n个模式**x_n**做错误分类，应校正权向量：**

$$w^{(k+1)} = w^{(k)} + C x_n$$

C为一个矫正量

若$x_n \in \omega_2$**，且**$w^{(k)^T} x_n \ge 0$**，同样错误分类，应校正权向量：**

$$w^{(k+1)} = w^{(k)} - C x_n$$

如果x_n**不符合上述情况，则分类正确：**

$$w^{(k+1)} = w^{(k)}$$

3. **反复迭代至所有样本正确分类。**

3.5 感知器算法

步长C：

当步长C较小时，收敛速度很慢。

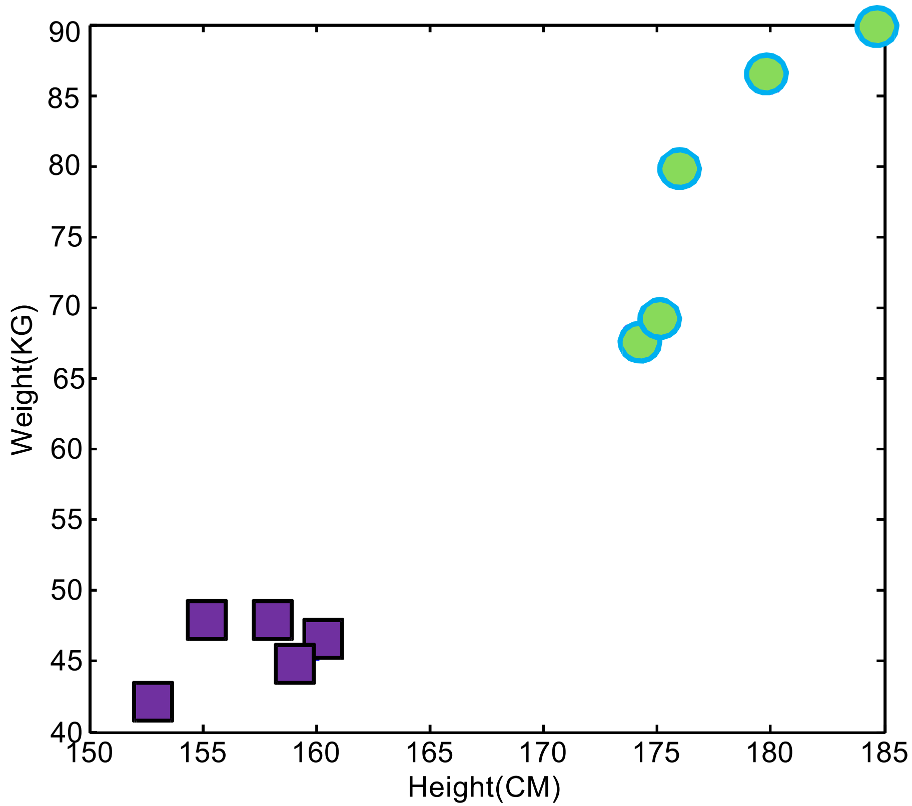
当步长C较大时，收敛速度快，但容易出现过极值震荡。

为了解决步长设定的问题，减小迭代次数，可以使用可变步长，比如绝对修正法就是对错分样本 x_n 采用下面的步长来调整权向量：

$$C^{(k+1)} = \frac{\left| {w^{(k)}}^T x_n \right|}{\left\| x_n \right\|^2}$$

3.5 感知器算法

已知一个班级所有同学的身高和体重数据，学号为{0,2,3,4,7}的同学为男生，学号为{1,5,6,8,9}的同学为女生，如何分开两类数据。



学号	0	1	2	3	4	5	6	7	8	9
身高（cm）	174	155	175	180	185	153	158	176	160	159
体重（kg）	68	48	69	86	90	42	48	80	46	45
性别	男	女	男	男	男	女	女	男	女	女

3.5 感知器算法

如图虚线为初始判别函数为0时对应的超平面，判别函数由右图可知：

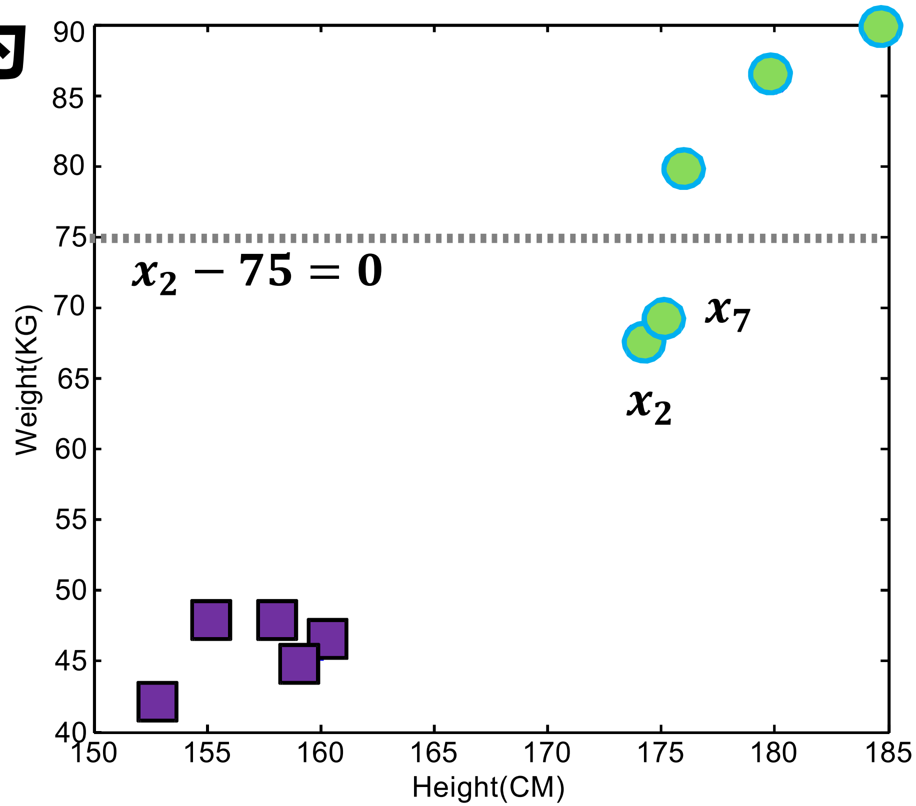
$$g^{(0)}(x) = x_2 - 75$$

该判别函数对应权向量

$$w = (0,1)^T$$

和偏差

$$w_0 = -75$$



因此，根据增广权向量的定义可知该判别函数对应的增广权向量为：

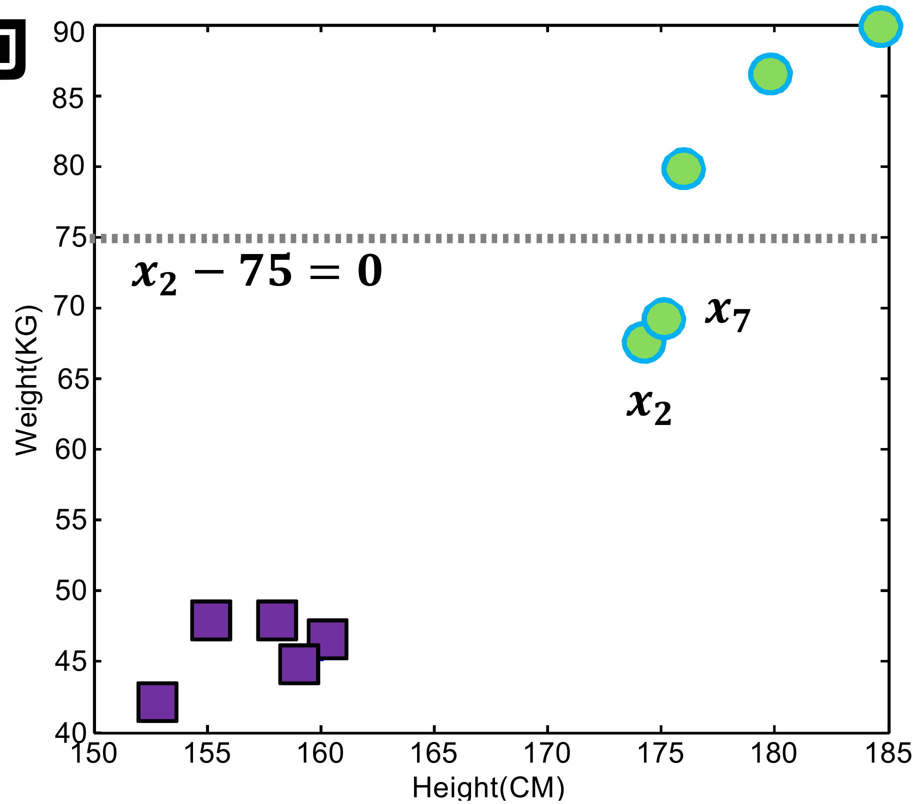
$$w = (0,1,-75)^T$$

3.5 感知器算法

记该权向量为初始增广权向量

$$w^{(0)} = (0,1,-75)^T$$

错分样本为x_2和x_7。

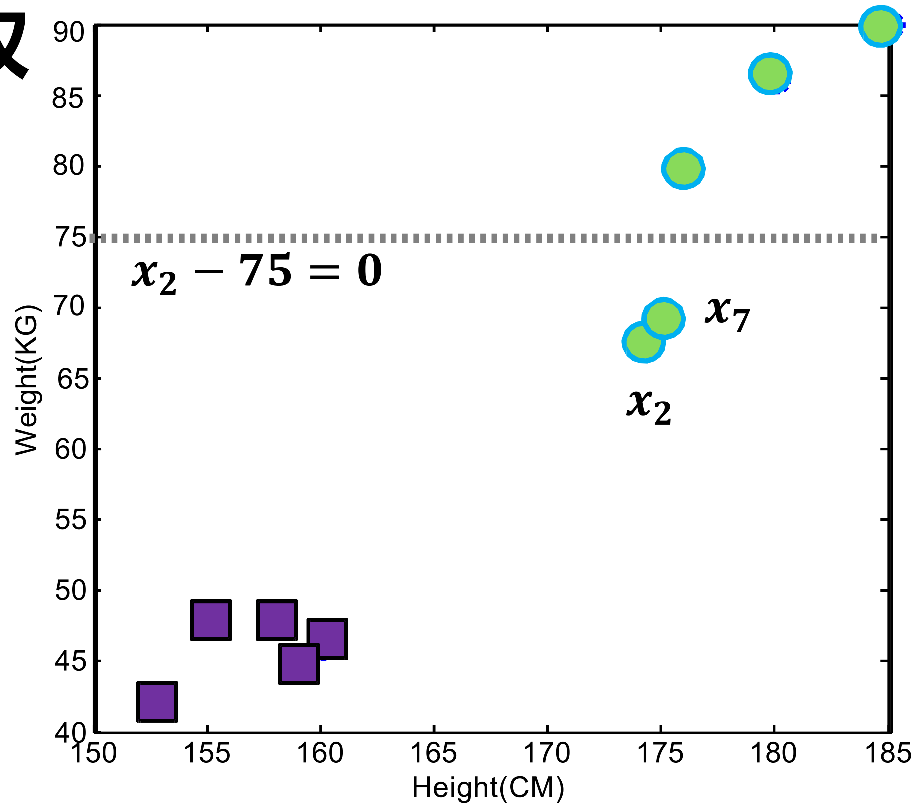


学号	0	1	2	3	4	5	6	7	8	9
身高（cm）	174	155	175	180	185	153	158	176	160	159
体重（kg）	68	48	69	86	90	42	48	80	46	45
性别	男	女	男	男	男	女	女	男	女	女

3.5 感知器算法

我们利用错分样本来调整权向量的取值：

$$w^{(1)} = w^{(0)} + Cx_2$$
$$= (0,1,-75)^T$$
$$+0.0001(175,69,1)^T$$
$$= (0.000175,1.000069,$$
$$-74.999999)^T$$



判别函数：

$$g^{(1)}(x) = 0.000175x_1 + 1.000069x_2 - 74.999999$$

该判别函数无法正确区分两类样本，继续迭代

3.5 感知器算法

步长C：

当步长C较小时，收敛速度很慢。

当步长C较大时，收敛速度快，但容易出现过极值震荡。

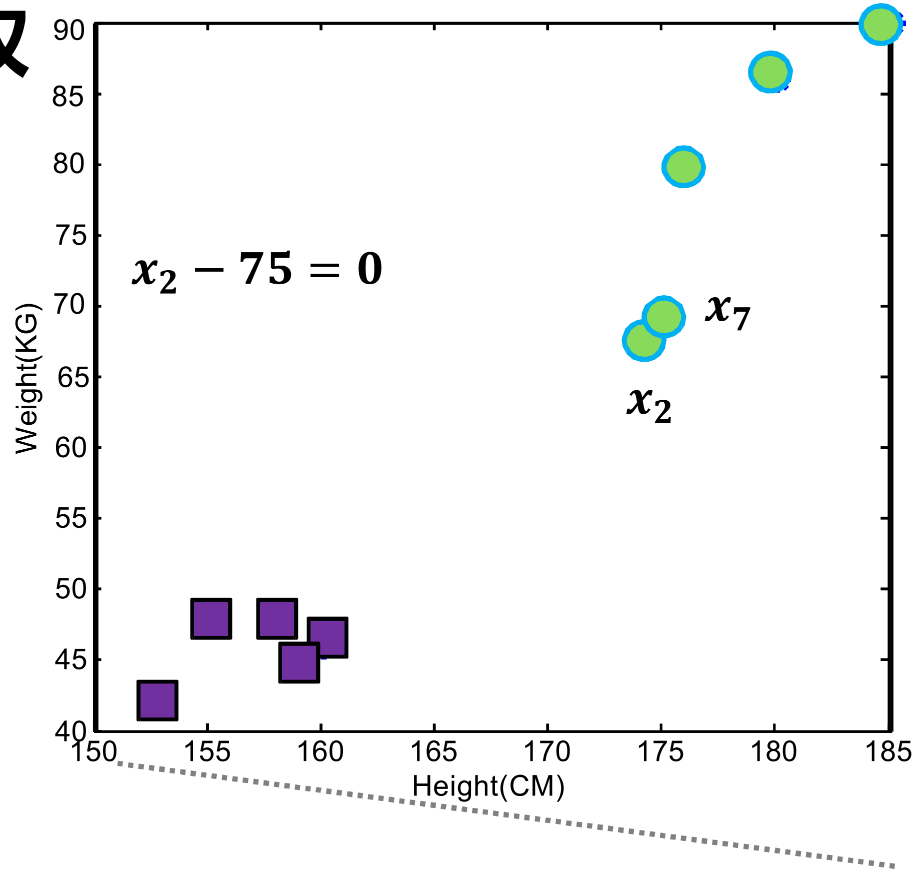
为了解决步长设定的问题，减小迭代次数，可以使用可变步长，比如绝对修正法就是对错分样本 x_n 采用下面的步长来调整权向量：

$$C^{(k+1)} = \frac{\left| {w^{(k)}}^T x_n \right|}{\|x_n\|^2}$$

3.5 感知器算法

我们利用错分样本来调整权向量的取值：

$$\boldsymbol{w}^{(1)} = \boldsymbol{w}^{(0)} + \boldsymbol{C}\boldsymbol{x}_2$$
$$= (0,1,-75)^T$$
$$+0.5(175,69,1)^T$$
$$= (87.5,33.5,-74.5)^T$$



判别函数：

$$g^{(1)}(x) = 87.5x_1 + 35.5x_2 - 74.5$$

该判别函数无法正确区分两类样本，继续迭代

3.5 感知器算法

步长C：

当步长C较小时，收敛速度很慢。

当步长C较大时，收敛速度快，但容易出现过极值震荡。

为了解决步长设定的问题，减小迭代次数，可以使用可变步长，比如绝对修正法就是对错分样本x_n采用下面的步长来调整权向量：

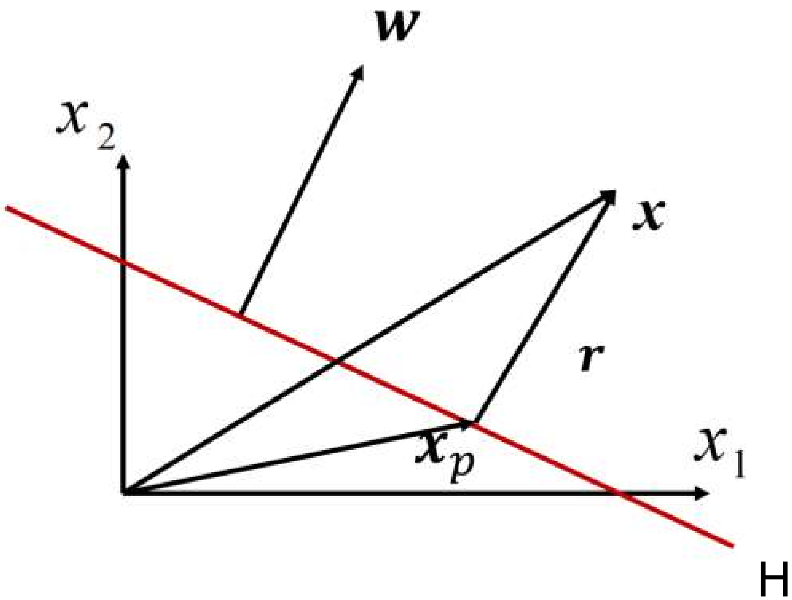
$$C^{(k+1)} = \frac{\left| {w^{(k)}}^T x_n \right|}{\|x_n\|^2}$$

超平面的几何性质

性质2：

$$\|\boldsymbol{r}\| = \frac{|g(\boldsymbol{x})|}{\|\boldsymbol{w}\|}$$

矢量x到H的距离$\|r\|$与$|g(x)|$值成正比



3.5 感知器算法

感知器算法的收敛证明：

感知器算法的收敛性，是指只要模式类别线性可分，就可以在有限的迭代次数之后，求出权向量的最优解。

证明方法：

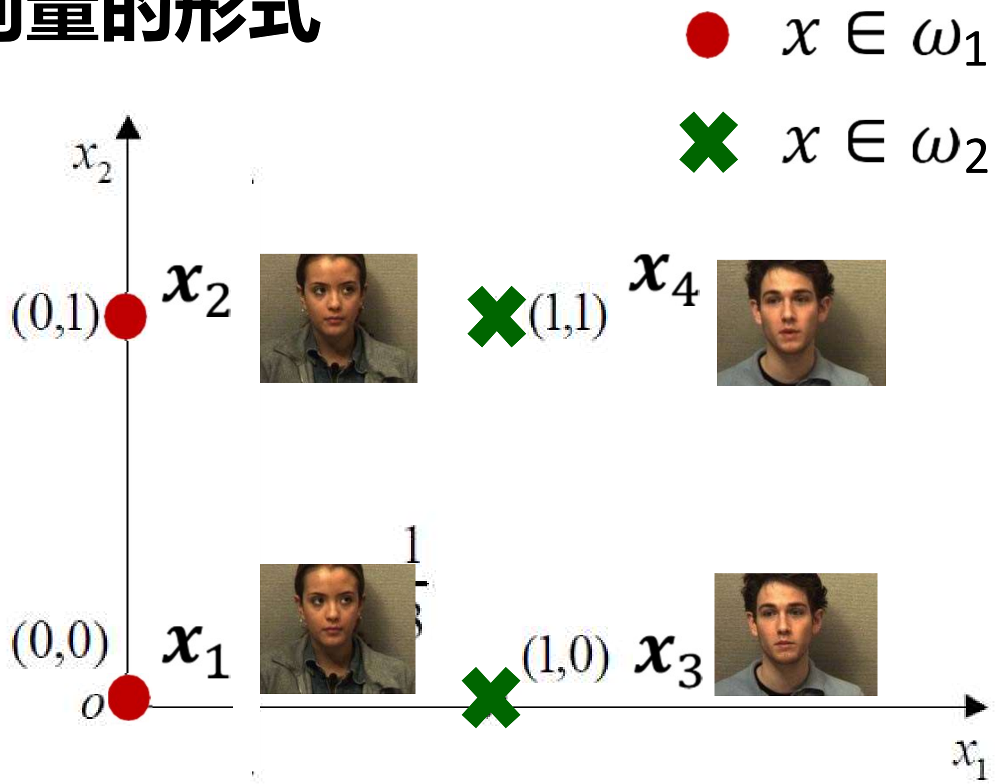
每一次迭代中$w^{(k+1)}$比$w^{(k)}$更接近最佳参数解。

3.5 感知器算法

例：试用感知器算法求出将图所示的模式分为两类的权向量解。显然，该模式是线性可分的，所以感知器算法分类肯定能成功。

将训练样本写成增广向量的形式

$\boldsymbol{x}_1 = (0, 0, 1)^{\mathrm{T}}$,
$\boldsymbol{x}_2 = (0, 1, 1)^{\mathrm{T}}$,
$\boldsymbol{x}_3 = (1, 0, 1)^{\mathrm{T}}$,
$\boldsymbol{x}_4 = (1, 1, 1)^{\mathrm{T}}$。

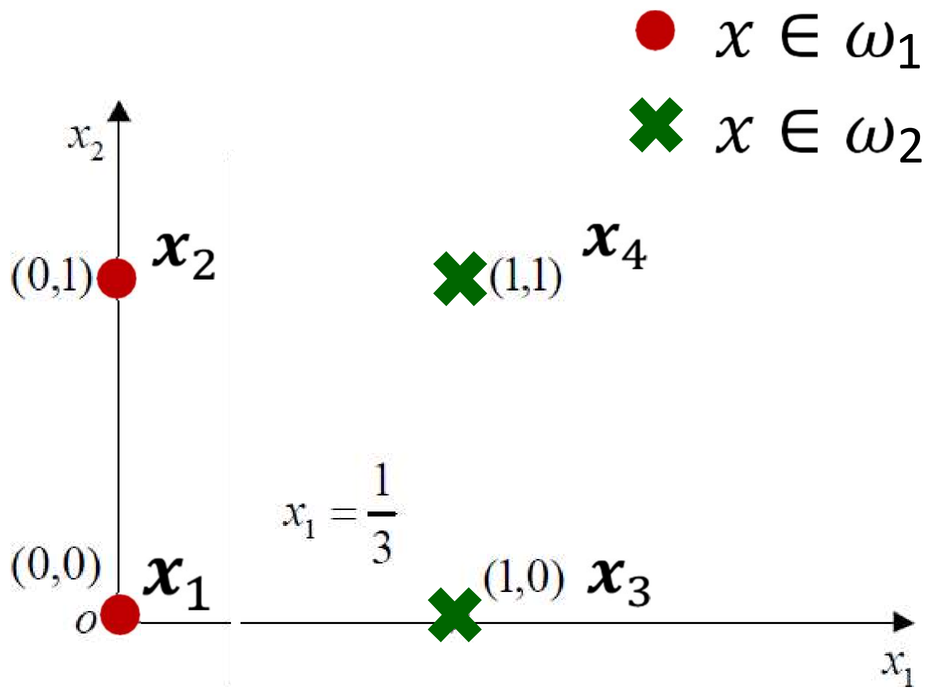


3.5 感知器算法

第一轮迭代取$C=1, w^{(0)}=0$，则迭代过程为：

$w^{(0)^T} x_1 = (0,0,0)^T (0,0,1) = 0$, 不大于**0**

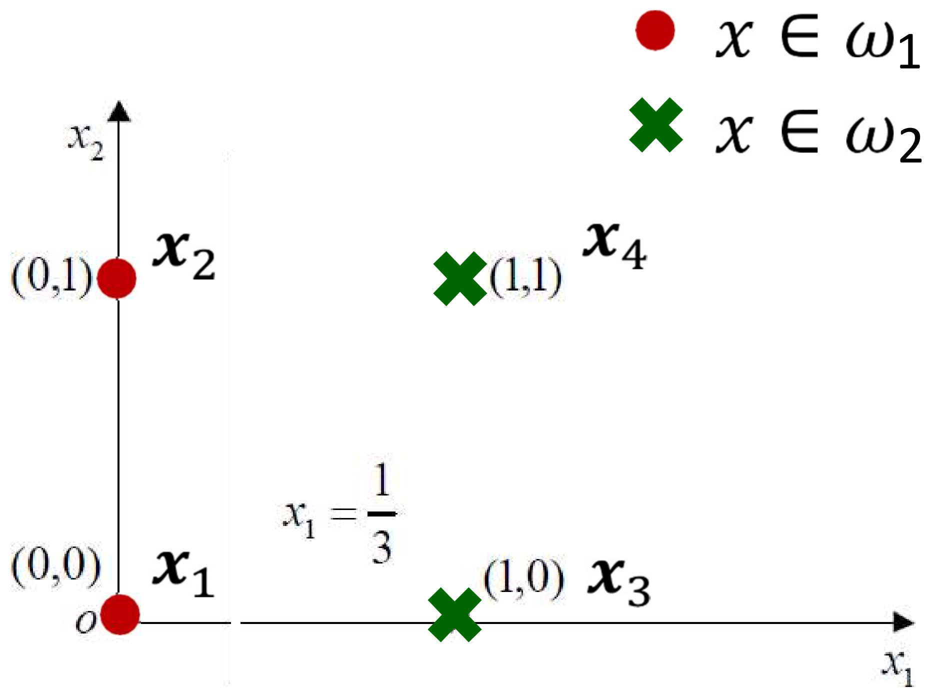
所以$w^{(1)} = w^{(0)} + x_1 = (0,0,1)^T$



3.5 感知器算法

$\boldsymbol{w}^{(1)^T}\boldsymbol{x}_2 = (0,0,1)^T(0,1,1) = 1$, 大于0

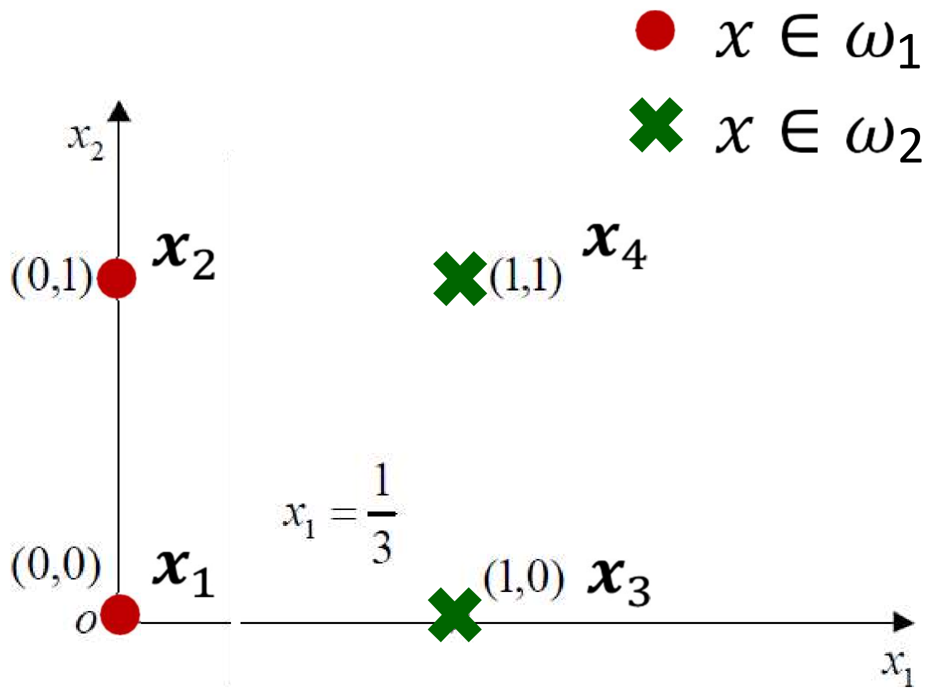
所以$\boldsymbol{w}^{(2)} = \boldsymbol{w}^{(1)} = (0,0,1)^T$



3.5 感知器算法

$\boldsymbol{w}^{(2)^T}\boldsymbol{x}_3 = (0,0,1)^T(1,0,1) = 1$，大于0

所以$\boldsymbol{w}^{(3)} = \boldsymbol{w}^{(2)} - \boldsymbol{x}_3 = (-1,0,0)^T$

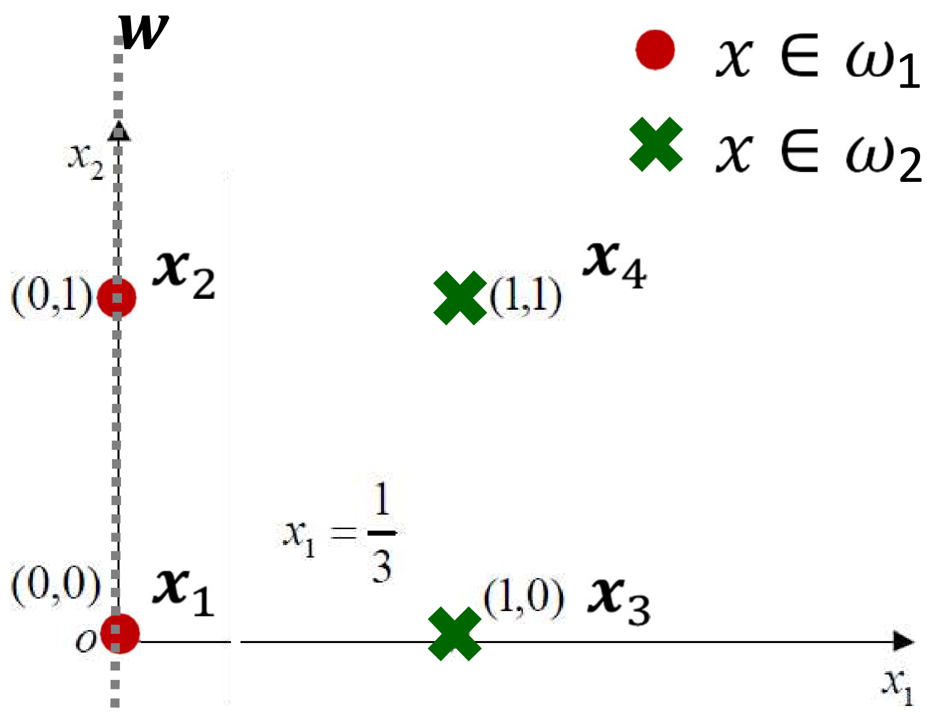


3.5 感知器算法

$w^{(3)^T} x_4 = (-1, 0, 0)^T (1, 1, 1) = -1$，小于0

所以$w^{(4)} = w^{(3)} = (-1, 0, 0)^T$

这里，第1和第3步为错误分类，应“罚”。

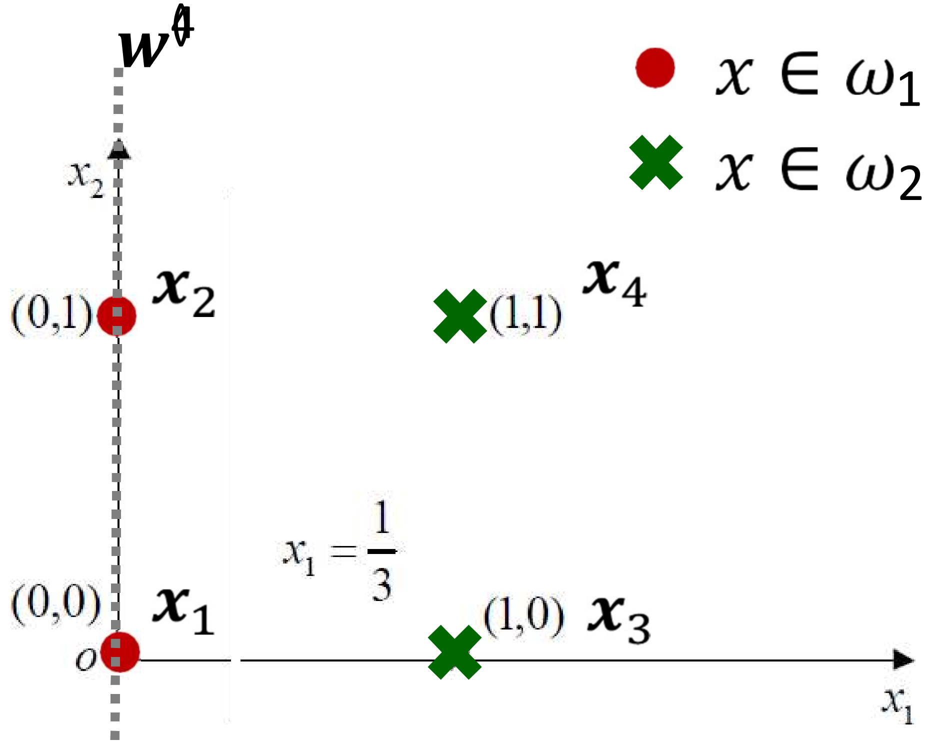


3.5 感知器算法

因为只有对全部模式都能正确判别的权向量才是正确的解，故需进行第二轮迭代。

$\boldsymbol{w}^{(4)^T}\boldsymbol{x}_1 = (-1,0,0)^T(0,0,1) = 0$, 不大于**0**

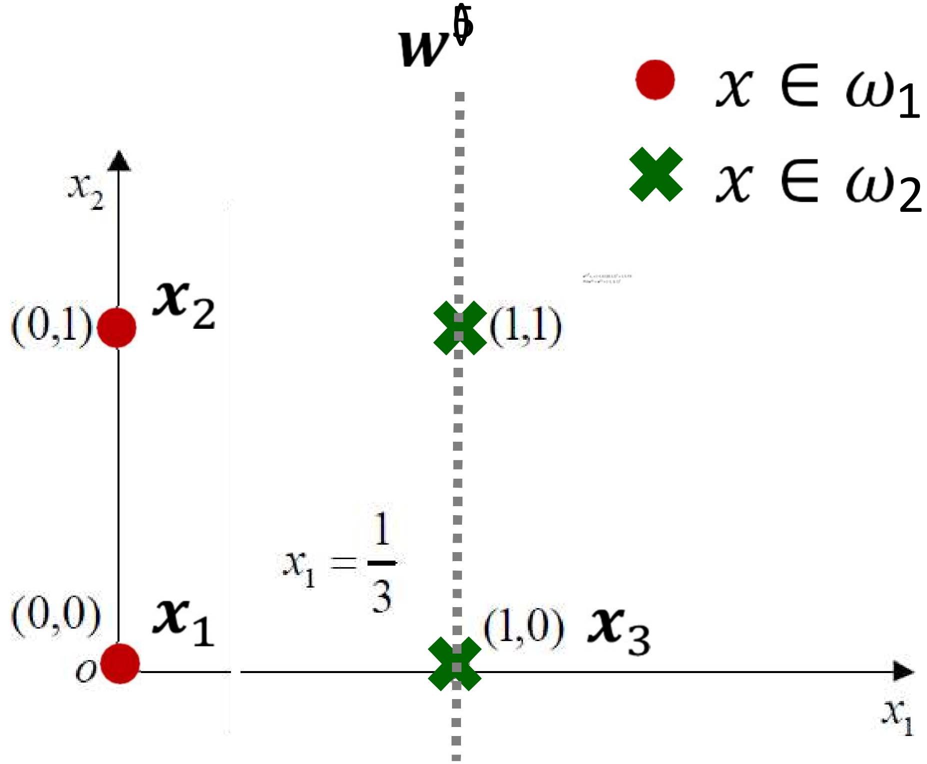
所以$\boldsymbol{w}^{(5)} = \boldsymbol{w}^{(4)} + \boldsymbol{x}_1 = (-1,0,1)^T$



3.5 感知器算法

$w^{(5)^T} x_2 = (-1, 0, 1)^T (0, 1, 1) = 1$，大于0

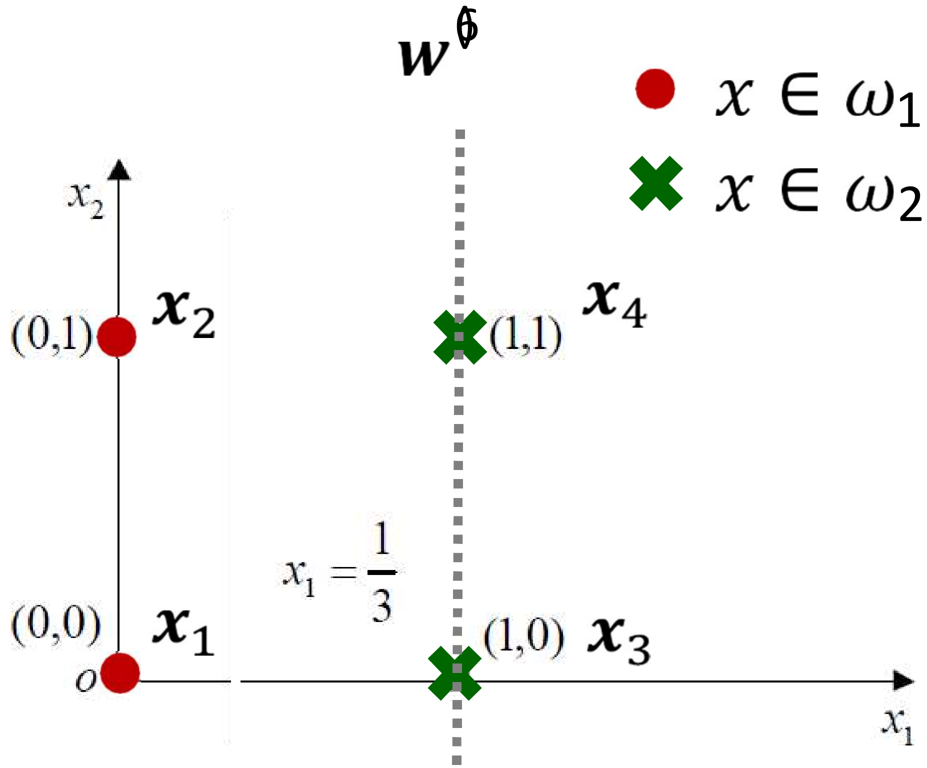
所以$w^{(6)} = w^{(5)} = (-1, 0, 1)^T$



3.5 感知器算法

$w^{(6)^T} x_3 = (-1,0,1)^T(1,0,1) = 0$, 不小于0

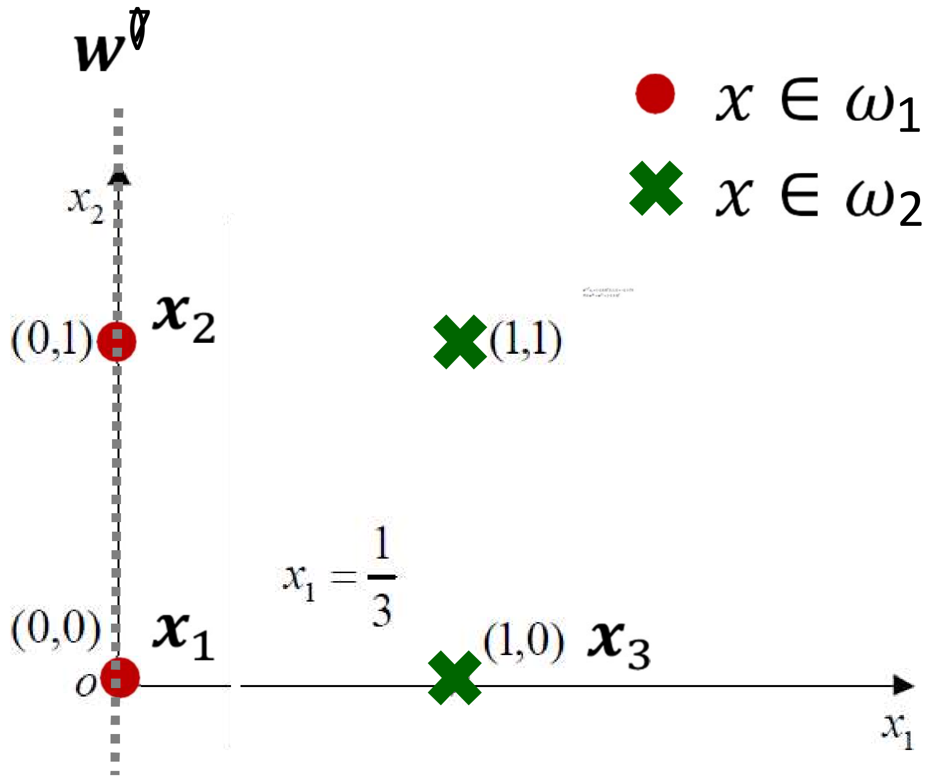
所以$w^{(7)} = w^{(6)} - x_3 = (-2,0,0)^T$



3.5 感知器算法

$w^{(7)^T} x_4 = (-2, 0, 0)^T (1, 1, 1) = -2$, 小于0

所以$w^{(8)} = w^{(7)} = (-2, 0, 0)^T$

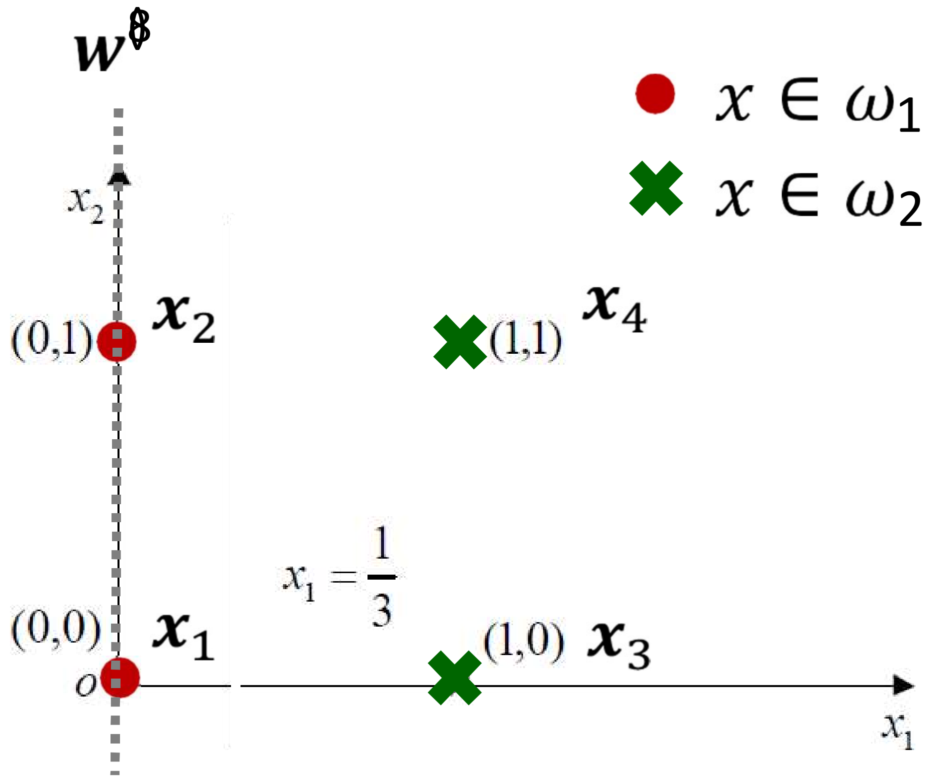


3.5 感知器算法

同样，需要进行第三轮迭代：

$\boldsymbol{w}^{(8)^T}\boldsymbol{x}_1=(-2,0,0)^T(0,0,1)=1$, 不大于**0**

所以$\boldsymbol{w}^{(9)}=\boldsymbol{w}^{(8)}+\boldsymbol{x}_1=(-2,0,1)^T$



3.5 感知器算法

$w^{(9)^T} x_2 = 1$，大于0

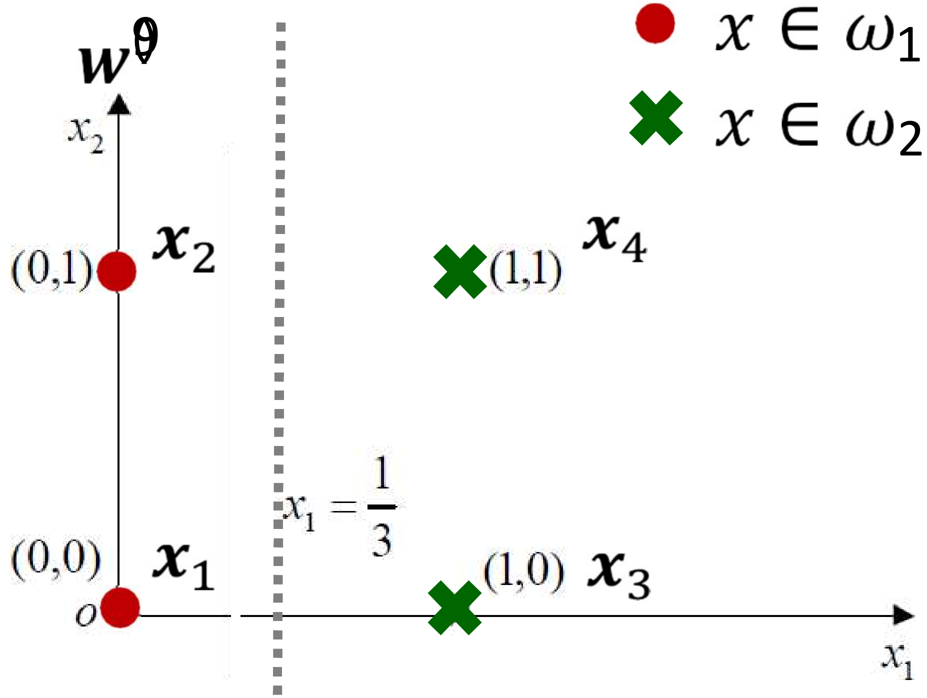
所以$w^{(10)} = w^{(9)} = (-2, 0, 1)^T$

$w^{(10)^T} x_3 = -1$，小于0

所以$w^{(11)} = w^{(10)} = (-2, 0, 1)^T$

$w^{(11)^T} x_4 = -1$，小于0

所以$w^{(12)} = w^{(11)} = (-2, 0, 1)^T$



3.5 感知器算法

第四轮迭代：

$\boldsymbol{w}^{(12)^T}\boldsymbol{x}_1 = 1$，所以$\boldsymbol{w}^{(13)} = \boldsymbol{w}^{(12)}$

$\boldsymbol{w}^{(13)^T}\boldsymbol{x}_2 = 1$，所以$\boldsymbol{w}^{(14)} = \boldsymbol{w}^{(13)}$

$\boldsymbol{w}^{(14)^T}\boldsymbol{x}_3 = -1$，所以$\boldsymbol{w}^{(15)} = \boldsymbol{w}^{(14)}$

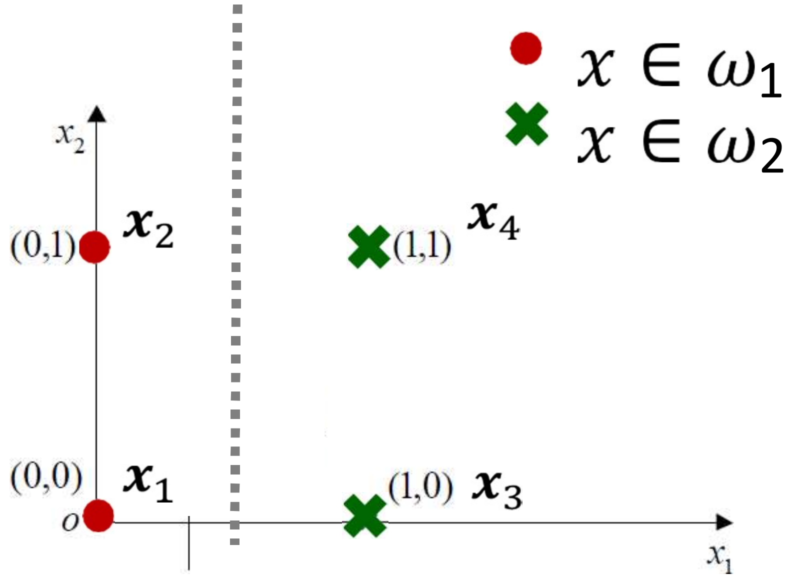
$\boldsymbol{w}^{(15)^T}\boldsymbol{x}_4 = -1$，所以$\boldsymbol{w}^{(16)} = \boldsymbol{w}^{(15)}$

该轮迭代的分类全部正确，

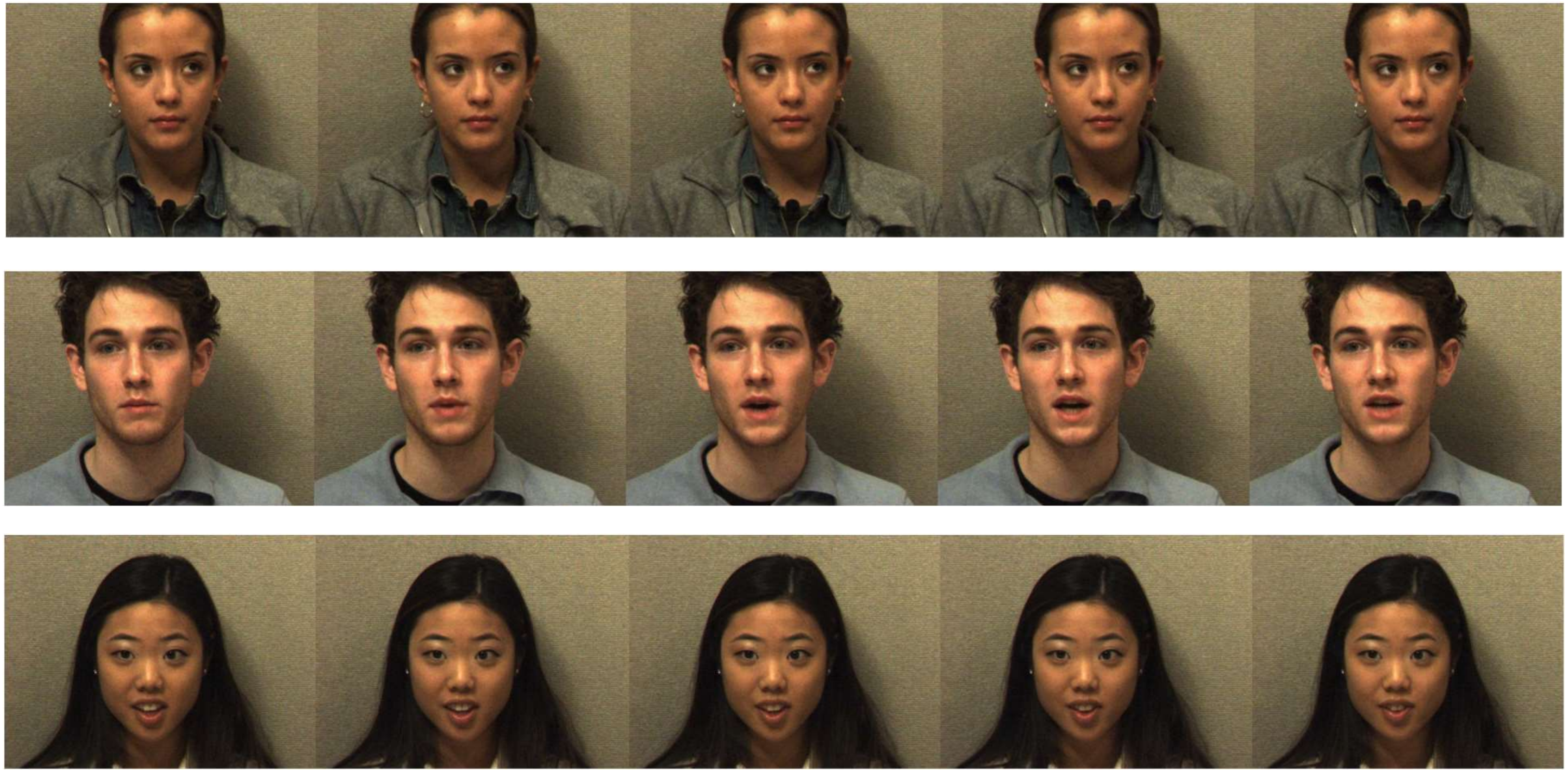
故解向量$\boldsymbol{w} = (-2, 0, 1)^T$

相应的判别函数为

$\boldsymbol{g}(\boldsymbol{x}) = -2\boldsymbol{x}_1 + 1$



线性判别函数



三类(人脸识别)

线性判别函数

情况3：

对于M类情况，有M个判别函数，即：

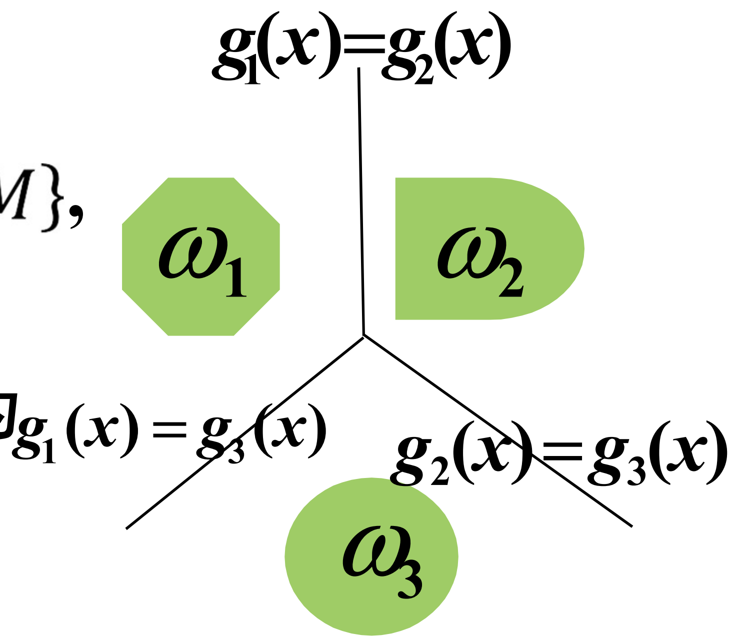
$g_i(x) = w_i^T x$ (增广形式)

如果

$g_k(x) = max\{g_i(x), i = 1,2, \ldots, M\}$,

则 x属于第k类。

这类分类特点是把M类情况分为M－１个两类问题，这类方法的 特点是不存在不确定区域。



3.5 感知器算法

采用感知器算法的多类模式分类：

设有M种模式类别$\omega_1,\omega_2,\ldots,\omega_M$ 若在训练过程的第k次迭代时，一个属于ω_i的模式样本x_n送入分类器，则应该先计算出M个线性判别函数：

$$g_j^{(k)}(\boldsymbol{x}_n) = {w_j^{(k)}}^T(\boldsymbol{x}_n) \quad j = 1,2,\ldots,M$$

如果：

$$g_i^{(k)}(\boldsymbol{x}_n) > g_j^{(k)}(\boldsymbol{x}_n), \quad j = 1,2,\ldots,M, j \neq i$$

的条件成立，则权向量不变

$$w_j^{(k+1)} = w_j^{(k)} \qquad j = 1,2,\ldots,M$$

3.5 感知器算法

如果其中第l个权向量使得 $g_i^{(k)}(\boldsymbol{x}_n) \ngtr g_j^{(k)}(\boldsymbol{x}_n)$

，则权向量应该作相应的调整，即：

$$w_i^{(k+1)} = w_i^{(k)} + C\boldsymbol{x}_n$$

$$w_l^{(k+1)} = w_l^{(k)} - C\boldsymbol{x}_n$$

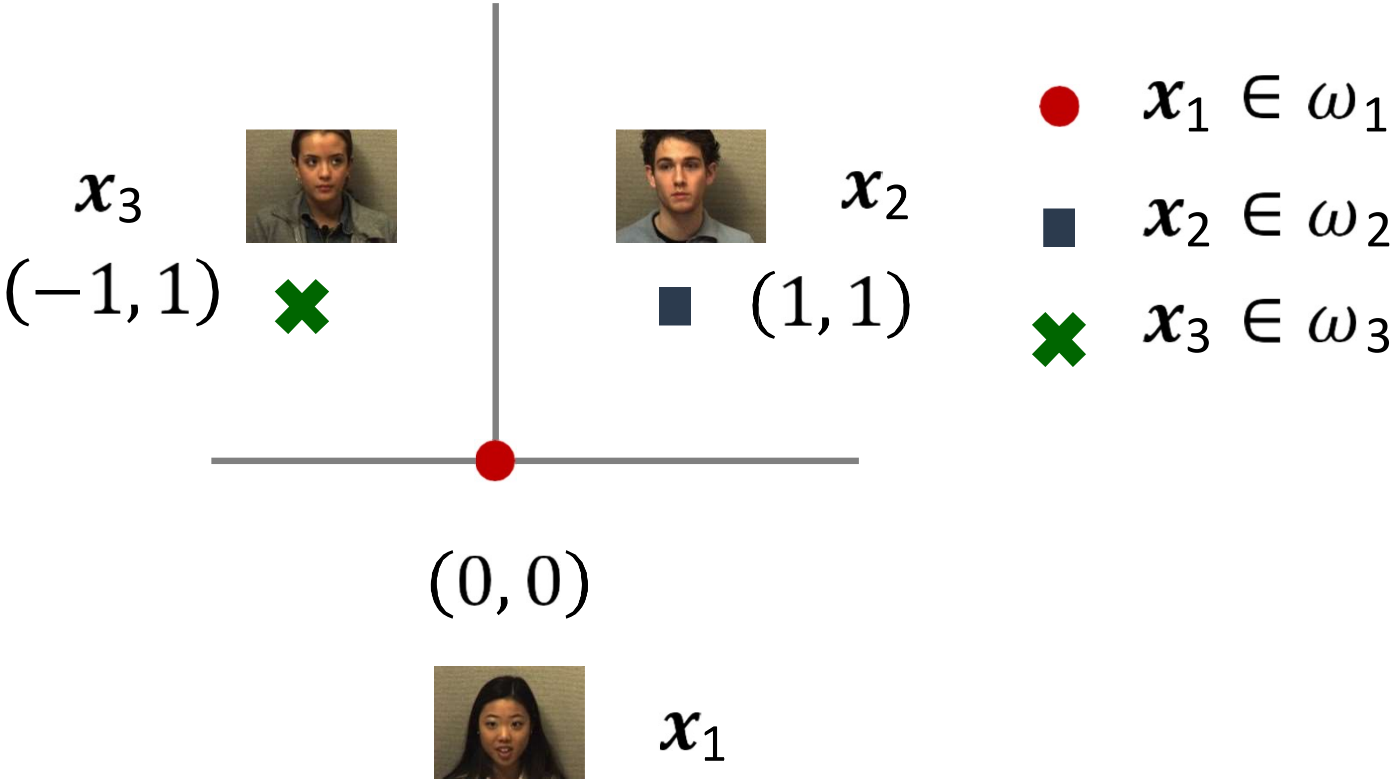
$$w_j^{(k+1)} = w_j^{(k)}, \mathrm{j} = 1,2, \ldots, \mathrm{M}, \mathrm{j} \neq i, i \neq l$$

其中C为一正常数。

3.5 感知器算法

给出三种模式，每类的训练样本为：

$$\omega_1: \{(0,0)^\mathrm{T}\}, \omega_2: \{(1,1)^\mathrm{T}\}, \omega_3: \{(-1,1)^\mathrm{T}\}$$



3.5 感知器算法

为采用一般化感知器算法，将模式样本写成增广形式，即

$$x_1 = (0,0,1)^{\mathrm{T}}$$
$$x_2 = (1,1,1)^{\mathrm{T}}$$
$$x_3 = (-1,1,1)^{\mathrm{T}}$$

取初始值$w_1^{(0)} = w_2^{(0)} = w_3^{(0)} = (0,0,0)^T$,取C=1，按上述步骤计算。

3.5 感知器算法

第一次迭代(k=0),以x_1作为**训练样本，即**

$$g_1^{(0)}(x_1) = {w_1^{(0)}}^T x_1 = 0$$

$$g_2^{(0)}(x_1) = {w_2^{(0)}}^T x_1 = 0$$

$$g_3^{(0)}(x_1) = {w_3^{(0)}}^T x_1 = 0$$

因$g_1^{(0)}(x_1) \ngtr g_2^{(0)}(x_1)$，$g_1^{(0)}(x_1) \ngtr g_3^{(0)}(x_1)$，**故**

$$w_1^{(1)} = w_1^{(0)} + x_1 = (0, 0, 1)^T$$

$$w_2^{(1)} = w_2^{(0)} - x_1 = (0, 0, -1)^T$$

$$w_3^{(1)} = w_3^{(0)} - x_1 = (0, 0, -1)^T$$

3.5 感知器算法

第二次迭代(k=1),以x_2=(1, 1, 1)^{T}作为训练样本，即

$g_1^{(1)}(x_2) = {w_1^{(1)}}^T x_2$=1

$g_2^{(1)}(x_2) = {w_2^{(1)}}^T x_2$=-1

$g_3^{(1)}(x_2) = {w_3^{(1)}}^T x_2$=-1

因$g_2^{(1)}(x_2) \ngtr g_1^{(1)}(x_2)$，$g_2^{(1)}(x_2) \ngtr g_3^{(1)}(x_2)$，**故**

$$w_1^{(2)} = w_1^{(1)} - x_2 = (-1, -1, 0)^T$$

$$w_2^{(2)} = w_2^{(1)} + x_2 = (1, 1, 0)^T$$

$$w_3^{(2)} = w_3^{(1)} - x_2 = (-1, -1, -2)^T$$

3.5 感知器算法

第三次迭代(k=2),以$x_3=(-1,1,1)^T$ 作为训练样本，即

$g_1^{(2)}(x_3)=w_1^{(2)T}x_3$=-1

$g_2^{(2)}(x_3)=w_2^{(2)T}x_3$=-1

$g_3^{(2)}(x_3)=w_3^{(2)T}x_3$=-1

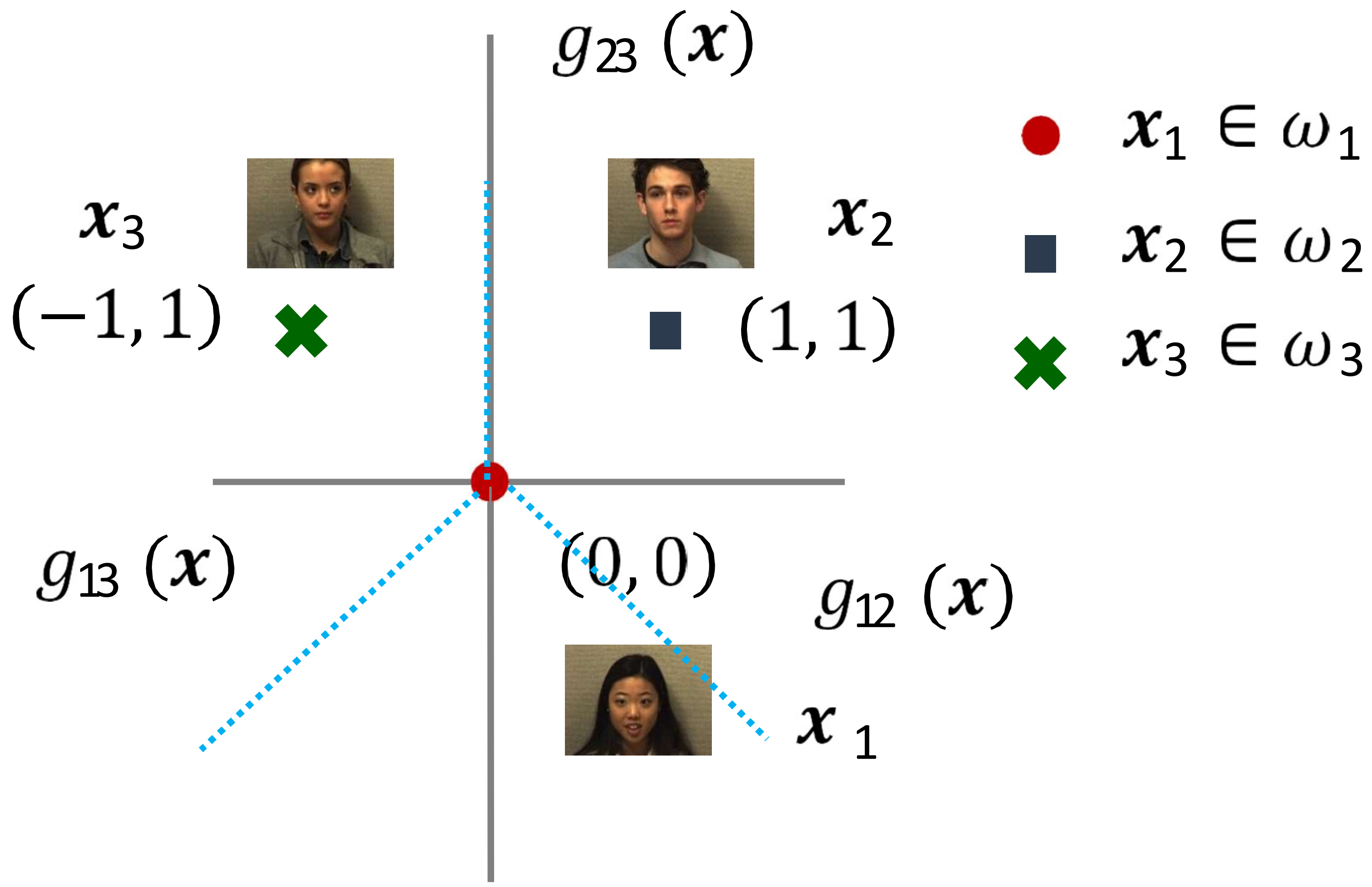
因$g_3^{(2)}(x_3) \ngtr g_1^{(2)}(x_3)$，$g_3^{(2)}(x_3) \ngtr g_2^{(2)}(x_3)$，**故**

$$w_1^{(3)}=w_1^{(2)}-x_3=(0,-2,-1)^T$$

$$w_2^{(3)}=w_2^{(2)}-x_3=(2,0,-1)^T$$

$$w_3^{(3)}=w_3^{(2)}+x_3=(-2,0,-1)^T$$

3.5 感知器算法


$g_{23}(\boldsymbol{x})$
$\boldsymbol{x}_1 \in \omega_1$
$\boldsymbol{x}_2 \in \omega_2$
$\boldsymbol{x}_3 \in \omega_3$
$\boldsymbol{x}_3$
$\boldsymbol{x}_2$
$(-1, 1)$
$(1, 1)$
$g_{13}(\boldsymbol{x})$
$(0, 0)$
$g_{12}(\boldsymbol{x})$
$\boldsymbol{x}_1$

3.5 感知器算法

第四次迭代(k=3),以$x_3=(0,0,1)^T$作为训练样本，即

$g_1^{(3)}(x_1)={w_1^{(3)}}^T x_1$=-1

$g_2^{(3)}(x_1)={w_2^{(3)}}^T x_1$=-1

$g_3^{(3)}(x_1)={w_3^{(3)}}^T x_1$=-1

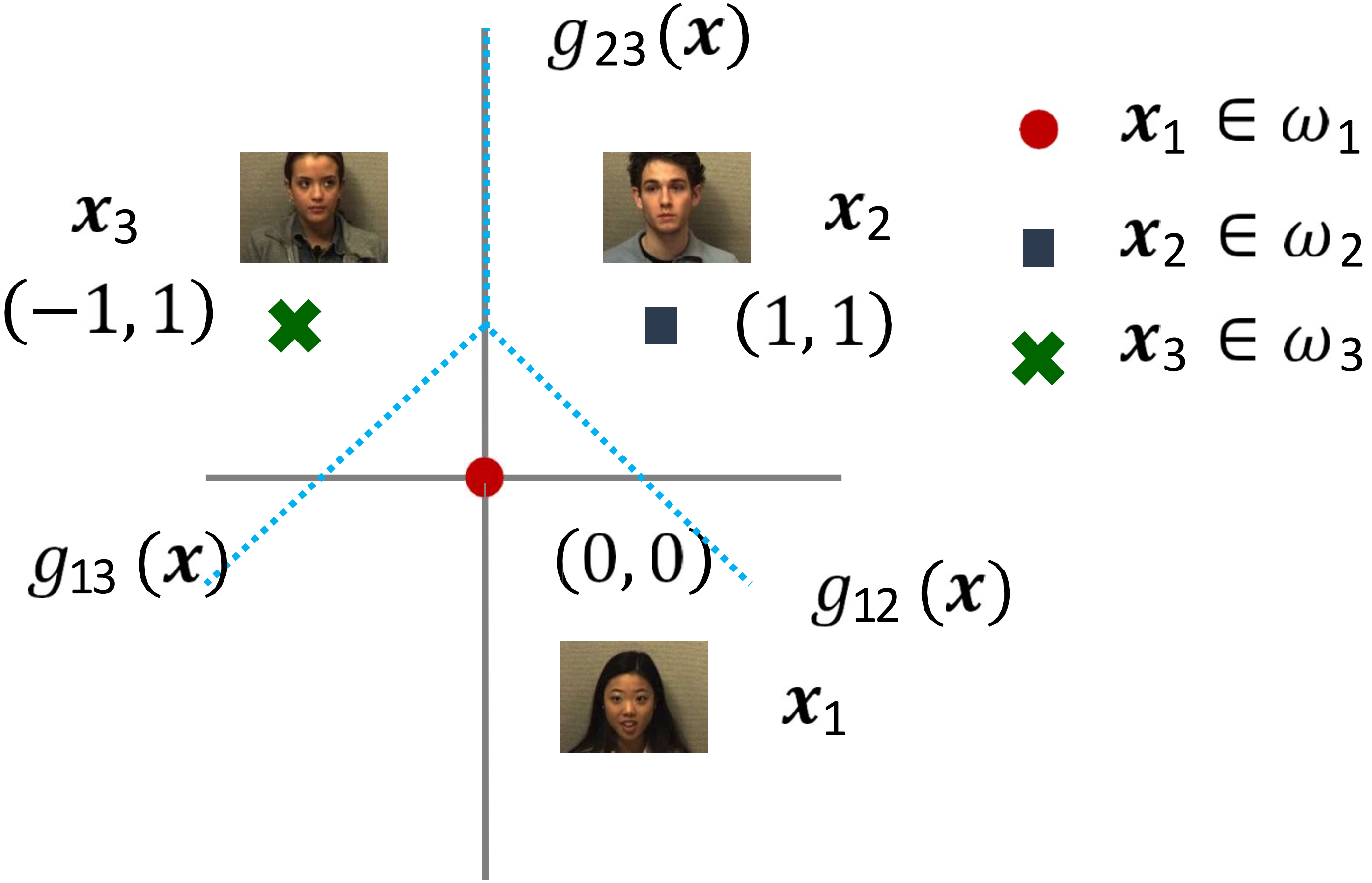
$x\in\omega_1$但$g_1^{(3)}(x_1)\ngtr g_2^{(3)}(x_1)$，$g_1^{(3)}(x_1)\ngtr g_3^{(3)}(x_1)$，**故**

$$w_1^{(4)}=w_1^{(3)}+x_1=(0,-2,0)^T$$

$$w_2^{(4)}=w_2^{(3)}-x_1=(2,0,-2)^T$$

$$w_3^{(4)}=w_3^{(3)}-x_{x_1}=(-2,0,-2)^T$$

3.5 感知器算法



$g_{23}(\boldsymbol{x})$
$\boldsymbol{x}_3$
$(-1, 1)$
$\boldsymbol{x}_2$
$(1, 1)$
$g_{13}(\boldsymbol{x})$
$(0, 0)$
$g_{12}(\boldsymbol{x})$
$\boldsymbol{x}_1$
$\boldsymbol{x}_1 \in \omega_1$
$\boldsymbol{x}_2 \in \omega_2$
$\boldsymbol{x}_3 \in \omega_3$

3.5 感知器算法

第五次迭代(k=4),以$x_2=(1,1,1)^T$作为训练样本，即

$g_1^{(4)}(x_2)={w_1^{(4)}}^T x_2$=-2

$g_2^{(4)}(x_2)={w_2^{(4)}}^T x_2$=0

$g_3^{(4)}(x_2)={w_3^{(4)}}^T x_2$=-4

因$g_2^{(4)}(x_2)>g_1^{(4)}(x_2)$，$g_2^{(4)}(x_2)>g_3^{(4)}(x_2)$，**故**

$$w_1^{(5)}=w_1^{(4)}$$
$$w_2^{(5)}=w_2^{(4)}$$
$$w_3^{(5)}=w_3^{(4)}$$

3.5 感知器算法

第六次迭代(k=5),以$x_3=(-1,1,1)^T$ 作为**训练样本，即**

$g_1^{(5)}(x_3)=w_1^{(5)^T}x_3$=-2

$g_2^{(5)}(x_3)=w_2^{(5)^T}x_3$=-4

$g_3^{(5)}(x_3)=w_3^{(5)^T}x_3$=0

模式已分类正确，故

$$w_1^{(6)}=w_1^{(5)}$$
$$w_2^{(6)}=w_2^{(5)}$$
$$w_3^{(6)}=w_3^{(5)}$$

3.5 感知器算法

第七次迭代(k=6),以$x_1=(0,0,1)^T$作为**训练样本，即**

$g_1^{(6)}(x_1)=w_1^{(6)^T}x_1$**=0**

$g_2^{(6)}(x_1)=w_2^{(6)^T}x_1$**=-2**

$g_3^{(6)}(x_1)=w_3^{(6)^T}x_1$**=-2**

模式已分类正确，故权向量不变。

3.5 感知器算法

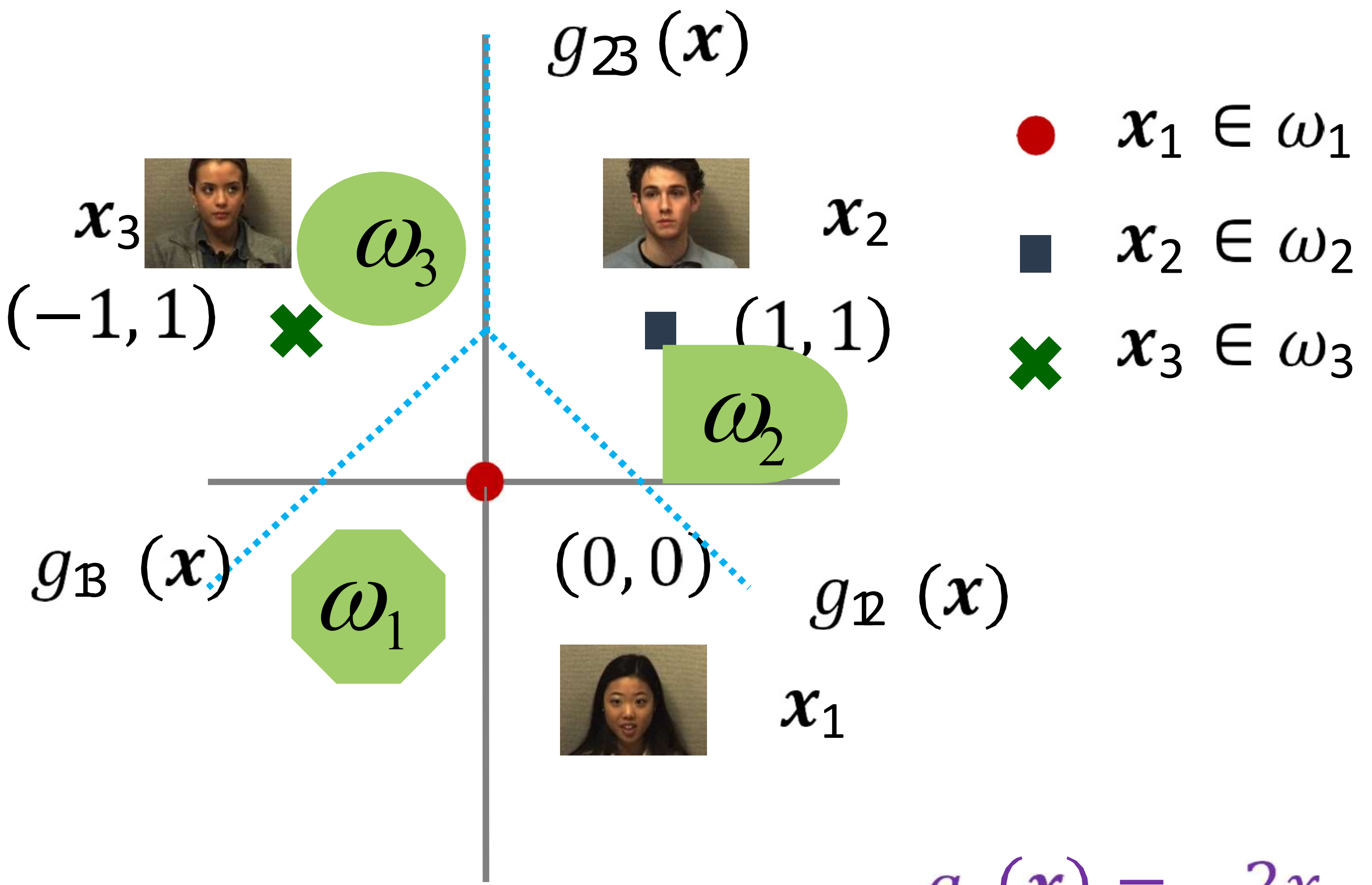
由于第五、六、七次迭代中x_1**，** x_2**，** x_3都已**正确分类，故权向量的解即为**

$$w_1 = w_1^{(6)} = w_1^{(5)} = w_1^{(5)} = (0, -2, 0)^T$$
$$w_2 = w_2^{(6)} = w_2^{(5)} = w_2^{(5)} = (2, 0, -2)^T$$
$$w_3 = w_3^{(6)} = w_3^{(5)} = w_3^{(5)} = (-2, 0 - 2)^T$$

因此得三个判别函数

$$g_1(x) = -2x_2$$
$$g_2(x) = 2x_1 - 2$$
$$g_3(x) = -2x_1 - 2$$

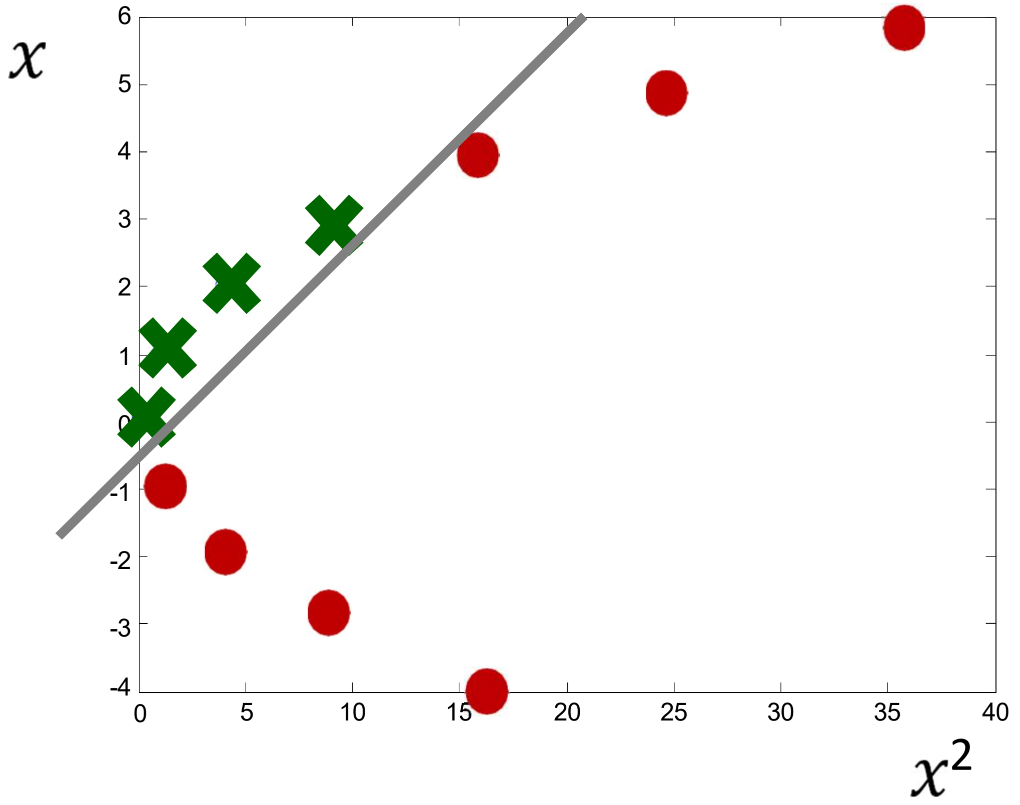
3.5 感知器算法

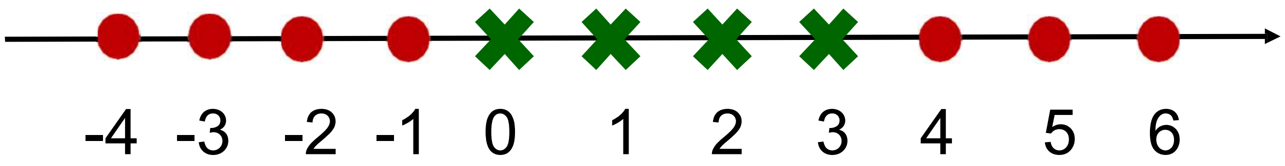


$$g_1(\boldsymbol{x}) = -2x_2$$
$$g_2(\boldsymbol{x}) = 2x_1 - 2$$
$$g_3(\boldsymbol{x}) = -2x_1 - 2$$

小结

3.4广义线性判别函数





3.4广义线性判别函数

设原模式空间中的模式为$x \in \mathbb{R}^D$，变换后空间中的模式为$f(x)$,

则$f(x) = [f_1(x), f_2(x), \cdots, f_K(x)]^T \in \mathbb{R}^H$

下面讨论广义线性判别函数的意义。

3.5 感知器算法

感知器的训练算法如下：

1. **已知训练样本集**$\{x_1, x_2, \dots, x_N\}$以及**它们所属类别**ω_1和ω_2，**同时已知权向量初始值为**$w^{(0)}$
2. **第k次训练中，若**$x_n \in \omega_1$**,但**$w^{(k)^T} x_n \le 0$**,则分类器对第n个模式**x_n**做错误分类，应校正权向量：**

$$w^{(k+1)} = w^{(k)} + C x_n$$ **C为一个矫正量**

若$x_n \in \omega_2$**，且**$w^{(k)^T} x_n \ge 0$**，同样错误分类，应校正权向量：**

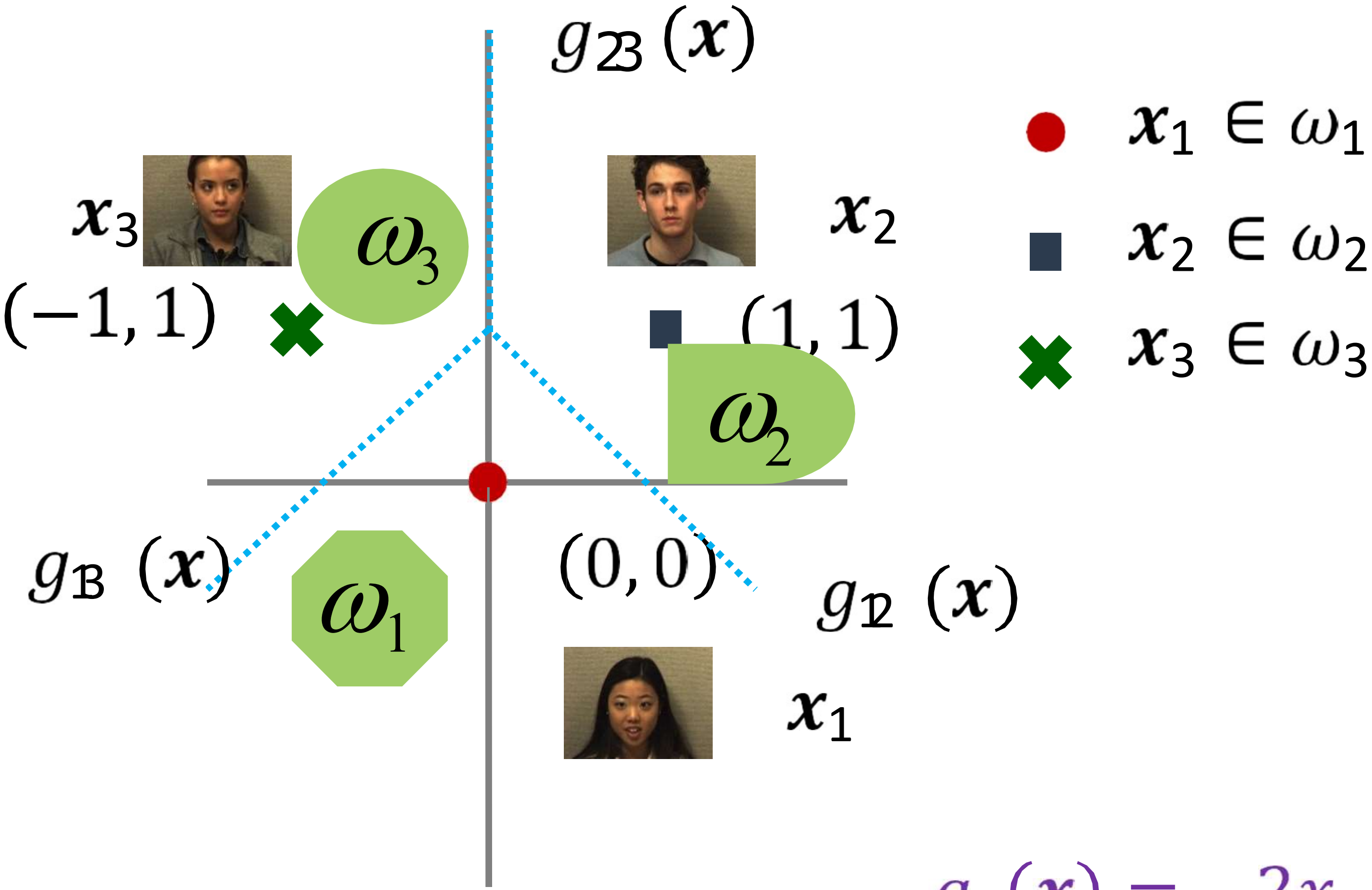
$$w^{(k+1)} = w^{(k)} - C x_n$$

如果x_n**不符合上述情况，则分类正确：**

$$w^{(k+1)} = w^{(k)}$$

3. **反复迭代至所有样本正确分类。**

3.5 感知器算法



$$g_1(\boldsymbol{x}) = -2x_2$$
$$g_2(\boldsymbol{x}) = 2x_1 - 2$$
$$g_3(\boldsymbol{x}) = -2x_1 - 2$$

End