第三章 判别域代数界面方程法

3.1 判别域界面方程分类的概念

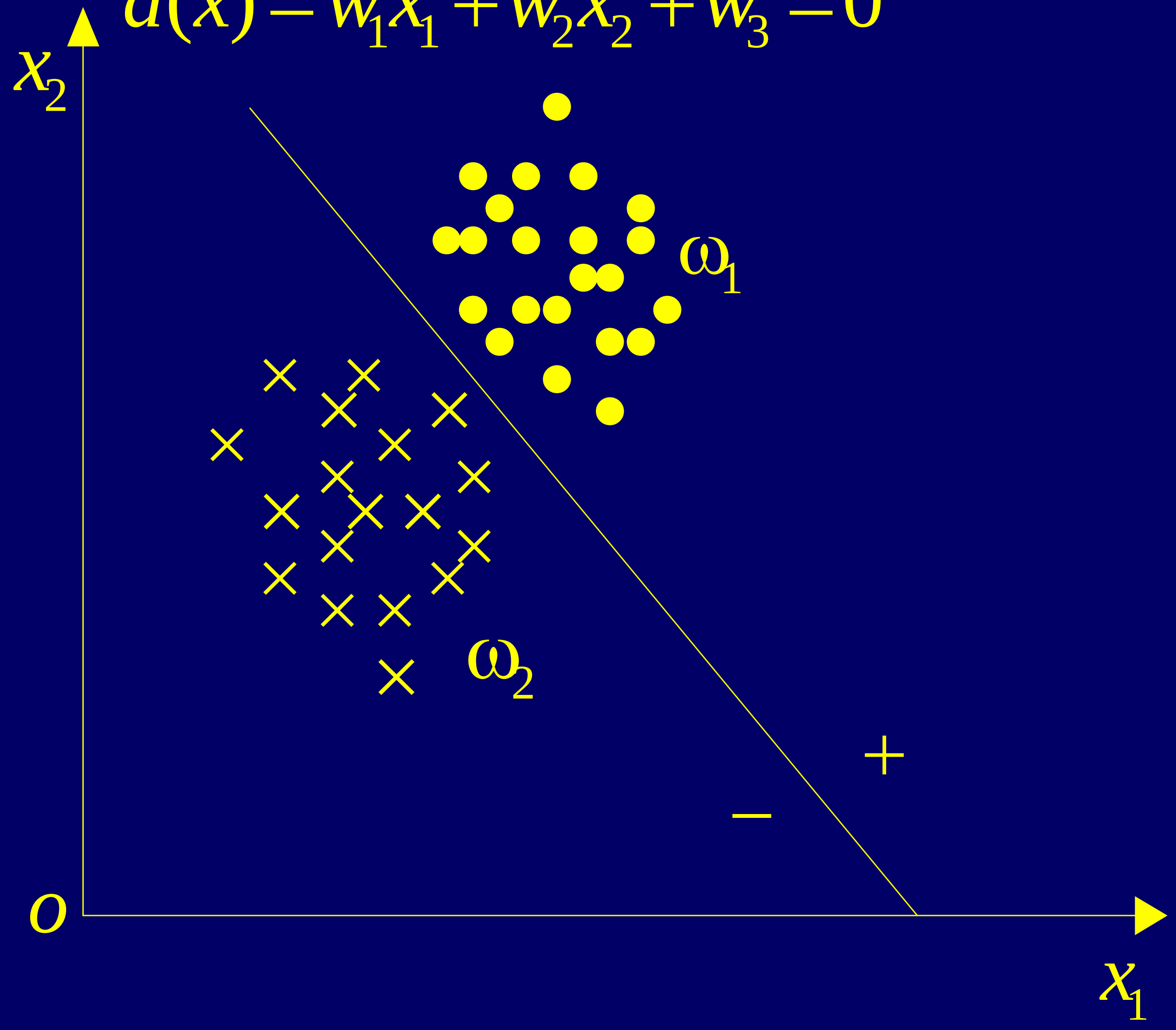
1. 分类的基本原理

不同模式对应特征点在不同的区域中散布。运用已知类别的训练样本进行学习，产生若干个代数界面 $d(\vec{x})=0$，将特征空间划分成一些互不重叠的子区域。

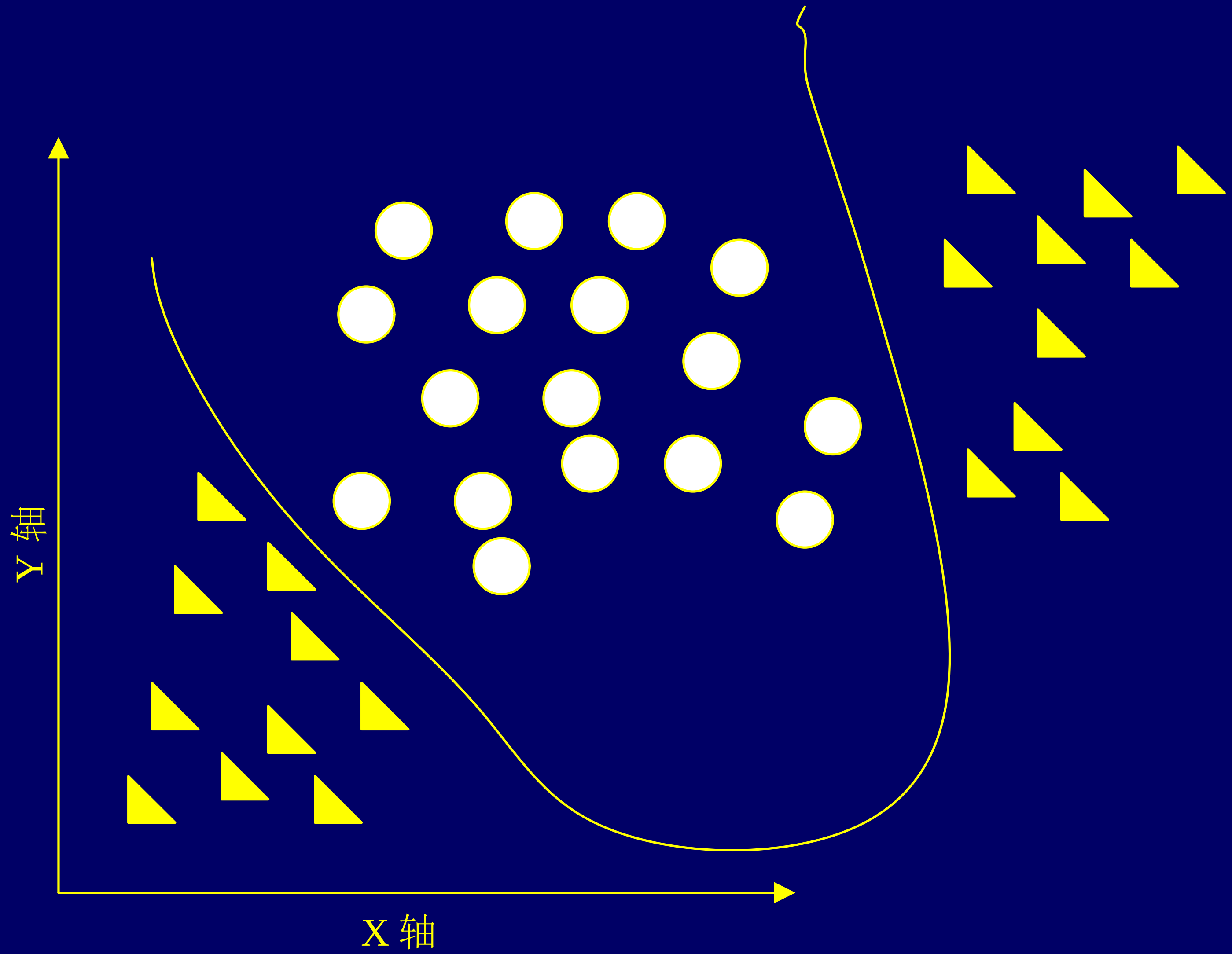
2. 判别函数

表示界面的函数 $d(\vec{x})$ 称为判别函数（Discriminant Function）。

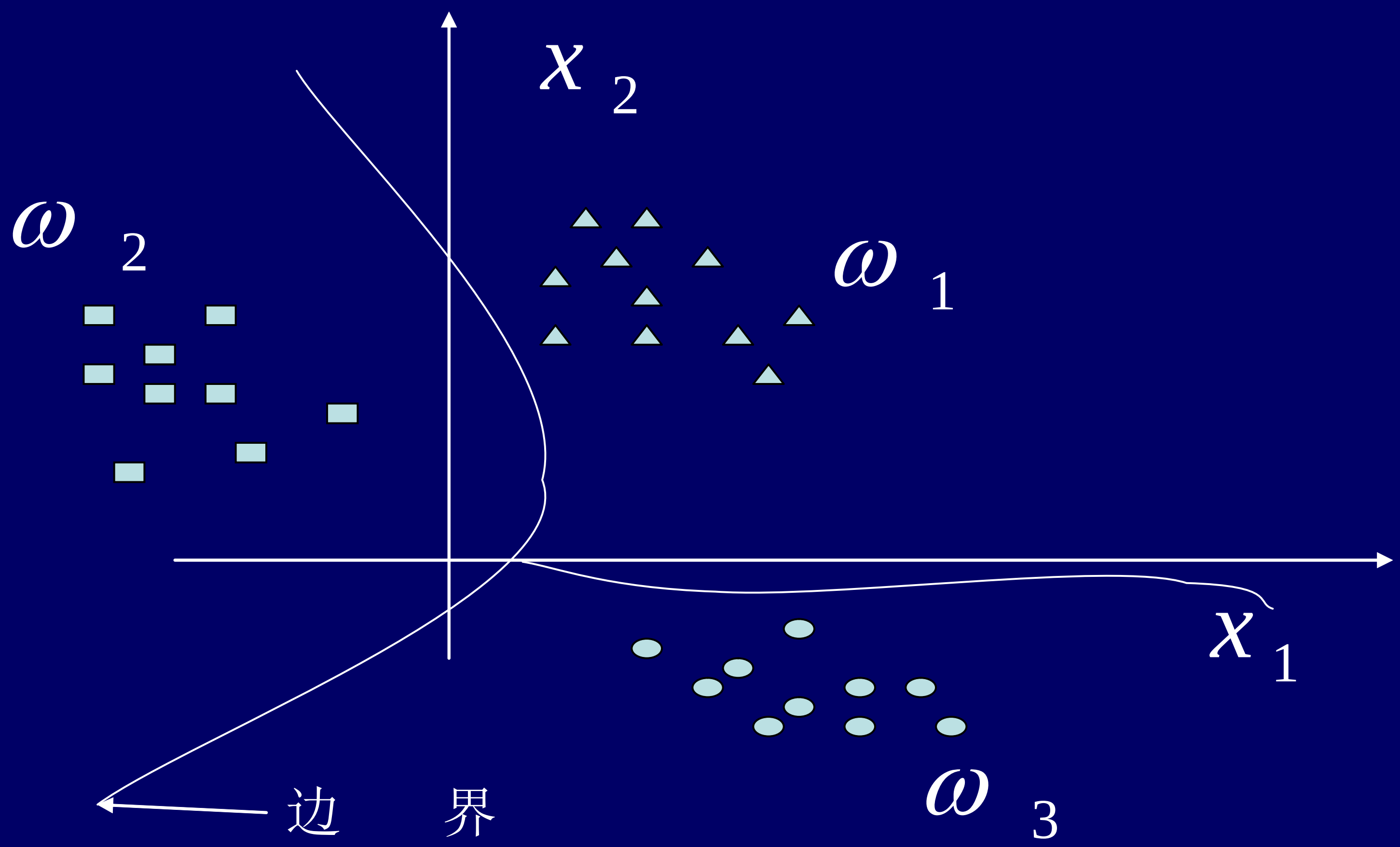
$$d(\vec{x}) = w_1 x_1 + w_2 x_2 + w_3 = 0$$



两类的分类问题，它们的边界线就是一个判别函数



两类问题中线性不可分的实例



三类的分类问题，它们的边界线也是一个判别函数

3. 线性可分的定义

对于来自两类的一组模式 $\vec{x}_1, \vec{x}_2, \ldots . \vec{x}_N$，如果能用一个线性判别函数正确分类，则称他们是线性可分的。

4. 本章分类方法的基本技术思路

第一步：利用训练样本求出分类器/判别函数

第二步：利用判别函数对未知类别样本分类

第三章 判别域代数界面方程法

3.1 判别域界面方程分类的概念

3.2 线性判别函数

3.3 判别函数值的鉴别意义、权空间及解空间

3.4 fisher线性判别

3.5 线性可分条件下判别函数权矢量算法

3.6 一般情况下的判别函数权矢量算法

3.7 非线性判别函数

3.8 最近邻方法

3.2 线性判别函数

在n维特征空间中，特征矢量为 $\vec{x}=(x_1,x_2,\cdots,x_n)'$，线性判别函数的一般形式是：

$$d(\vec{x})=w_1x_1+w_2x_2+\cdots+w_nx_n+w_{n+1}\triangleq\vec{w}_0'\vec{x}+w_{n+1}$$

其中 $\vec{w}_0=(w_1,w_2,....w_n)'$ 称为权矢量或系数矢量。

3.2 线性判别函数

在n维特征空间中，特征矢量为 $\vec{x}=(x_1,x_2,\cdots,x_n)'$，线性判别函数的一般形式是：

$$d(\vec{x})=w_1x_1+w_2x_2+\cdots+w_nx_n+w_{n+1}\triangleq\vec{w}_0'\vec{x}+w_{n+1}$$

其中 $\vec{w}_0=(w_1,w_2,\dots.w_n)'$ 称为权矢量或系数矢量。

简化为： $d(\vec{x})\triangleq\vec{w}'\vec{x}$

$$\vec{x}=(x_1,x_2,\dots x_n,1)' \qquad \vec{w}=(w_1,w_2,\dots.w_n,w_{n+1})'$$

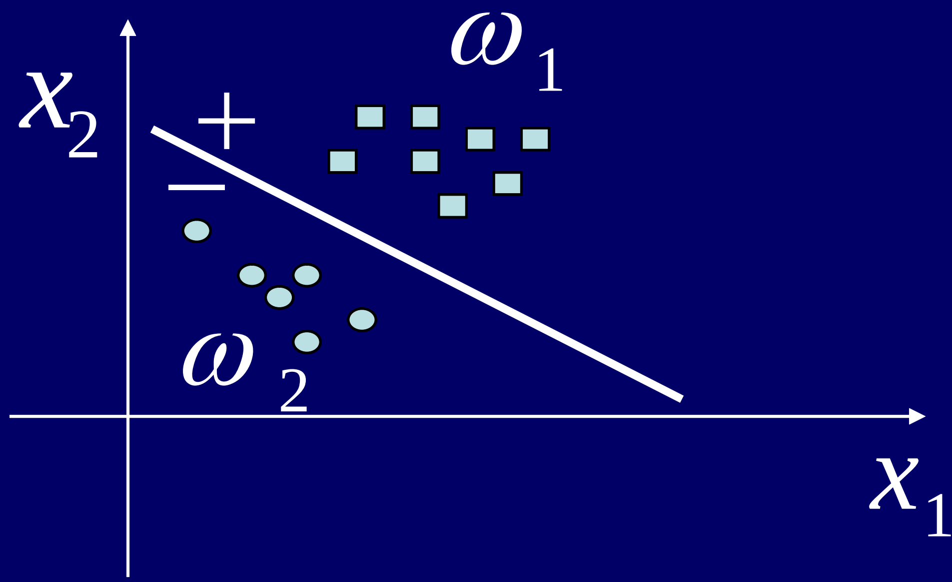
$\vec{x}$和$\vec{w}$分别称为增广特征矢量和增广权矢量。

3.2 线性判别函数

一、两类问题

$$d(\vec{x}) = w_1x_1 + w_2x_2 + w_3$$

对于两类问题，待识别模式增广特征矢量 $\vec{x}$ 可通过下面的判别规则进行分类识别：设 $d(\vec{x})$ 为判别函数，



$$d(\vec{x}) = \vec{w}'\vec{x} \begin{cases} > 0 \Rightarrow \vec{x} \in \omega_1 \\ < 0 \Rightarrow \vec{x} \in \omega_2 \\ = 0 \Rightarrow \vec{x} \in \omega_i \text{或拒判} \end{cases}$$

二、多类问题

处理多类问题主要有以下几种方法:

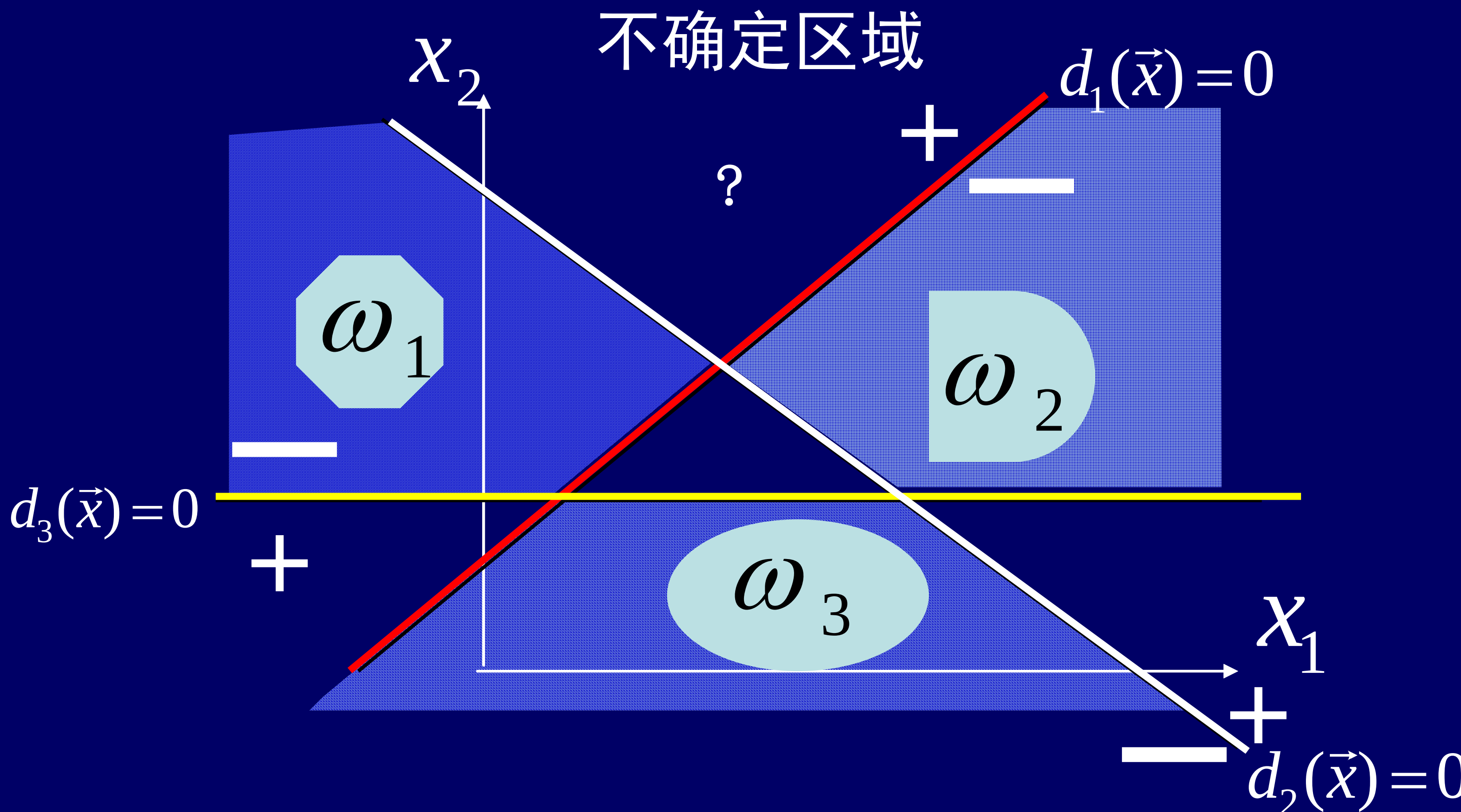
1. $\omega_i/\bar{\omega}_i$ 两分法（第一种情况）

基本思想:将属于 ω_i 类和不属于 ω_i 类的模式分划开。c 类问题转化为 $c-1$ 个两类问题。可建立 c 个判别函数。

$$d_i(\vec{x}) = \vec{w}_i^{'}\vec{x} \qquad (i=1,2,\cdots,c)$$

通过训练, 其中每个判别函数都具有下面的性质:

$$d_i(\vec{x})\begin{cases} >0, 若\vec{x}\in\omega_i, \\ <0, 若\vec{x}\in\bar{\omega}_i, \end{cases} \qquad (i=1,2,\cdots,c)$$



多类问题图例（第一种情况）

3.2　线性判别函数

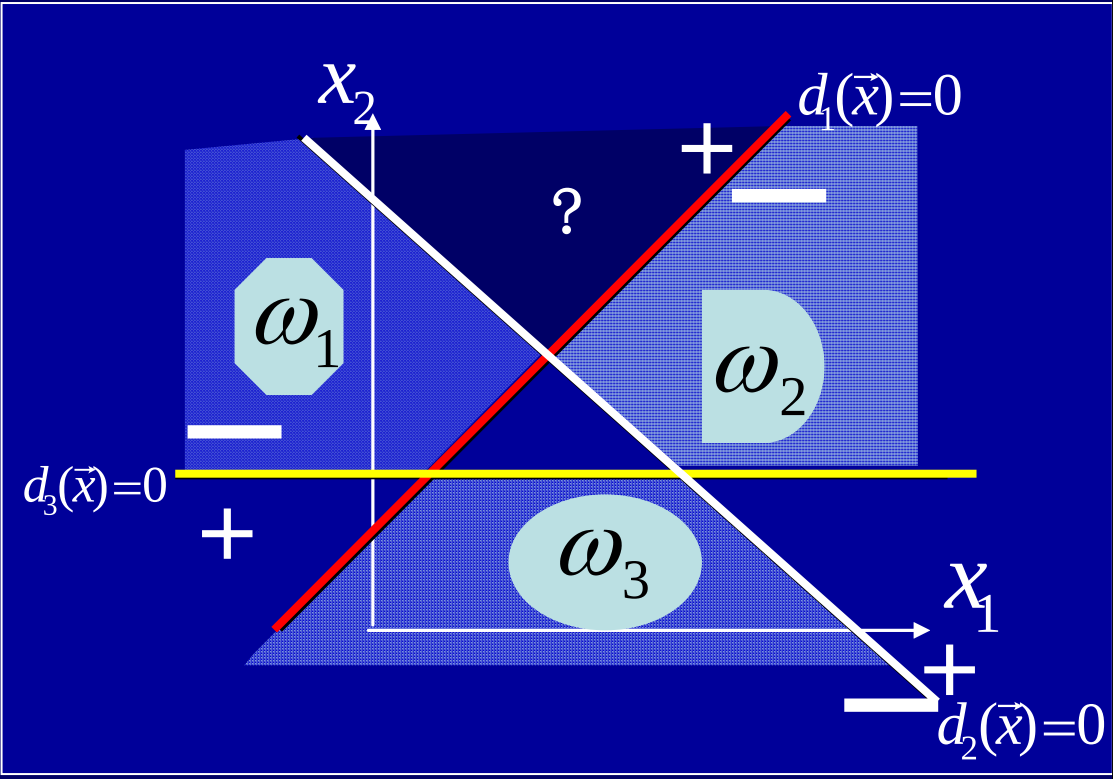
判别规则为：

如果 $\begin{cases} d_i(\vec{x}) > 0 \\ d_j(\vec{x}) \le 0 \qquad \forall j \ne i \end{cases}$ 则判 $\vec{x} \in \omega_i$

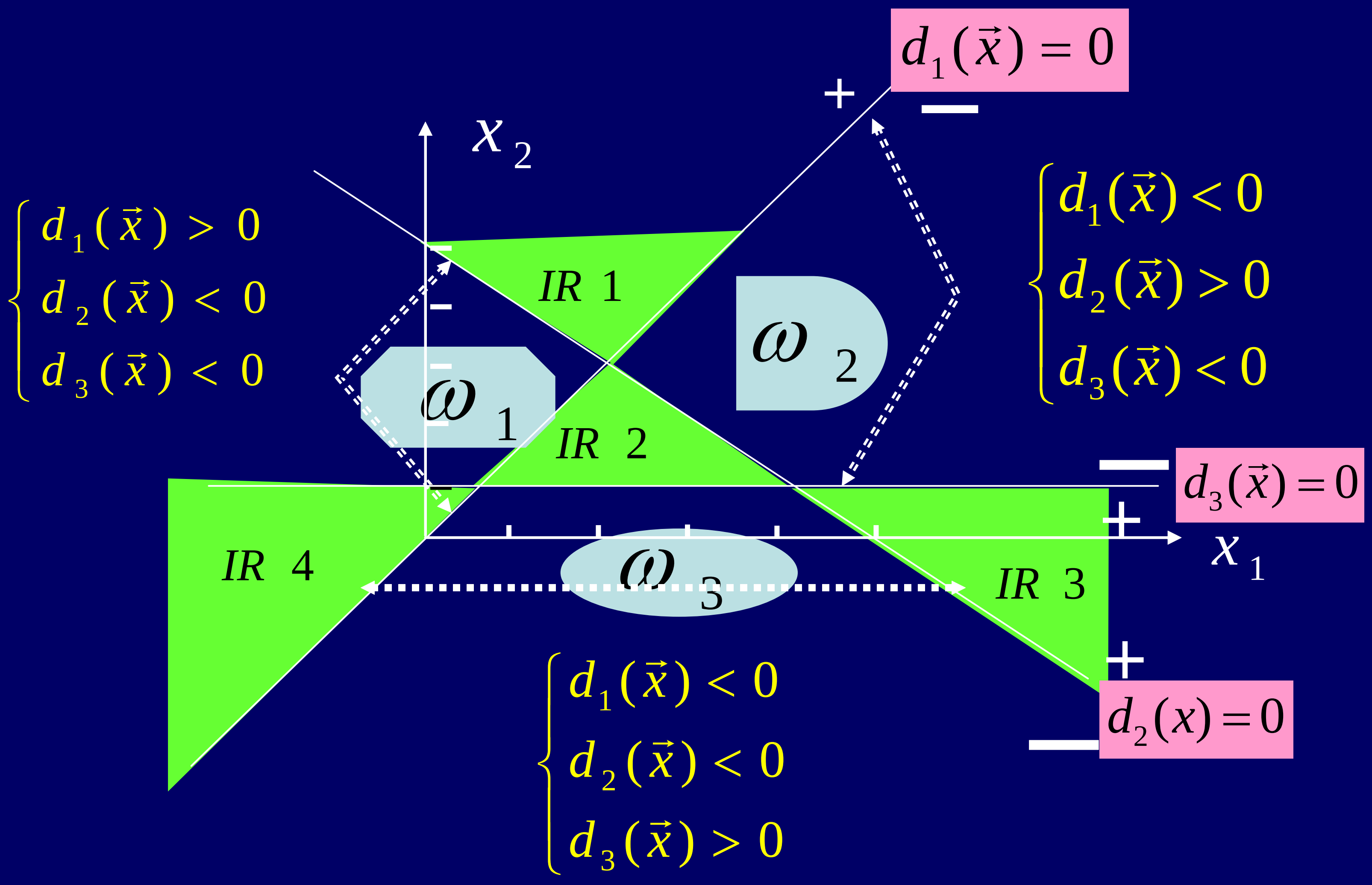
比如对图的三类问题，如果对于任一模式 $\vec{x}$ 如果它的

$$d_1(\vec{x}) > 0$$
$$d_2(\vec{x}) < 0$$
$$d_3(\vec{x}) < 0$$

则该模式属于ω₁类。



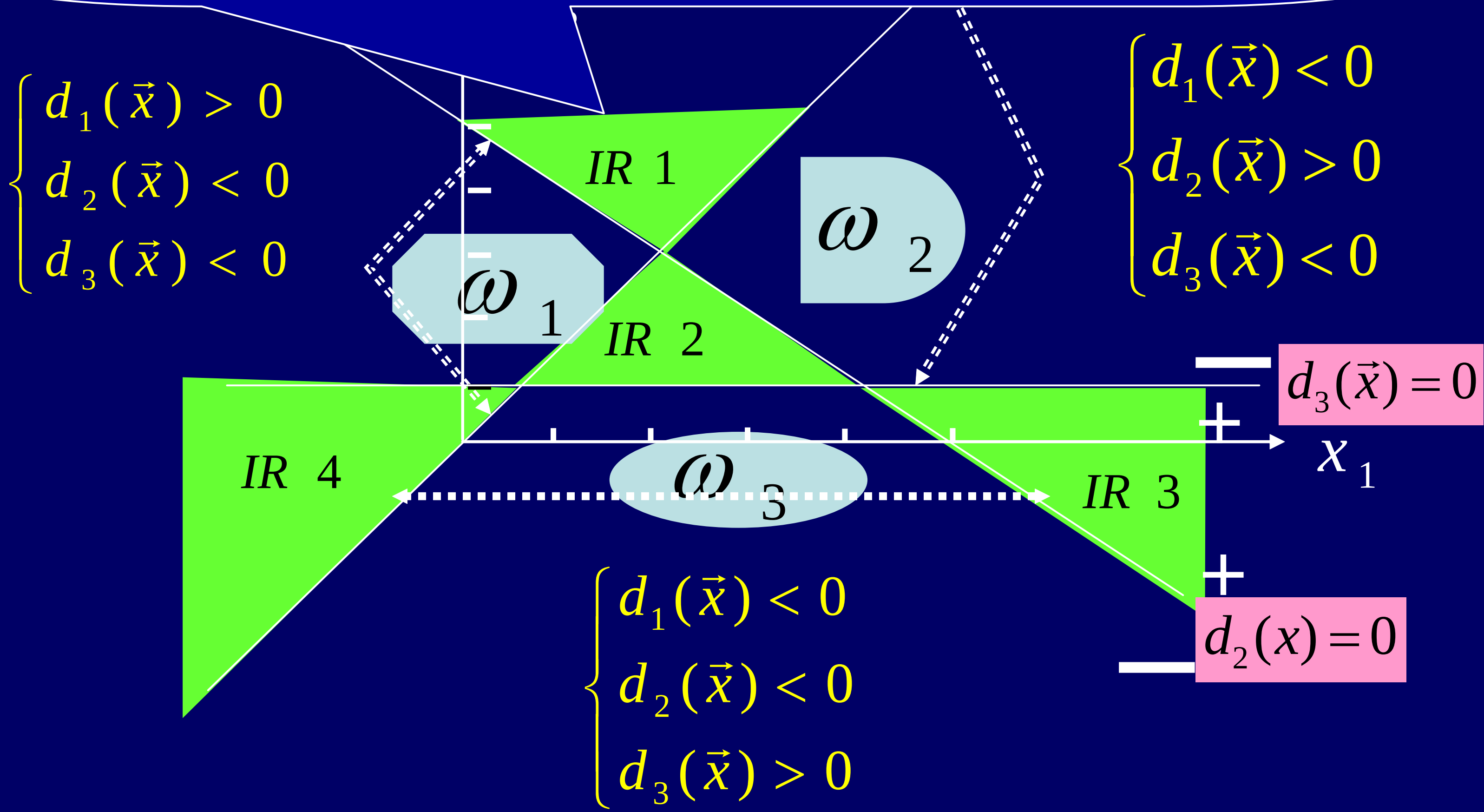
3.2 线性判别函数



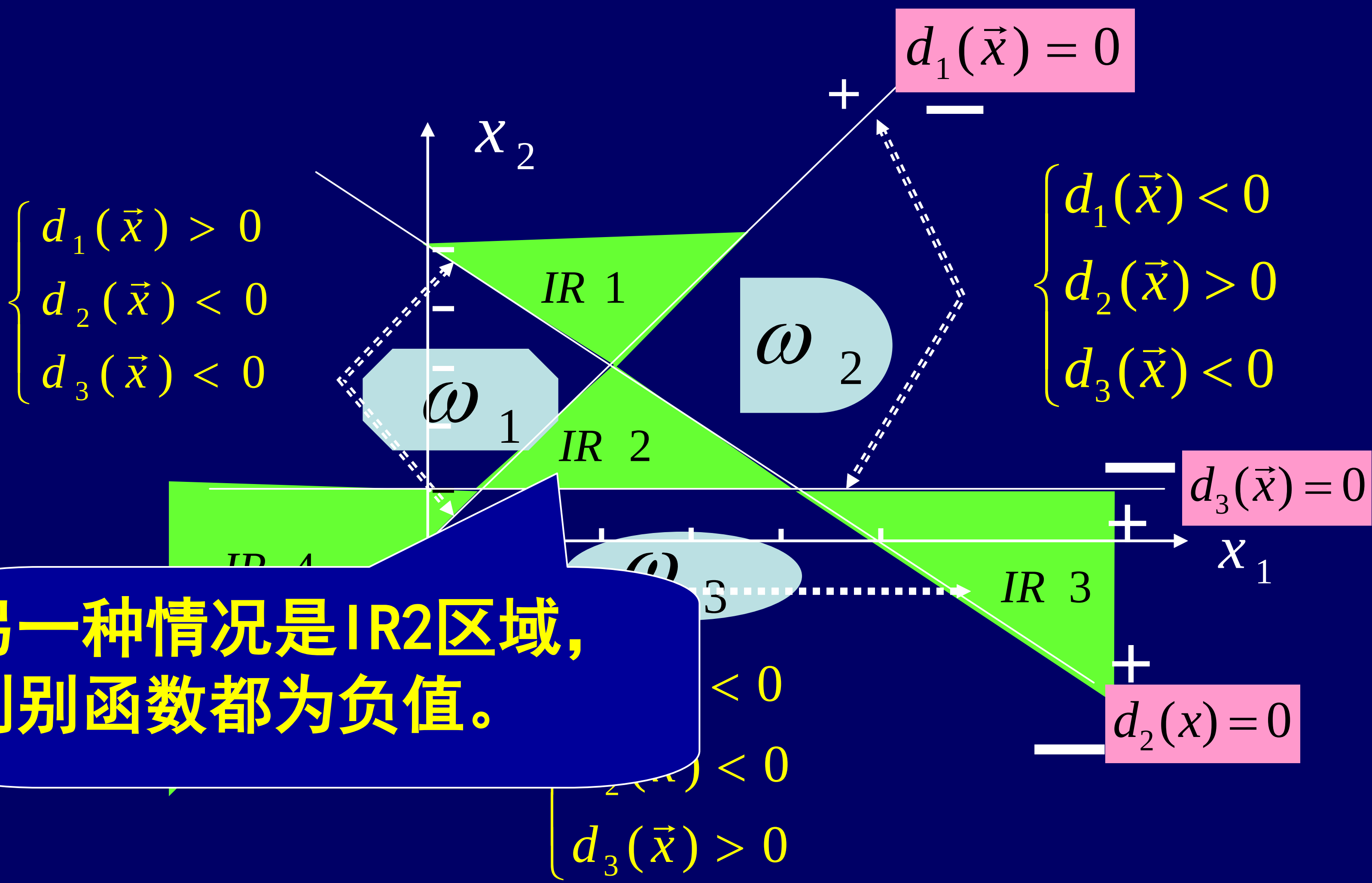
$d_1(\vec{x}) = 0$
$d_3(\vec{x}) = 0$
$d_2(x) = 0$
x_2
x_1
IR 1
IR 2
IR 3
IR 4
ω_1
ω_2
ω_3
$\begin{cases} d_1(\vec{x}) > 0 \\ d_2(\vec{x}) < 0 \\ d_3(\vec{x}) < 0 \end{cases}$
$\begin{cases} d_1(\vec{x}) < 0 \\ d_2(\vec{x}) > 0 \\ d_3(\vec{x}) < 0 \end{cases}$
$\begin{cases} d_1(\vec{x}) < 0 \\ d_2(\vec{x}) < 0 \\ d_3(\vec{x}) > 0 \end{cases}$

3.2　线性判别函数

如果某个x使二个以上的判别函数 $d_i>0$ 。则此模式x就无法作出确切的判决。如图



3.2 线性判别函数



另一种情况是IR2区域，判别函数都为负值。

例：已知三类ω$_1$，ω$_2$，ω$_3$的判别函数分别为：

$$\begin{cases} d_1(\vec{x}) = -x_1 + x_2 \\ d_2(\vec{x}) = x_1 + x_2 - 5 \\ d_3(\vec{x}) = -x_2 + 1 \end{cases}$$

求：当 $\vec{x} = (x_1, x_2)' = (6,5)'$ **时属于哪一类?**

例：已知三类ω$_1$，ω$_2$，ω$_3$的判别函数分别为：

$$\begin{cases} d_1(\vec{x}) = -x_1 + x_2 \\ d_2(\vec{x}) = x_1 + x_2 - 5 \\ d_3(\vec{x}) = -x_2 + 1 \end{cases}$$

求：当 $\vec{x} = (x_1, x_2)' = (6,5)'$ **时属于哪一类?**

解： 三个判别边界分别为：

$$\begin{cases} d_1(\vec{x}) = -x_1 + x_2 = 0 \\ d_2(\vec{x}) = x_1 + x_2 - 5 = 0 \\ d_3(\vec{x}) = -x_2 + 1 = 0 \end{cases}$$

将 $\vec{x}=(x_1,x_2)'=(6,5)'$ 代入方程组:

$$\begin{cases} d_1(\vec{x}) = -x_1 + x_2 \\ d_2(\vec{x}) = x_1 + x_2 - 5 \\ d_3(\vec{x}) = -x_2 + 1 \end{cases}$$

得：

$$d_1(\vec{x}) = -1, d_2(\vec{x}) = 6, d_3(\vec{x}) = -4.$$

结论：　因为

$$d_1(x) < 0, d_2(x) > 0, d_3(x) < 0$$

所以它属于ω_2类。

2. ω_i/ω_j 两分法（第二种情况）

对 c 类中的任意两类 ω_i 和 ω_j 都分别建立一个判别函数，这个判别函数将属于 ω_i 的模式与属于 ω_j 的模式区分开。此函数对其他模式分类不提供信息，因此总共需要 $c(c-1)/2$ 个这样的判别函数。

通过训练得到区分两类 ω_i 和 ω_j 的判别函数为：

$$d_{ij}(\vec{x}) = \vec{w}_{ij}^{'}\vec{x} \qquad (i, j = 1, 2, \cdots, c; i \neq j)$$

2. ω_i / ω_j 两分法（第二种情况）

$$d_{ij}(\vec{x}) = \vec{w}_{ij}^{'}\vec{x}\begin{cases} > 0, 若\vec{x} \in \omega_i \\ < 0, 若\vec{x} \in \omega_j \end{cases}$$

它具有性质： $d_{ij}(\vec{x}) = -d_{ji}(\vec{x})$

2. ω_i / ω_j 两分法（第二种情况）

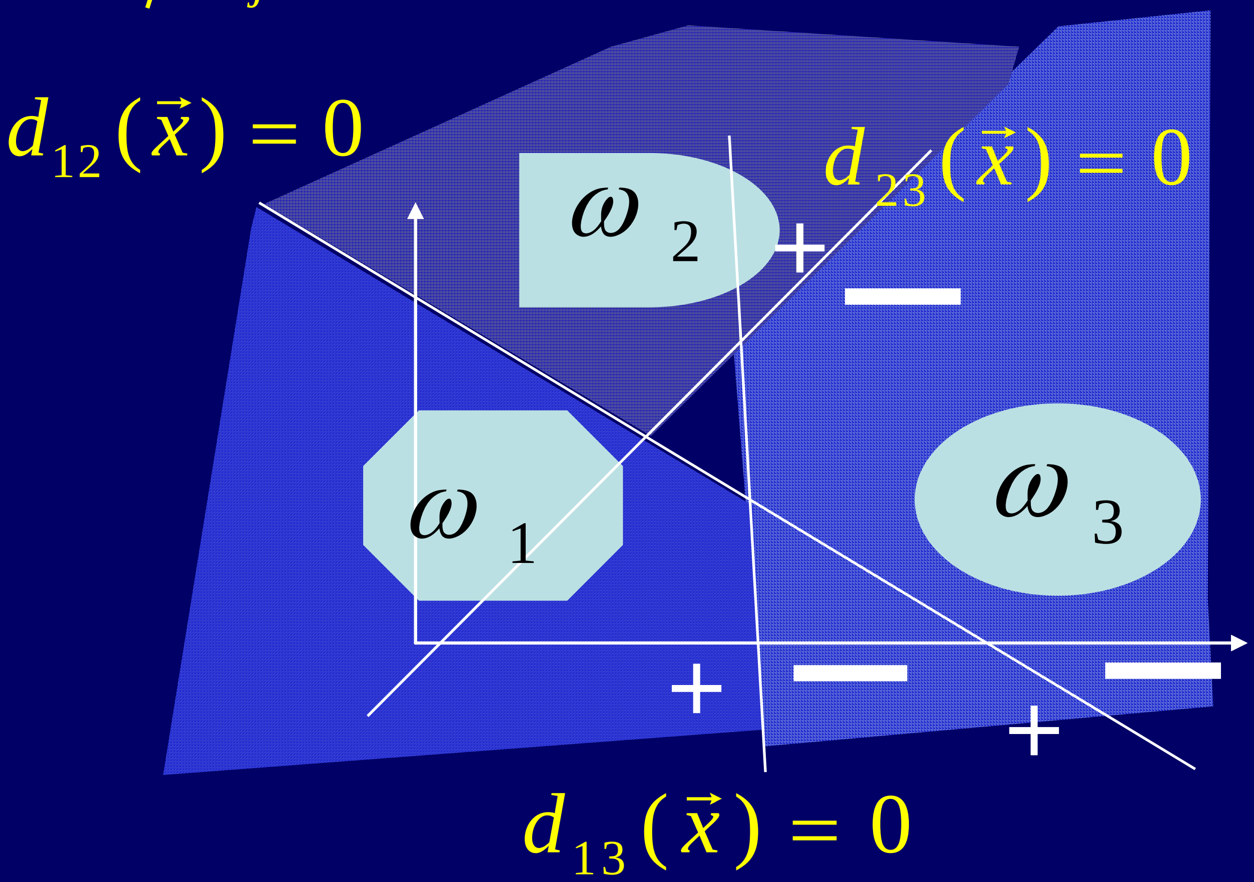
$$d_{ij}(\vec{x}) = \vec{w}_{ij}^{'}\vec{x}\begin{cases} > 0, 若\vec{x} \in \omega_i \\ < 0, 若\vec{x} \in \omega_j \end{cases}$$

它具有性质： $d_{ij}(\vec{x}) = -d_{ji}(\vec{x})$

判别规则是:

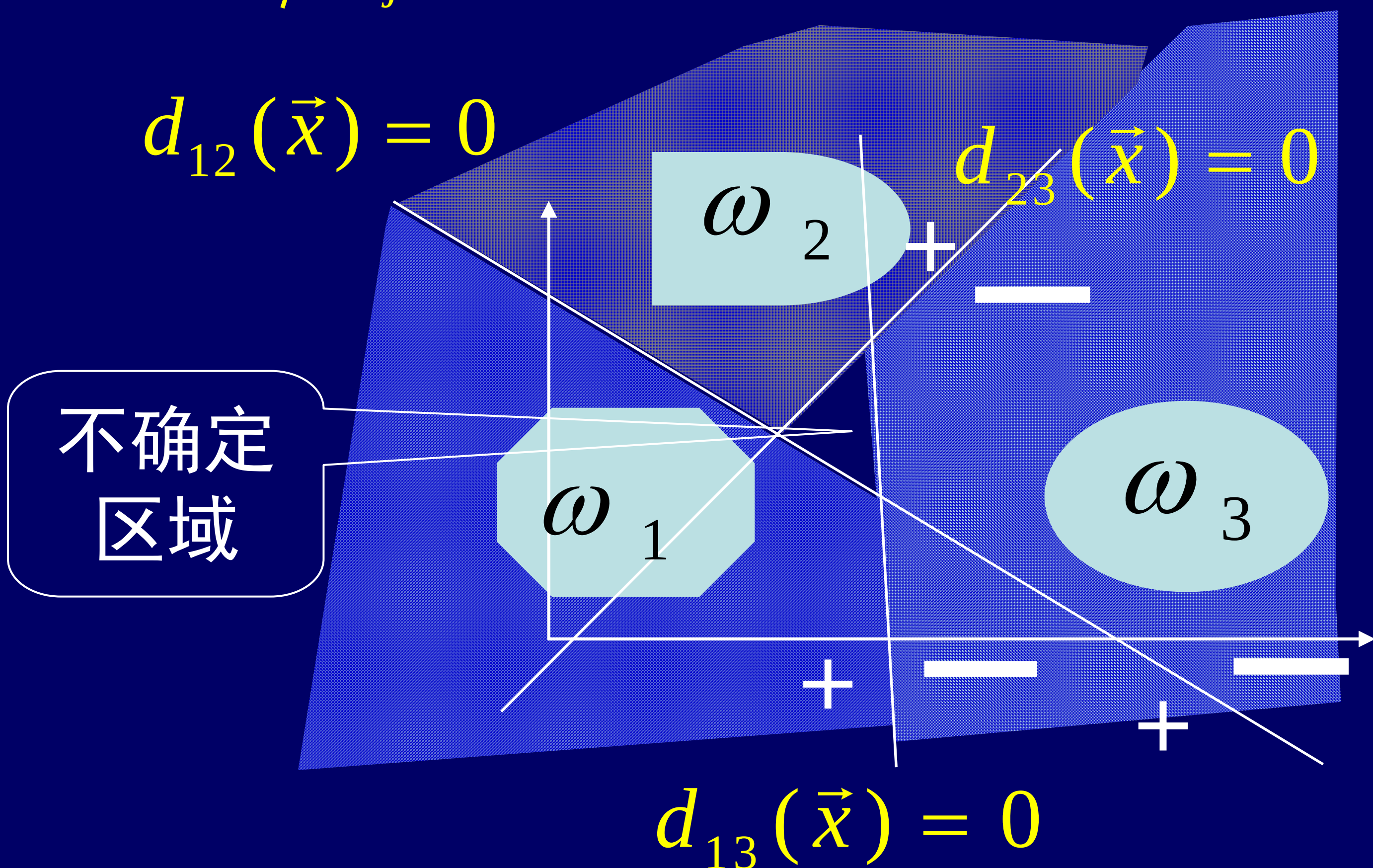
如果： $d_{ij}(x) > 0, \forall j \neq i$ **则判** $x \in \omega_i$

2. ω_i / ω_j **两分法（第二种情况）**



多类问题图例（第二种情况）

2.　ω_i/ω_j 两分法（第二种情况）



多类问题图例（第二种情况）

例：设有一个二维三类问题，三个判别函数为：

$$d_{12}(\vec{x}) = -x_1 - x_2 + 5 \quad d_{13}(\vec{x}) = -x_1 + 3 \quad d_{23}(\vec{x}) = -x_1 + x_2$$

求模式 $\vec{x} = (4,3)'$ **属于哪一类?**

例：设有一个二维三类问题，三个判别函数为：

$d_{12}(\vec{x}) = -x_1 - x_2 + 5 \quad d_{13}(\vec{x}) = -x_1 + 3 \quad d_{23}(\vec{x}) = -x_1 + x_2$

求模式 $\vec{x} = (4,3)'$ **属于哪一类?**

解：将模式 $\vec{x} = (4,3)'$代入$d_{ij}(\vec{x})$有：

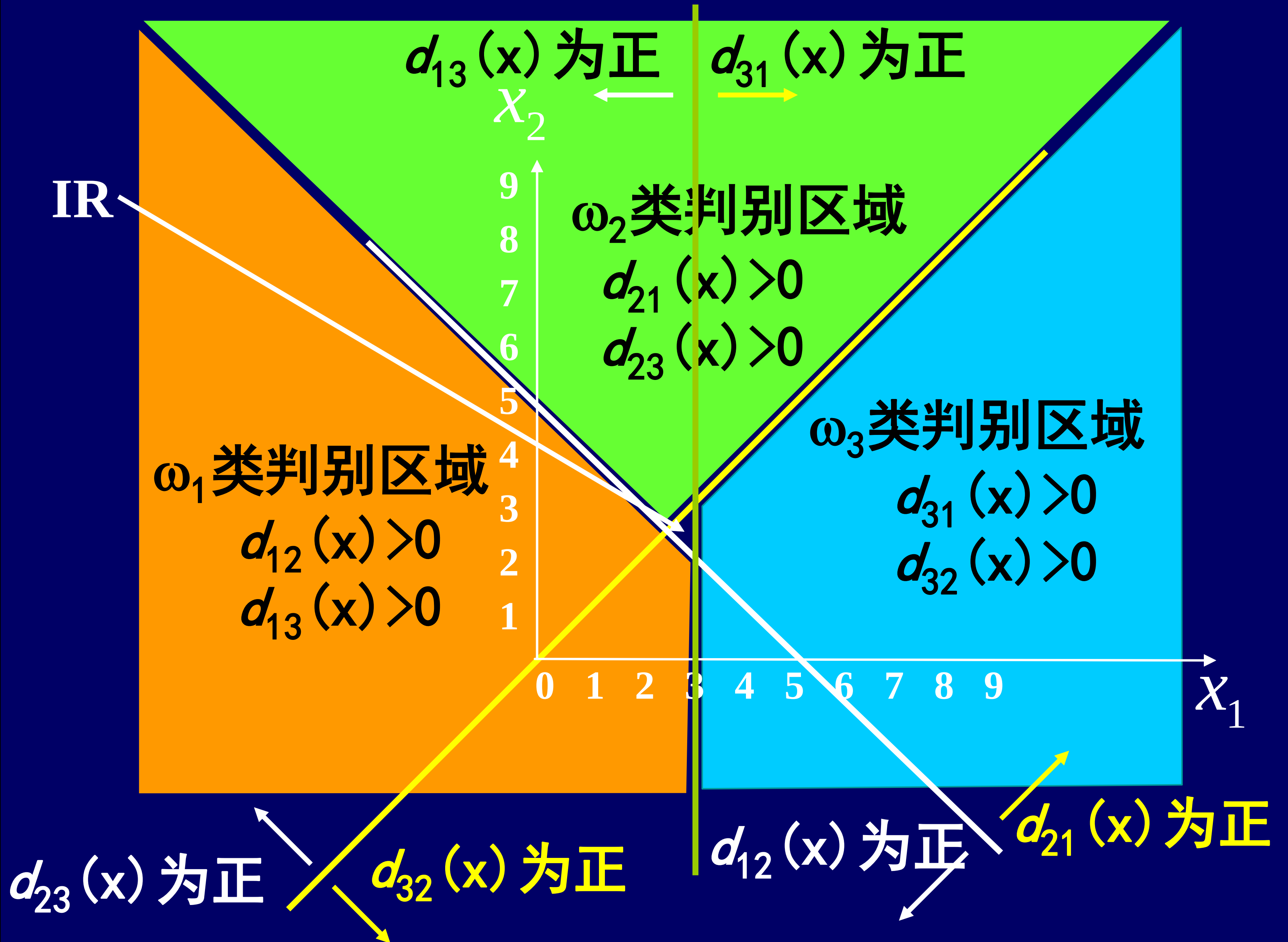
$$d_{12}(\vec{x}) = -2 \Rightarrow \vec{x} \notin \omega_1$$

$$d_{23}(\vec{x}) = -1 \Rightarrow \vec{x} \notin \omega_2$$

$$d_{13}(\vec{x}) = -1 \Rightarrow \vec{x} \notin \omega_1$$

上面三式等效为：$d_{21}(\vec{x}) = 2 \quad d_{31}(\vec{x}) = 1 \quad d_{32}(\vec{x}) = 1$

由于 $d_{3j}(\vec{x}) > 0$ （$j = 1,2$），所以判 $\vec{x} \in \omega_3$



ω_i/ω_j两分法例题图示

3. 没有不确定区的 ω_i/ω_j 两分法（第三种情况）

令方法2中的判别函数为：

$$d_{ij}(\vec{x}) = d_i(\vec{x}) - d_j(\vec{x}) = (\vec{w}_i - \vec{w}_j)'\vec{x}$$

则 $d_{ij}(\vec{x}) > 0$ 等价于 $d_i(\vec{x}) > d_j(\vec{x})$, 于是对每一类 ω_i 均建立一个判别函数 $d_i(\vec{x})$, C 类问题有 C 个判别函数

$$d_i(\vec{x}) = \vec{w}_i'\vec{x} \qquad i = 1, 2, \cdots, c$$

3. 没有不确定区的 ω_i/ω_j 两分法（第三种情况）

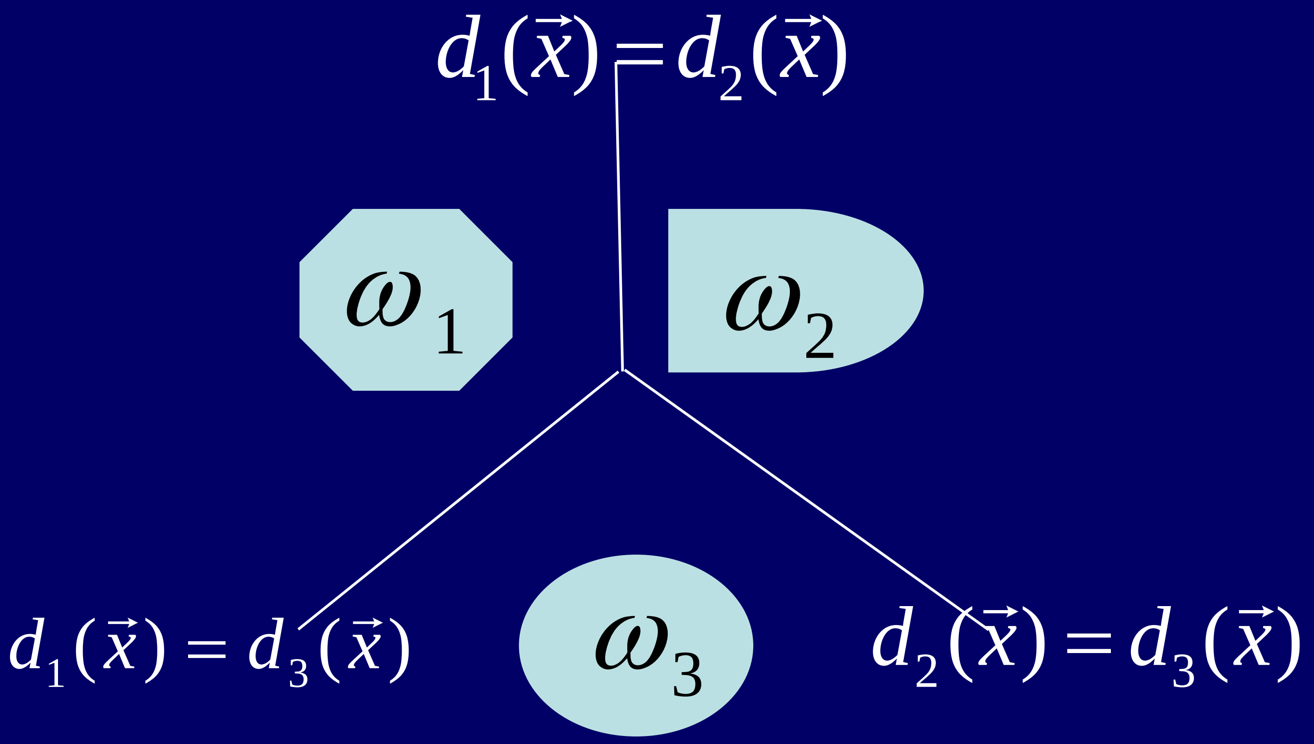
判决规则成为：

如果 $d_i(\vec{x}) > d_j(\vec{x}) \quad \forall j \neq i$ 则判 $\vec{x} \in \omega_i$

判决规则的另一种表达形式

如果 $d_i(\vec{x}) = \max_j \left[d_j(\vec{x}) \right]$ 则判 $\vec{x} \in \omega_i$

3. 没有不确定区的 ω_i/ω_j 两分法（第三种情况）



多类问题图例（第三种情况）

例：设有一个二维三类问题，三个判别函数为：

$$d_1(\vec{x}) = -x_1 + x_2 \quad d_2(\vec{x}) = x_1 + x_2 - 1 \quad d_3(\vec{x}) = -x_2$$

求模式 $\vec{x} = (1,1)'$ **属于哪一类?**

例：设有一个二维三类问题，三个判别函数为：

$$d_1(\vec{x}) = -x_1 + x_2 \quad d_2(\vec{x}) = x_1 + x_2 - 1 \quad d_3(\vec{x}) = -x_2$$

求模式 $\vec{x} = (1,1)'$ **属于哪一类?**

解：将模式 $\vec{x} = (1,1)'$ **代入上面各式得：**

$$d_1(\vec{x}) = 0 \qquad d_2(\vec{x}) = 1 \qquad d_3(\vec{x}) = -1$$

由于 $\begin{cases} d_2(\vec{x}) > d_3(\vec{x}) \\ d_2(\vec{x}) > d_1(\vec{x}) \end{cases}$ **所以** $\vec{x} \in \omega_2$

上述三种方法小结:

当 $c>3$ 时，ω_i/ω_j 法比 $\omega_i/\overline{\omega}_i$ 法需要更多的判别函数式，这是一个缺点。

但是 $\omega_i/\overline{\omega}_i$ 法是将 ω_i 类与其余的 $c-1$ 类区分开，而 ω_i/ω_j 法是将 ω_i 类和 ω_j 类分开，显然 ω_i/ω_j 法使模式更容易线性可分，这是它的优点。

方法(3)判别函数的数目和方法(1)相同，但没有不确定区，分析简单，是最常用的一种方法。

谢谢！