第二讲

# 模式识别

(Pattern Recognition)


武汉大学计算机学院

Email: zhanglefei@whu.edu.cn

# 第2章 线性判别分析

2.1 判别函数
2.2 线性判别函数
2.3 线性判别函数的性质
2.4 广义线性判别函数
2.5 线性分类器设计
2.6 分段线性分类器设计(*:不要求)

- 统计模式识别在模式识别学科方向中研究历史最长，是模式识别的主要理论。

  统计模式识别分为判别函数法和聚类分析法两大类。判别函数法属于监督分类(Supervised Classification)，聚类分析法属于非监督分类(Unsupervised Classification)。

  判别函数法需要有足够的先验知识，首先利用已知类别的训练样本集确定出判别函数，然后再利用训练好的判别函数对未知的模式进行识别分类。

  在判别函数法中，又可分为线性判别函数法、非线性判别函数法和统计决策方法等，其中线性判别函数法和非线性判别函数法是几何分类法，用于研究确定性事件的识别分类；统计决策方法是概率分类法，用于研究随机事件的识别分类。

- 统计模式识别
  - 判别函数法
    - 线性判别函数法
    - 非线性判别函数法
    - 统计决策方法 ⟹ 概率分类法
  - 聚类分析法

线性判别函数法、非线性判别函数法 } 几何分类法

几何分类法: 研究确定性事件的分类

概率分类法: 研究随机事件的分类

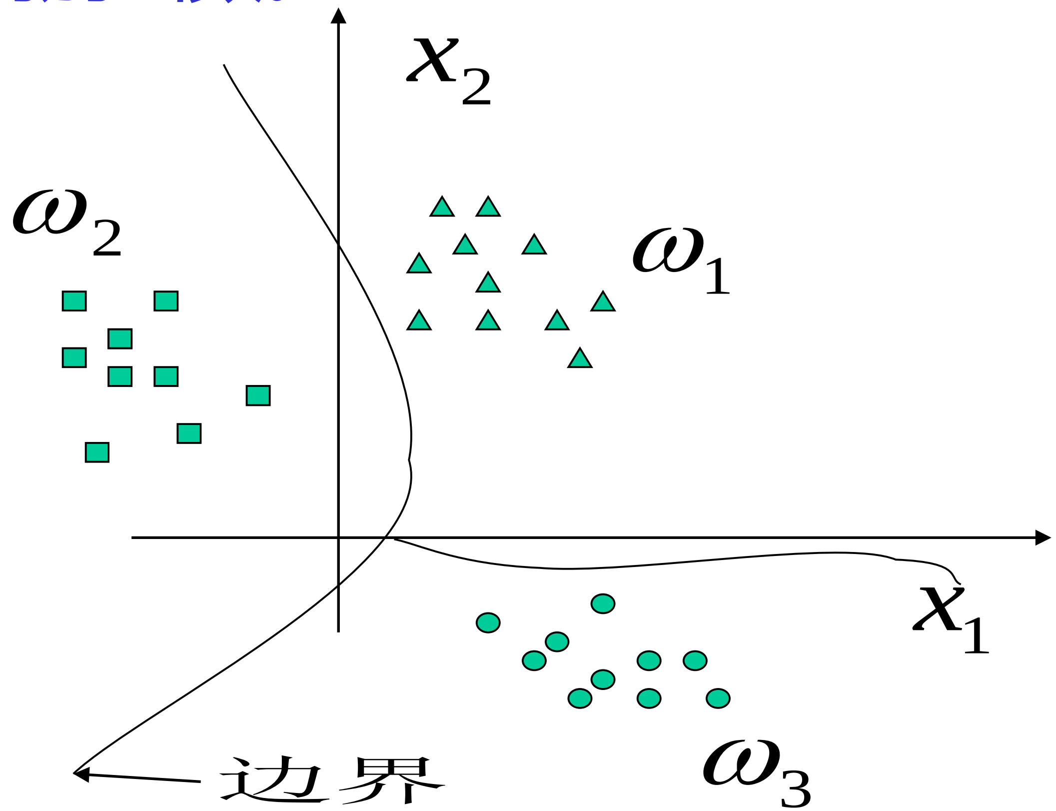
图: 统计模式识别方法的分类

- 在科研、生产和日常生活中，经常会遇到各种各样的判别问题，即要根据某个体的特征的观察值，判断它属于几个类别中的哪一个类。医生根据病人的指标（体温、血压、白细胞数目等）来判断病人生何种病等就属此类。其实，在生物学分类、医疗诊断、地质找矿、石油钻探、天气预报等诸多领域，判别分析方法已经成为一种有效的统计推断方法。

- 简单地说，所谓判别分析问题，就是指在已有给定的若干个模式类（若干个总体）的观察资料的基础上，构造出一个或多个判别函数，并能由此函数对未知其所属类别的新的模式作出判断，决定其应属于哪个模式类。

# 2.1 判别函数

如下图：三类的分类问题，它们的边界线就是一个判别函数。



# 2.1　判别函数

假设对一模式X已提取n个特征，表示为:

$$X=(x_1,x_2,x_3,\ldots,x_n)^T$$

$X$是n维空间的一个向量

模式识别问题就是根据模式***X***的**n**个特征(指标)来判别模式属于ω1 ,ω2 , … , ωm 类中的哪一类。

# 2.1 判别函数

- 判别函数包含两类：

  一类是线性判别函数：

  - 线性判别函数
  - 广义线性判别函数
    - （所谓广义线性判别函数就是把非线性判别函数映射到另外一个空间变成线性判别函数）
  - 分段线性判别函数

  另一类是非线性判别函数

# 2.2 线性判别函数

分别对两类问题和多类问题进行讨论。

➢(一) 两类问题 即：$\omega_i = (\omega_1, \omega_2)^T, M = 2$

1. 二维情况：取两个特征向量

$$X = (x_1, x_2)^T, n = 2$$

这种情况下判别函数：

$$g(x) = w_1 x_1 + w_2 x_2 + w_3$$

$w$为参数，$x_1, x_2$为坐标向量

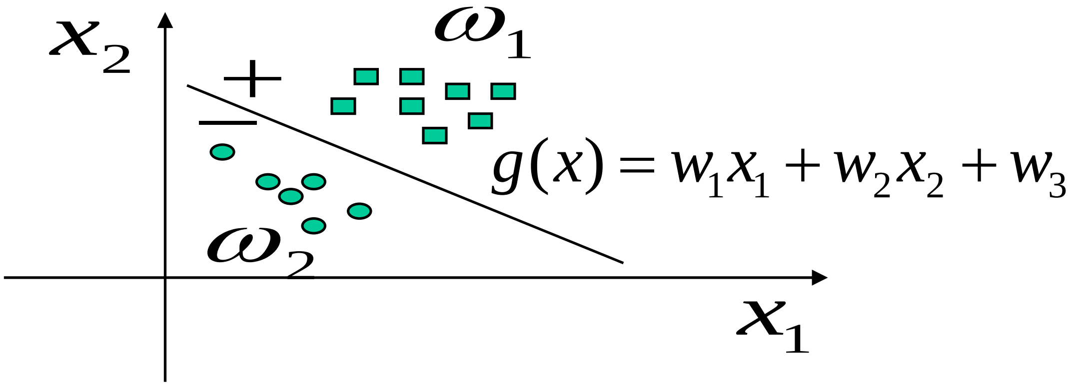
# 2.2 线性判别函数

- 1. 二维情况

在两类别情况下，判别函数 $g(x)$ 具有以下性质：

$$g(x)=\begin{cases}>0, X\in\omega_1\\<0, X\in\omega_2\end{cases}\qquad g(x)=0, X\text{不定}$$

这是二维情况下判别由判别边界分类。情况如图：



# 2.2 线性判别函数

➢2. n维情况

现提取n个特征为： $X=(x_1,x_2,x_3,\ldots x_n)^T$

判别函数：$g(x)=w_1x_1+w_2x_2+\ldots\ldots+w_nx_n+w_{n+1}$

$W_0=(w_1,w_2,\ldots,w_n)^T$ 为权向量，

$X=(x_1,x_2,\ldots,x_n)^T$ 为模式向量。

另外一种表示方法：$g(x)=W^TX$

$W=(w_1,w_2,\ldots,w_n,w_{n+1})^T$ 为增值权向量，

$X=(x_1,x_2,\ldots,x_n,1)^T$ 为增值模式向量。

# 2.2 线性判别函数

***回顾：多元线性回归(Multivariate linear regression)**

给定数据集$D=\{(\mathbf{x}_1,y_1),(\mathbf{x}_2,y_2),\ldots,(\mathbf{x}_m,y_m)\}$,

其中:

$\mathbf{x}_i=(x_{i1},x_{i2},\ldots,x_{id},1)^T\ (i=1,2,\ldots,m)$,

即给定$m$个模式, 各模式样本有$d$个特征, $y_i\in R$.

$\mathbf{w}=(w_1,w_2,\ldots,w_d,w_{d+1})^T$

我们试图从数据集$D=\{(\mathbf{x}_i,y_i)\}_{i=1}^m$学得

$g(\mathbf{x}_i)=\mathbf{w}^T\mathbf{x}_i$, 使得$g(\mathbf{x}_i)\approx y_i$的$\mathbf{w}$.

**w**参数学习/训练出来之后，所确定的g(**x**)模型也叫预测模型。

若用g(**x**)预测的是离散值，此类学习任务称为“**分类**(Classification)”；若用g(**x**)预测的是连续值，此类学习任务称为“**回归**(regression)”。由于这里g(**x**)是线性的，故称为多元线性回归(本科生概率统计课程中，学过多元正态线性回归)。

# 2.2 线性判别函数

## ➢2. n维情况

❖模式分类：

$$g(x)=W^T X \begin{cases} >0, X \in \omega_1 \\ <0, X \in \omega_2 \end{cases}$$

❖$g(x)=W^TX=0$为判别边界。当$n$=2时，二维情况的判别边界为一直线；当$n$=3时，判别边界为一平面；当$n$>3时，则判别边界为一超平面。

# 2.2 线性判别函数

## (二) 多类问题

对于多类问题，模式有 ω1 ,ω2 , … , ωm 个类别。可分三种情况：

1.第一种情况：每一模式类与其他模式类间可用单个判别平面把一个类分开。这种情况，M类可有M个判别函数，且具有以下性质：

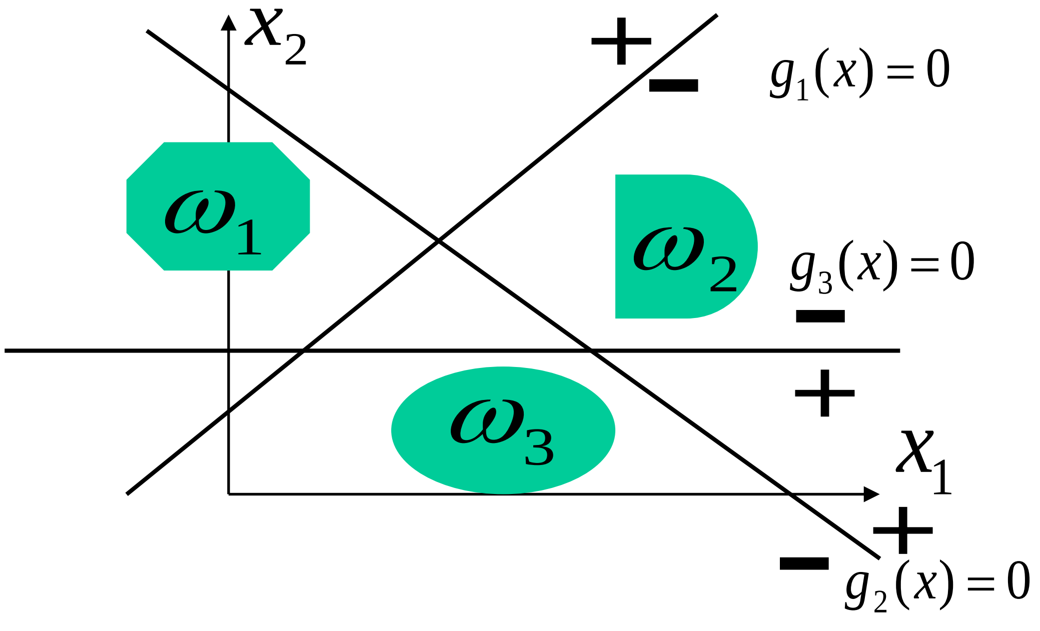
$$g_i(x) = W_i^T X \begin{cases} > 0, X \in \omega_i \\ < 0, \text{其他}, i = 1, 2, \dots, M。 \end{cases}$$

式中 $W_i = (w_{i1}, w_{i2}, \dots, w_{in}, w_{in+1})^T$ 为第 $i$ 个判别函数的权向量。

# 1. 第一种情况

- 下图所示，每一类别可用单个判别边界与其他类别相分开 。

若一模式$X$属于$\omega_1$，则由图可清楚看出：这时$g_1(x)>0$而$g_2(x)<0$， $g_3(x)<0$ 。 $\omega_1$ 类与其他类之间的边界由$g_1(x)=0$确定。



# 1. 第一种情况(Cont.)

- 例：已知三类ω1,ω2,ω3的判别函数分别为

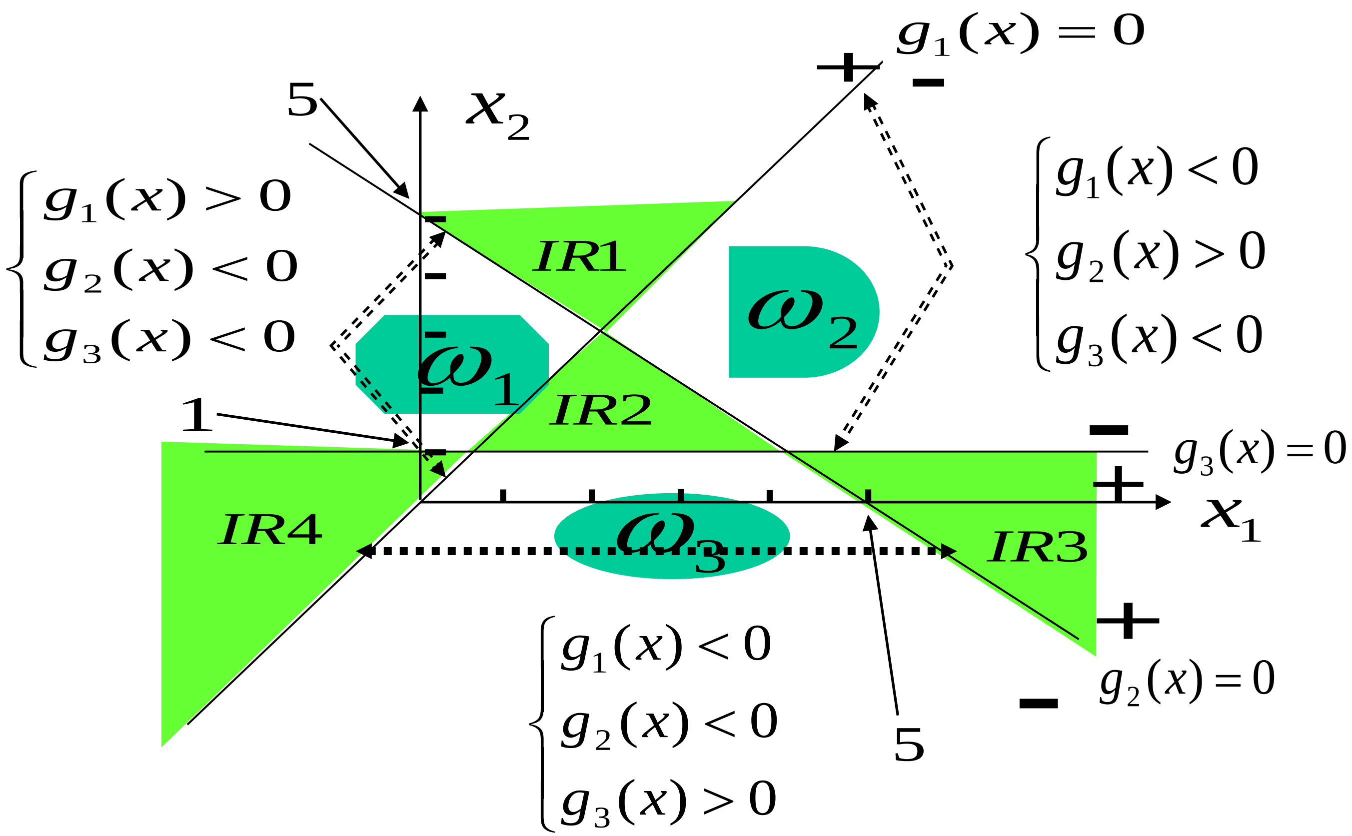
$$\begin{cases} g_1(x) = -x_1 + x_2 \\ g_2(x) = x_1 + x_2 - 5 \\ g_3(x) = -x_2 + 1 \end{cases}$$

则，三个判别边界为

$$\begin{cases} g_1(x) = -x_1 + x_2 = 0 \\ g_2(x) = x_1 + x_2 - 5 = 0 \\ g_3(x) = -x_2 + 1 = 0 \end{cases}$$

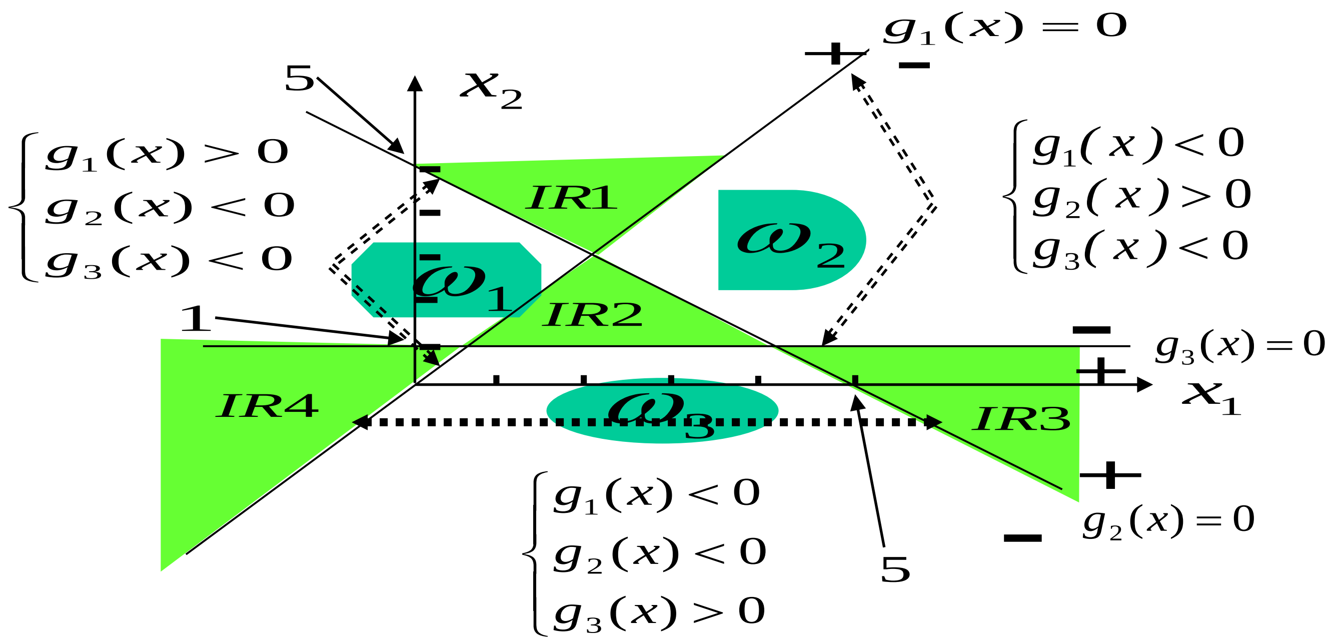
# 1. 第一种情况(Cont.)

作图如下：



# 1. 第一种情况(Cont.)

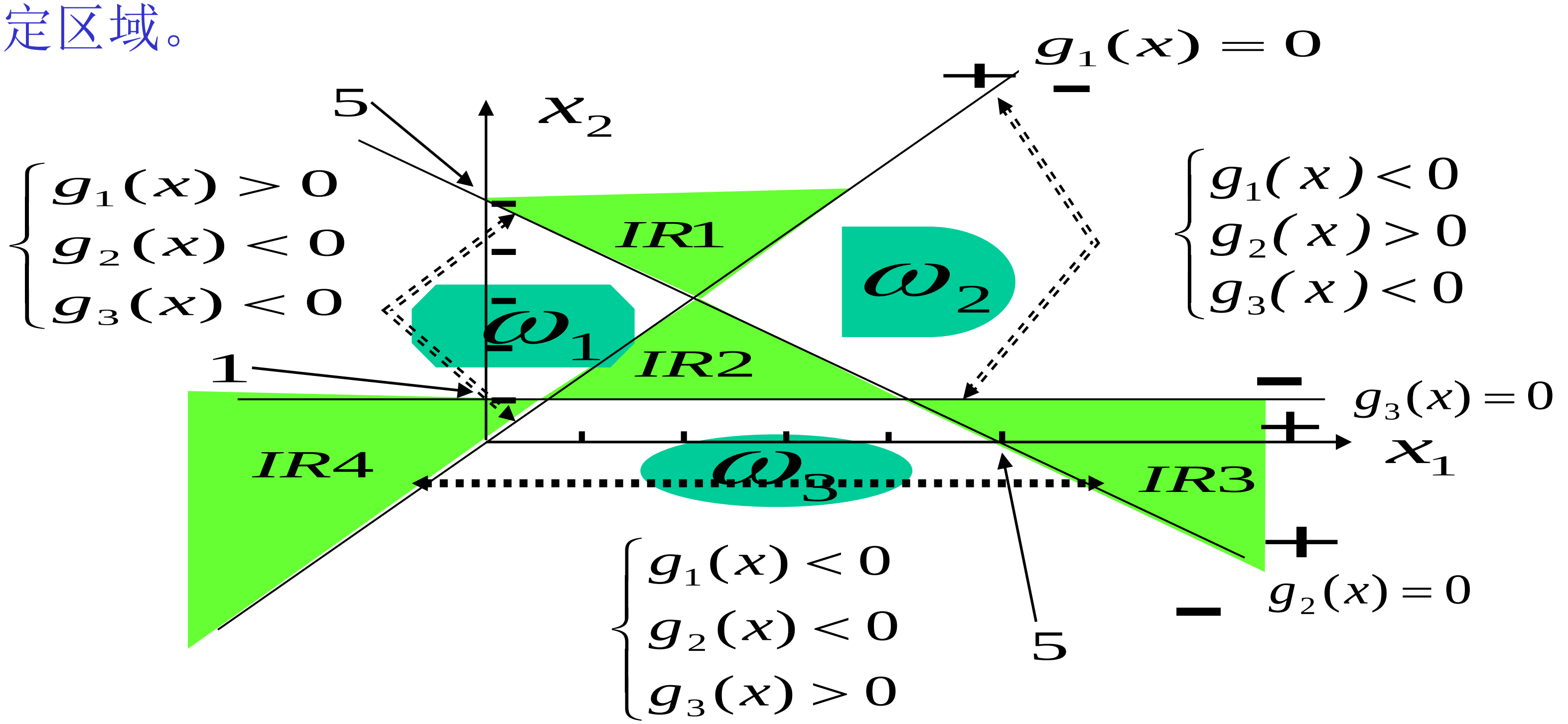
- 对于任一模式**X**，若它的 $g_1(x)>0$， $g_2(x)<0$， $g_3(x)<0$
- 则该模式属于$\omega_1$类。相应$\boldsymbol{\omega_1}$类的区域由直线$-x_2+1=0$的正边、直线$-x_1+x_2-5=0$ 和直线$-x_1+x_2=0$的负边来确定。



# ➢1. 第一种情况(Cont.)

❖必须指出，如果某个***X***使二个以上的判别函数 $g_i(x)>0$ ，则此模式***X***就无法作出准确的判决。如图中 **IR1,IR3**，**IR4**区域。

❖另一种情况是**IR2**区域，判别函数都为负值。**IR1**，**IR2**，**IR3**，**IR4**都为不确 定区域。



## 1. 第一种情况(Cont.)

- 问当$x=(x_1,x_2)^T=(6,5)^T$时属于那一类

代入判别函数方程组：

$$\begin{cases} g_1(x) = -x_1 + x_2 \\ g_2(x) = x_1 + x_2 - 5 \\ g_3(x) = -x_2 + 1 \end{cases}$$

得：

$$g_1(x) = -1, g_2(x) = 6, g_3(x) = -4.$$

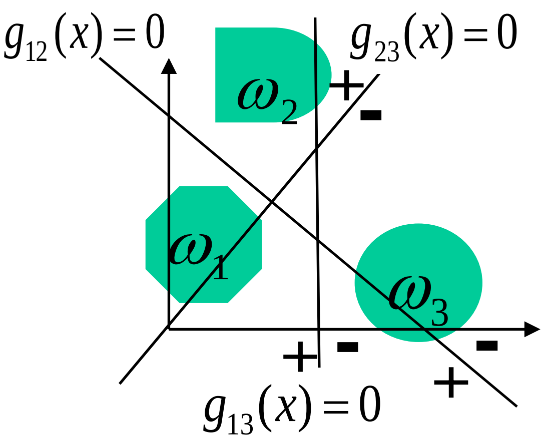
- 结论：$g_1(x) < 0$，$g_2(x) > 0$，$g_3(x) < 0$所以它属于$\omega_2$类

# 2. 第二种情况

- 每两个模式类间可用判别平面分开。

- 这样有 $C_M^2 = M(M-1)/2$ 个判别平面。
- 对于两类问题，$M$=2，则有一个判别平面。
- 同理，三类问题则有三个判别平面。

- 判别函数：$g_{ij}(x) = W_{ij}^T X$
- 判别边界：$g_{ij}(x) = o$




- 判别条件：$g_{ij}(x)\begin{cases} >0 \rightarrow 当X\in\omega_i \\ <0 \rightarrow 当X\in\omega_j \end{cases} \quad i \neq j$

- 判别函数性质：$g_{ij}(x) = -g_{ji}(x)$

# 2. 第二种情况(Cont.)

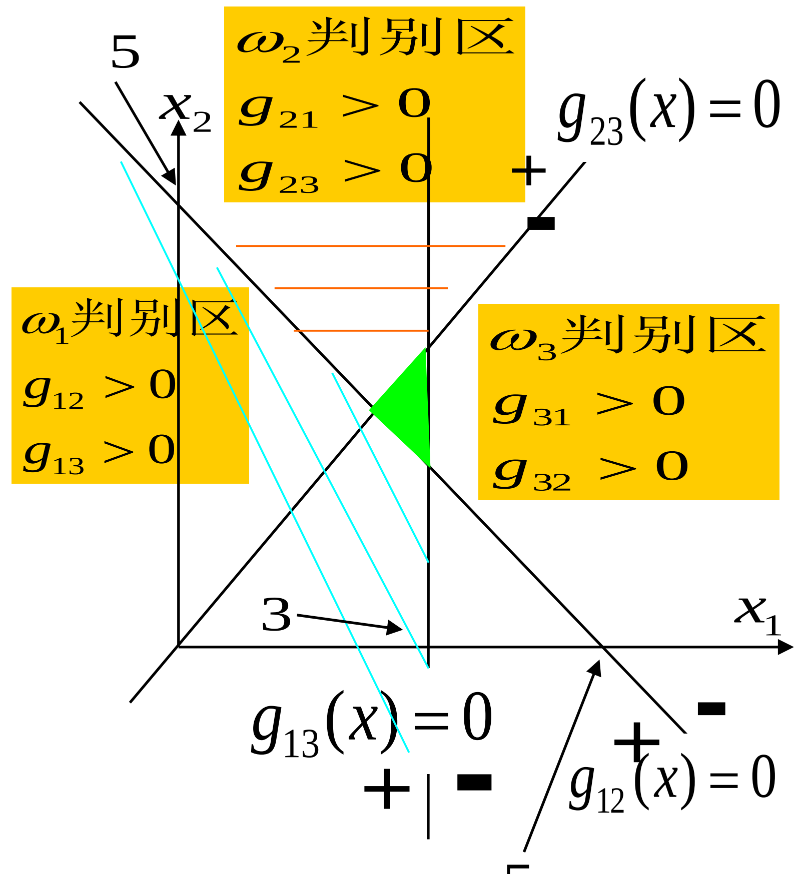
- 假设判别函数为：

$$\begin{cases} g_{12}(x) = -x_1 - x_2 + 5 \\ g_{13}(x) = -x_1 + 3 \\ g_{23}(x) = -x_1 + x_2 \end{cases}$$

- 判别边界为：

$$\begin{cases} g_{12}(x) = -x_1 - x_2 + 5 = 0 \\ g_{13}(x) = -x_1 + 3 = 0 \\ g_{23}(x) = -x_1 + x_2 = 0 \end{cases}$$

- 用上面方程式作出右图：



# ➢2. 第二种情况(Cont.)

❖结论：判别区间增大，不确定区间减小，比第一种情况小的多。

❖问:未知模式X=(x1, x2)$^T$=(4, 3)$^T$属于哪一类?

代入判别函数可得:

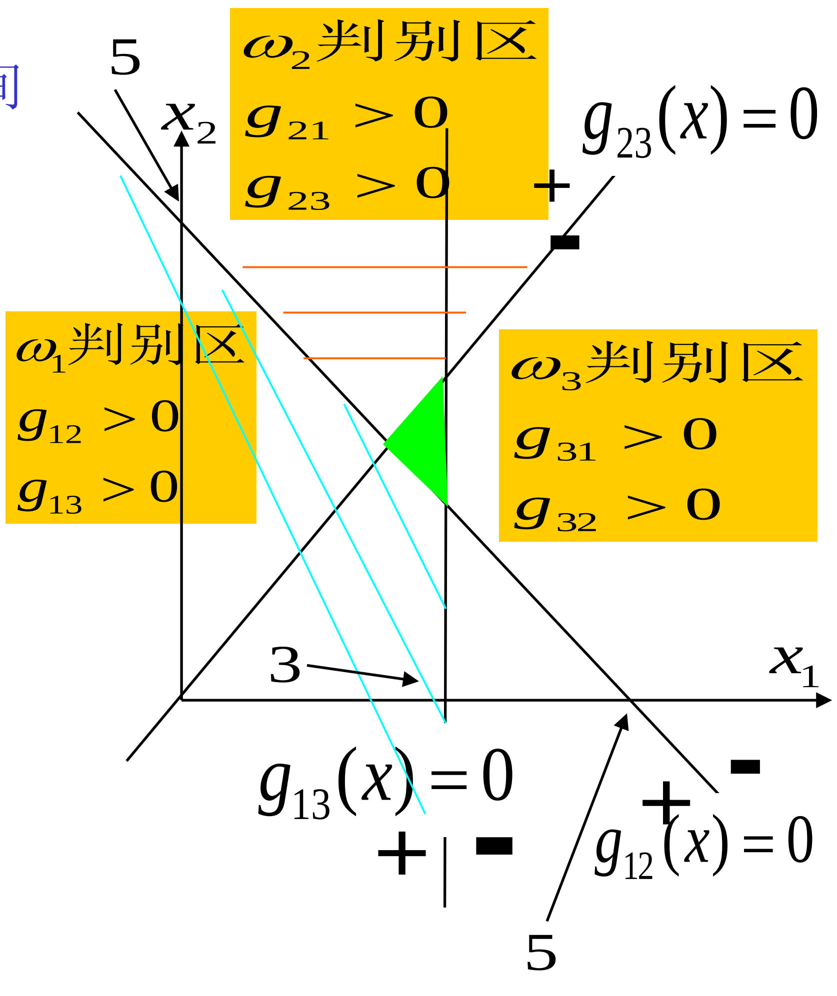
$$g_{12}(x) = -2, g_{13}(x) = -1, g_{23}(x) = -1$$

把下标对换取负，可得:

$$g_{21}(x) = 2, g_{31}(x) = 1, g_{32}(x) = 1$$

因为$g_{3j}(x) > 0$

所以$X$ 属于 $\omega_3$类



## 3. 第三种情况

- 每类都有一个判别函数,存在$M$个判别函数
- 判别函数: $g_i(x)=W_i X \qquad i=1,2,\ldots,M$
- 
- 判别规则: $g_i(x)=W_i^T X \begin{cases} \text{最大,当} X\in\omega_i \\ \text{小,其他} \end{cases}$
- 判别边界: $g_i(x)=g_j(x)$ 或$g_i(x)-g_j(x)=0$
- 就是说,要判别模式$X$属于哪一类,先把$\boldsymbol{X}$代入**M**个判别函数中,判别函数最大的那个类别就是$\boldsymbol{X}$所属类别。类与类之间的边界可由$g_i(x)=g_j(x)$ 或$g_i(x)-g_j(x)=0$来确定。
- 与第一种情况的区别:此情况可有多个判别函数值大于0(而第一种情况只有一个判别函数的值大于0)

## ➢3. 第三种情况(Cont.)

❖下图所示是$M$=3 的例子。

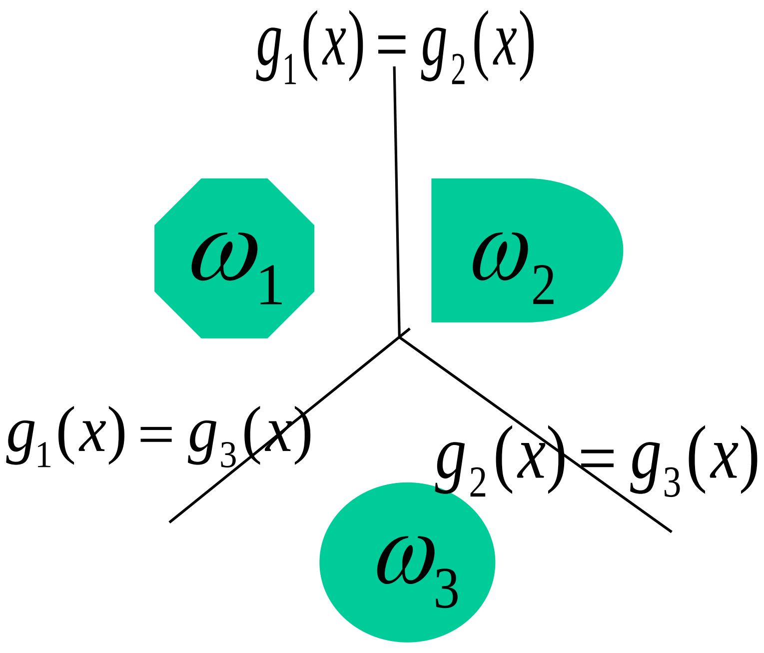
对于ω1类模式，必然满足$g_1(x)>g_2(x)$ 和 $g_1(x)>g_3(x)$ 。

假设判别函数为：

$$\begin{cases} g_1(x) = -x_1 + x_2 \\ g_2(x) = x_1 + x_2 - 1 \\ g_3(x) = -x_2 \end{cases}$$

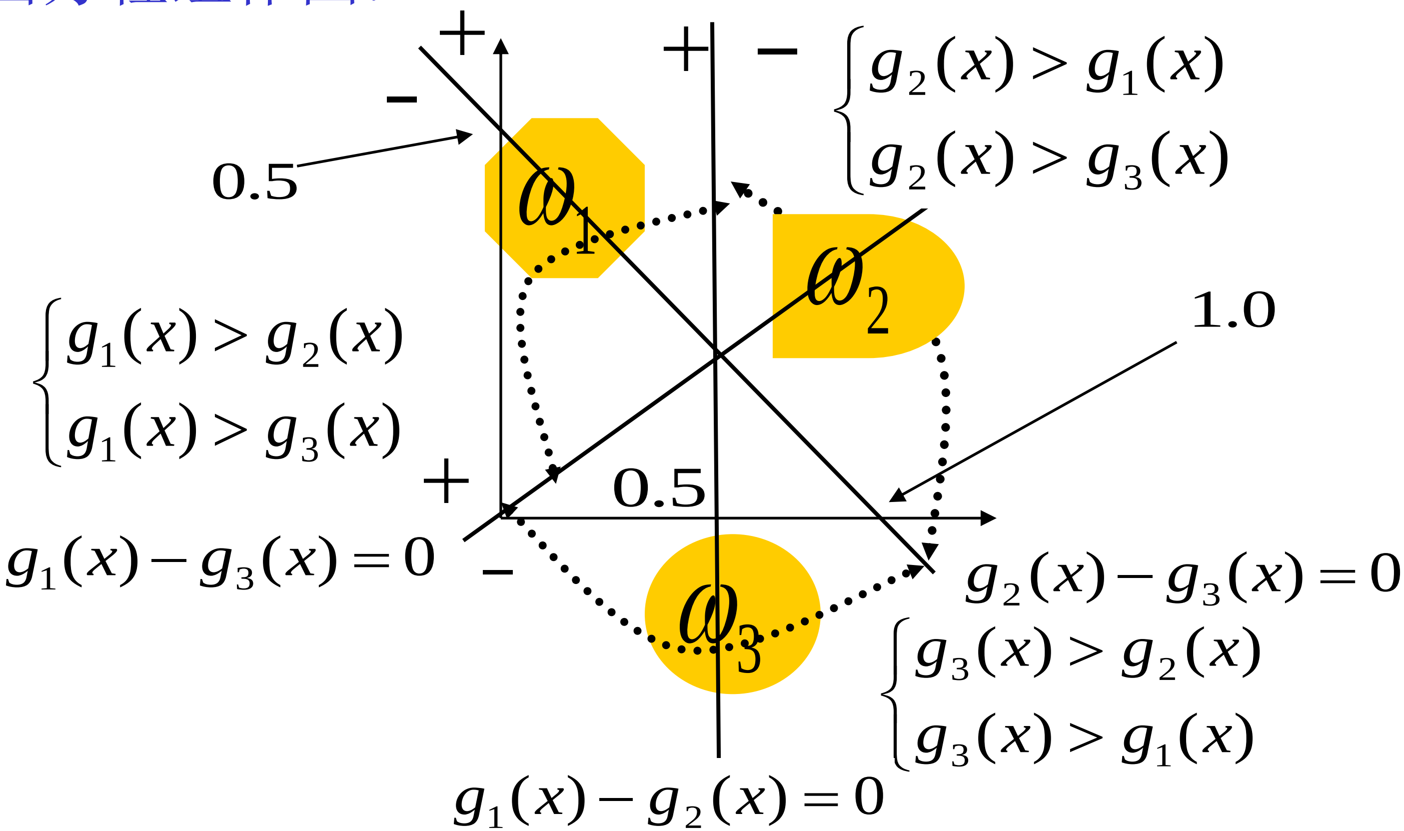
则判别边界为：

$$\begin{cases} g_1(x) - g_2(x) = -2x_1 + 1 = 0 \\ g_1(x) - g_3(x) = -x_1 + 2x_2 = 0 \\ g_2(x) - g_3(x) = x_1 + 2x_2 - 1 = 0 \end{cases}$$



## ➢3. 第三种情况(Cont.)

❖用上面方程组作图：




❖ 结论：不确定区间没有了，所以这种是最好的情况。

## 3. 第三种情况(Cont.)

- 问：假设未知模式$x=(x_1,x_2)^T=(1,1)^T$，则$x$属于哪一类。

  把$x$代入判别函数：$g_1(x), g_2(x), g_3(x).$

  得判别函数为：$g_1(x)=0, g_2(x)=1, g_3(x)=-1$

  因为 $g_2(x)>g_3(x), g_2(x)>g_1(x)$

  所以模式$x=(1,1)^T$属于$\omega_2$类。

# 2.3　线性判别函数的性质

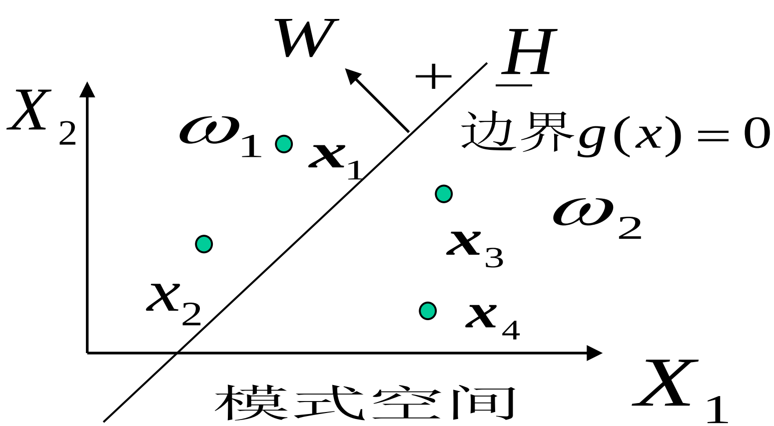
➢1. 模式空间与加权空间　　$g(x)=W_i^T X$

.模式空间：由 $X=(x_1,x_2,x_3,\ldots,x_n)^T$ 构成的n维欧氏空间。

.W是此空间的加权向量，它决定模式的分界面H，W与H正交。

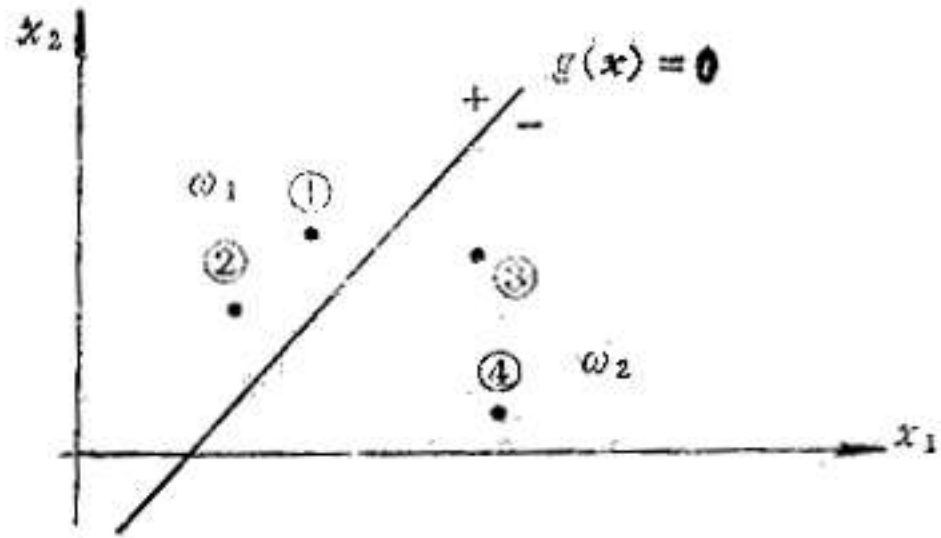
.加权空间：以 $w_1,w_2,\ldots,w_{n+1}$ 为变量构成的欧氏空间。

.模式空间与加权空间的几何表示如下图：



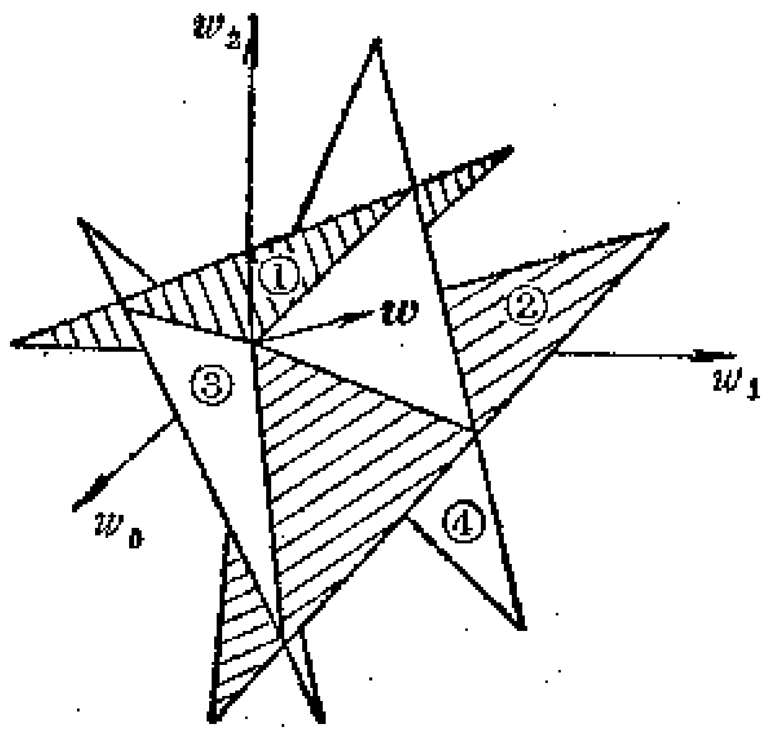

由于假设权向量W与模式向量X的内积为零(g(x)=0)，故W与分界面H正交

# 模式空间



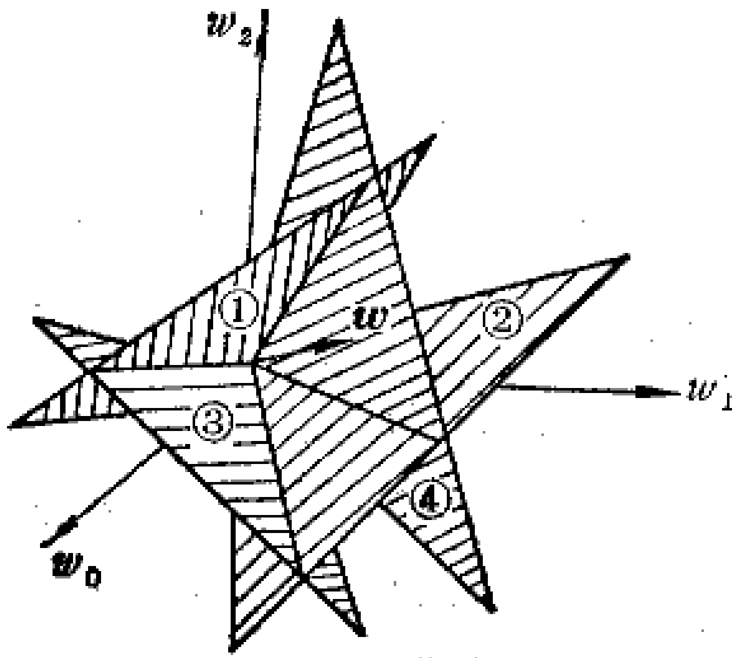

(a)

# 加权空间判别界面




(b)

正规化后的加权空间判别界面




(c)

# 1. 模式空间与加权空间(Cont.)

- 加权空间的构造：$g(x) = w_1 x_1 + w_2 x_2 + w_3$
- 设 $x_1 = (x_{11}, x_{12})^T$ 是加权空间分界面上的一点，代入上式得：
  $g(x_1) = w_1 x_{11} + w_2 x_{12} + w_3 = 0$，这是加权空间的边界。
- 该式表示一个通过加权空间原点的平面，此平面就是加权空间图中的平面①,同样令$g(x_2)$ =$g(x_3)$ =$g(x_4)$=0，分别作出通过加权空间原点的平面②③④；图中用阴影表示的部分是各平面的正侧。

设：$x_1, x_2 \in \omega_1$

$x_3, x_4 \in \omega_2$

$$\because g(x) \begin{matrix} > \\ < \end{matrix} 0 \quad \begin{matrix} x \in \omega_1 \\ x \in \omega_2 \end{matrix}$$

$$\left.\begin{matrix} w_1 x_{11} + w_2 x_{12} + w_3 > 0 \\ w_1 x_{21} + w_2 x_{22} + w_3 > 0 \end{matrix}\right\} \omega_1$$

$$\left.\begin{matrix} w_1 x_{31} + w_2 x_{32} + w_3 < 0 \\ w_1 x_{41} + w_2 x_{42} + w_3 < 0 \end{matrix}\right\} \omega_2$$

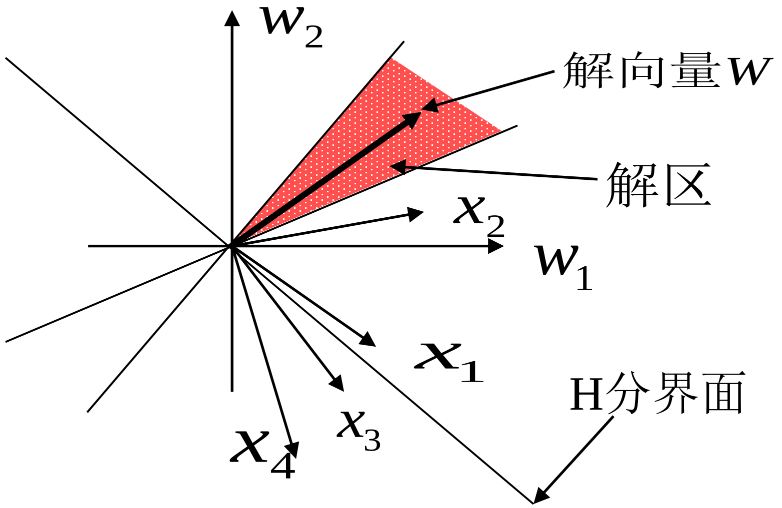
最终形成凸多面锥

## 1. 模式空间与加权空间(Cont.)

- 这是一个不等式方程组，它的解 $W=(w_1,w_2,w_3)^T$ 处于由 $\omega_1$类所有模式决定的平面的正边和由 $\omega_2$类所有模式决定的平面的负边，它的解区即为凸多面锥。
- 如图所示：(b)为加权空间，(c)为正规化后的加权空间。
- 由上可以得出结论：加权空间的所有分界面都通过坐标原点。这是加权空间的性质。
- 为了更清楚，下面用二维权空间来表示解向量和解区。

## 2. 解向量和解区

- 在三维空间里，令 $w_3 = 0$ 则为二维权空间。如右图：
- 给定一个模式 $X$，就决定一条直线：$g(x) = W^T X = 0$
- 即分界面H，W与H正交，W称为解向量。
- 解向量的变动范围称为解区。
- 因 $x_1, x_2 \in \omega_1$，$x_3, x_4 \in \omega_2$ 由图可见 $x_1, x_3$ 离的最近，所以分界面H可以是 $x_1, x_3$ 之间的任一直线，由垂直于这些直线的W就构成解区，解区为一扇形平面，即红色阴影区域。
- 如右图：



解向量与解区

# 2.解向量和解区(Cont.)

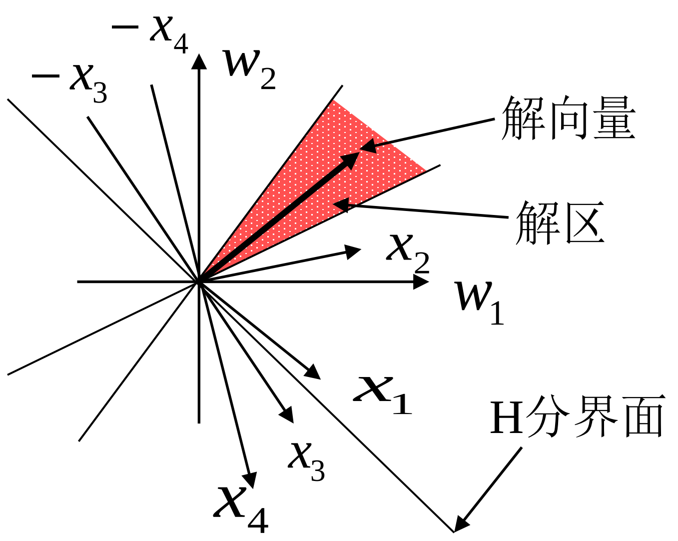
❖将不等式方程正规化：

$$\begin{cases} w_1 x_{11} + w_2 x_{12} + w_3 > 0 \\ w_1 x_{21} + w_2 x_{22} + w_3 > 0 \\ -w_1 x_{31} - w_2 x_{32} - w_3 > 0 \\ -w_1 x_{41} - w_2 x_{42} - w_3 > 0 \end{cases}$$

❖正规化：

$$g(x) = W_i^T X > 0$$

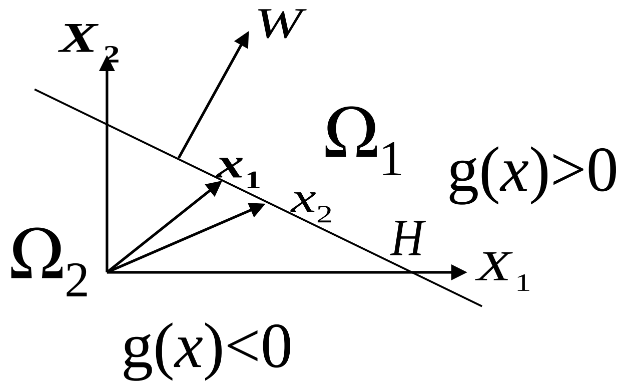
$$W = (w_1, w_2, \ldots, w_n, w_{n+1})$$



正规化

## 3.超平面的几何性质

- $g(x)=W^TX=0$决定一个决策界面，当$g(x)$为线性时，这个决策界面便是一个超平面***H***，并有以下性质：
- 性质①：***W***与***H***正交（见右图）




. 假设$x_1$,$x_2$是***H***上的两个向量

$$W^T x_1 + w_{n+1} = W^T x_2 + w_{n+1} = 0$$

所以$W^T(x_1 - x_2) = 0$, $(x_1 - x_2)$向量一定在$H$上

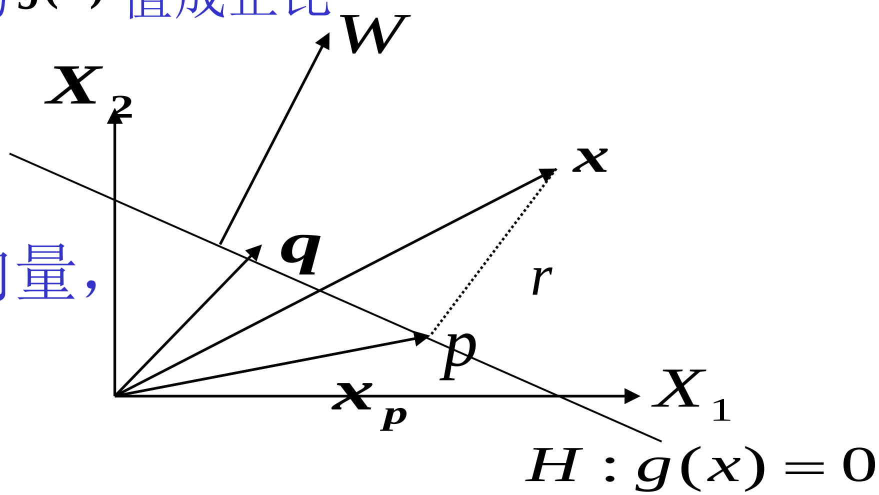
. ***W***与$(x_1$-$x_2)$ 垂直，即***W***与***H***正交。

. 一般说，超平面***H***把特征空间分成两个半空间。即$\Omega_1$,$\Omega_2$空间，当$x$在$\Omega_1$空间时$g(x)>0$,***W***指向$\Omega_1$，为***H***的正侧，反之为***H***的负侧。

# 3.超平面的几何性质(Cont.)

- 性质 ②：$\|r\| = \frac{g(x)}{\|W\|}$
- 向量$x$到超平面$H$的正交投影 $\|r\|$ 与$g(x)$ 值成正比

$$x = x_p + r = x_p + |r|\frac{W}{\|W\|}$$




- 其中： $x_p$为 $x$在$H$ 的投影向量，
- $r$是$x$ 到$H$的垂直距离。
- $\frac{W}{\|W\|}$是$W$方向的单位向量。

教材：P46

# 3.超平面的几何性质(Cont.)

- 另一方面：

$$g(x)=W^T x+w_{n+1}=W^T(x_p+r)+w_{n+1}$$

$$\because p在H上, \therefore W^T x_p+w_{n+1}=0$$

$$\therefore g(x)=W^T r=W^T(\|r\|\frac{W}{\|W\|})=\|r\|\frac{W^T W}{\|W\|}=\|r\|\|W\|$$

$$\therefore \|r\|=\frac{g(x)}{\|W\|}, \|r\|是投影的绝对值(W^T W=\|W\|^2)$$

- 这是超平面的第二个性质，向量$x$到超平面H的正交投影正比于g($x$)的函数值。

## 3.超平面的几何性质(Cont.)

- 性质③:

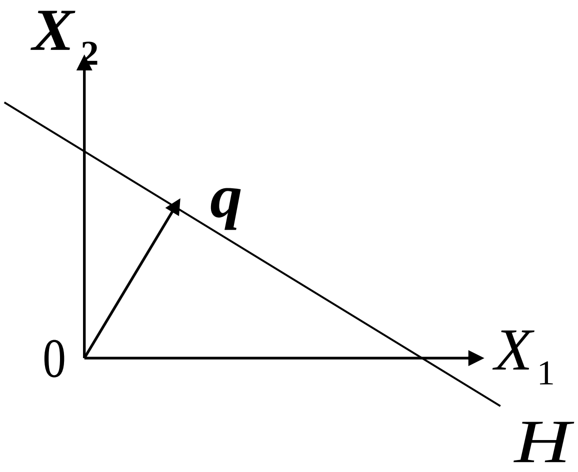
$$\|q\| = \frac{W_{n+1}}{\|W\|}$$,原点到超平面$H$的距离与$W_{n+1}$成正比

$\because g(x) = W^T x + w_{n+1} = w_{n+1}$(因原点$x = 0$)

$$\therefore \|r\| = \frac{g(x)}{\|W\|} = \frac{W_{n+1}}{\|W\|} = \|q\|$$

$\because x = 0$时$x$到$H$的投影为$\|r\| = \|q\|$

$$\therefore \|q\| = \frac{W_{n+1}}{\|W\|}$$



# 3.超平面的几何性质(Cont.)

- 性质④:

若$W_{n+1}>0$,则$H$在原点正侧，若$W_{n+1}<0$,则$H$在原点负侧。

若$W_{n+1}=0$,则$g(x)=W^T x$，说明超平面$H$通过原点。

结论:

$(a)$超平面$H$的平面与$W$正交，方向由$W$决定。

$(b)$超平面$H$的位置由$W_{n+1}$来决定。

$(c)$$g(x)$正比于$x$到$H$代数距离。

$x$在$H$的正侧，$g(x)>0$;否则，反之。

# 2.4 广义线性判别函数

- 判别函数的一般形式：

$$g(x)=w_1 f_1(x)+w_2 f_2(x)+\ldots+w_k f_k(x)+w_{k+1}$$

$$=\sum_{i=1}^{k+1} w_i f_i(x), i=1,2,\ldots,k \qquad \text{式中} f_i(x) \text{是单值函数,} f_{k+1}(x)=1$$

- 这样一个非线性判别函数通过映射，变换成线性判别函数。

$$g(x)=\sum_{i=1}^{k+1} w_i f_i(x) \xrightarrow{x\text{空间}\to\text{变换}\to y\text{空间}}$$

$$W^T Y=g(Y)\begin{cases} >0, x\in\omega_1 \\ <0, x\in\omega_2 \end{cases}$$

# 2.4 广义线性判别函数(Cont.)

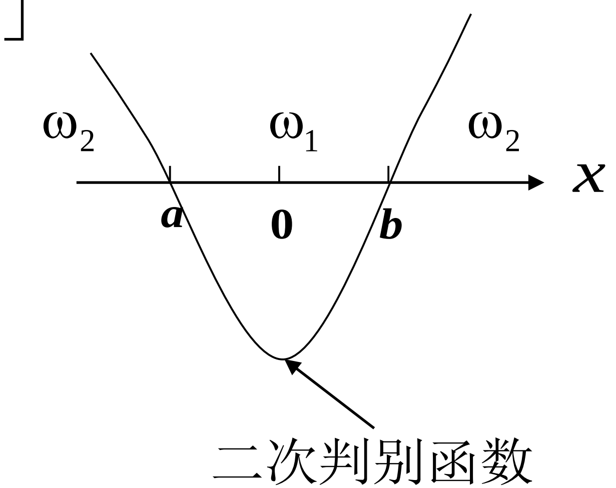
$$g(x)=\sum_{i=1}^{k+1} w_i f_i(x) \xrightarrow{x\text{空间}\to\text{变换}\to y\text{空间}} W^T Y = g(Y)\begin{cases} >0, x\in\omega_1 \\ <0, x\in\omega_2 \end{cases}$$

其中： $W=\begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_k \end{bmatrix}$(广义权向量)。$Y=\begin{bmatrix} f_1(x) \\ f_2(x) \\ \cdots \\ f_k(x) \end{bmatrix}$(增广模式向量)

❖例：见右图。

$x>a, or, x<b$, 则$x\in\omega_1$

$b<x<a$, 则$x\in\omega_2$



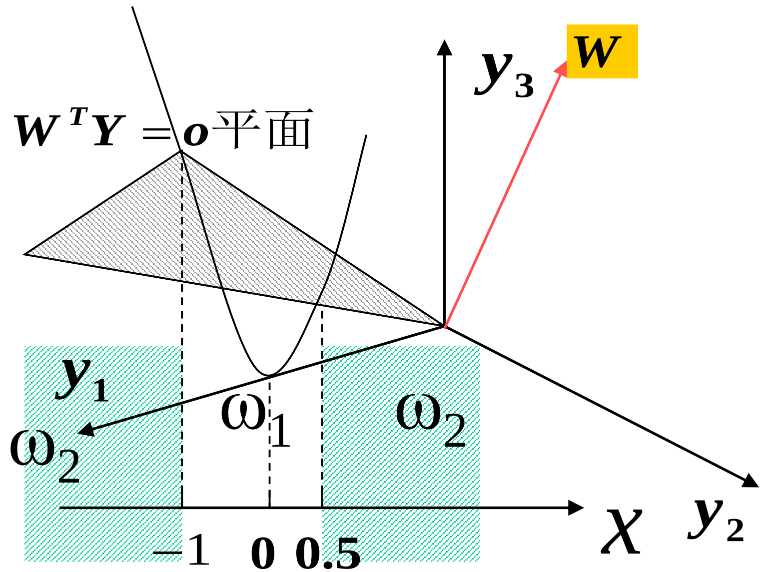
# 2.4 广义线性判别函数(Cont.)

- 要用二次判别函数才可将二类分开：

$$g(x)=a_1+a_2x+a_3x^2\begin{cases}>0, x\in\omega_1\\<0, x\in\omega_2\end{cases}$$

映射：$g(x)=W^TY=g(Y)\begin{cases}>0, x\in\omega_1\\<0, x\in\omega_2\end{cases}$

$$W=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}, Y=\begin{bmatrix}1\\x\\x^2\end{bmatrix}$$




$$y_1=(1,0,0), y_2=(1,0.5,0.25)$$
$$y_3=(1,-1,1)$$

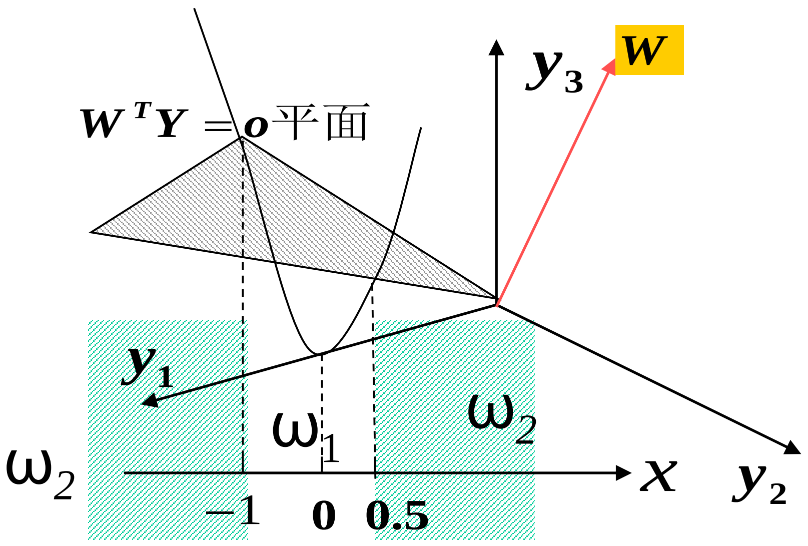
# 2.4 广义线性判别函数(Cont.)

讨论在$x$空间它的判别边界：设$a_1=-1, a_2=1, a_3=2$

即：$W=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}-1\\1\\2\end{bmatrix}, Y=\begin{bmatrix}1\\x\\x^2\end{bmatrix}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$

$g(x)=2x^2+x-1=0 \xrightarrow{\text{推出}} \begin{cases}x_1=0.5\\x_2=-1\end{cases}$

$Y$空间判别平面：$W^TY=0$




- 从图可以看出：在阴影上面是$\omega_1$类，在阴影下面是$\omega_2$类。
- 结论：在X空间的非线性判别函数通过变换到Y空间成为线性的，但X变为高维空间。

## 课堂思考(课后写在作业本上):

1.若X=[2 3], Y=[3 1], 求COV(X,Y).

2.$已知X=\begin{bmatrix} 4 & 5 & 1 \\ 3 & 1 & 0 \\ 2 & 3 & 2 \end{bmatrix}, 求X^{-1}, COV(X)$

3.有一个二类问题，其判别函数为 $g(X)=3x_1+5x_2-6x_3-2$。试将下面三个模式分别进行分类：$X1=[4\ 7\ 1]^T$,$X2=[1\ -5\ 2]^T$,$X3=[4\ 4\ 5]^T$ 。

# 1.分析：对于二维随机变量(X,Y)

$Definition:$

$$\operatorname{cov}(X,Y)=E[(X-E(X))(Y-E(Y)]=E(XY)-E(X)E(Y)$$

$$\operatorname{cov}(X,Y)=\begin{bmatrix} s_{11}^2 & s_{12}^2 \\ s_{21}^2 & s_{22}^2 \end{bmatrix}$$

$$s_{11}^2=\sum(x_i-\bar{x})^2\Big/(n-1)$$

$$s_{12}^2=s_{21}^2=\sum(x_i-\bar{x})(y_i-\bar{y})\Big/(n-1)$$

$$=\sum x_i y_i-\sum x_i\sum y_i\Big/n$$

$$s_{22}^2=\sum(y_i-\bar{y})^2\Big/(n-1)$$

$$r_{ij}=\frac{s_{ij}^2}{\sqrt{s_{ii}^2 s_{jj}^2}}$$，$MATLAB$中求相关使用$corrcoef(X,Y)$函数

# 解：

$$n=2$$

$$\because \operatorname{cov}(X,Y)=\begin{bmatrix} s_{11}^2 & s_{12}^2 \\ s_{21}^2 & s_{22}^2 \end{bmatrix}$$

$$s_{11}^2=\{(2-2.5)^2+(3-2.5)^2\}/(n-1)=0.5$$

$$s_{12}^2=s_{12}^2=\{(2-2.5)(3-2)+(3-2.5)(1-2)\}/(n-1)=-1$$

$$s_{22}^2=\{(3-2)^2+(1-2)^2\}/(n-1)=2$$

$$\therefore \operatorname{cov}(X,Y)=\begin{bmatrix} 0.5 & -1 \\ -1 & 2 \end{bmatrix}$$

## 2.分析:

对于n维随机变量$(X_1, X_2, ...., X_n)$,定义

$\sigma_{ij} = COV(X_i, X_j) = E(X_i - EX_i)(X_j - EX_j)$,则称

$$\Sigma = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \cdots & \sigma_{nn}^2 \end{pmatrix}$$

为$(X_1, X_2, ...., X_n)$的协方差矩阵

$$对于矩阵X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

$$有协方差矩阵\Sigma=\begin{pmatrix} s_{11}^2 & s_{12}^2 & \cdots & s_{1n}^2 \\ s_{21}^2 & s_{22}^2 & \cdots & s_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ s_{n1}^2 & s_{n2}^2 & \cdots & s_{nn}^2 \end{pmatrix}$$

$$s_{11}^2=\sum(x_{i1}-\overline{x}_1)^2\big/(n-1)$$

$$s_{12}^2=\sum(x_{i1}-\overline{x}_1)(x_{i2}-\overline{x}_2)\big/(n-1)$$

$$=\sum x_{i1}x_{i2}-\sum x_{i1}\sum x_{i2}\big/n$$

……

$$s_{ij}^2=\sum_{k=1}^{n}(x_{ki}-\overline{x}_i)(x_{kj}-\overline{x}_j)\big/(n-1)$$

$$=\sum_{k=1}^{n}x_{ki}x_{kj}-\sum_{k=1}^{n}x_{ki}\sum_{k=1}^{n}x_{kj}\big/n$$

$$其中,\ \overline{\mathrm{x}}_{\mathrm{j}}=\sum_{k=1}^{n}x_{kj}$$

……

$$相关系数r_{ij}=\frac{s_{ij}^2}{\sqrt{s_{ii}^2s_{jj}^2}}$$

提示:MATLAB中, 求X的逆$X^{-1}$使用$INV(X)$, 求X矩阵的协方差矩阵使用$COV(X)$

答案:$COV(X)=\begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 4 & 1 \\ -\frac{1}{2} & 1 & 1 \end{bmatrix}$

3.答案：X1属于ω1类；X2属于第ω2类；无法判断X3属于二类中的哪一类。