2. FUNCIONES DE DECISIÓN

1.2	Introducción	18
2.2	Funciones de decisión lineales	19
	2.2.1. Ejemplo	28
2.3	Funciones de decisión generalizadas	33
2.4	Funciones de varias variables	34
	2.4.1 Definiciones	34
	2.4.2 Construcción de funciones multivariadas	36
	2.4.3 Sistema de fuciones ortogonales y ortonormales	36
	2.4.4 Ejemplo	40

2. FUNCIONES DE DECISIÓN

2.1. INTRODUCCIÓN

La principal función de un sistema de reconocimiento de patrones es proveer funciones de decisión sobre la pertenencia de un elemento al patrón con el cual es confrontado.
A fin de llevar a cabo esta tarea, es necesario establecer algunas reglas que fundamenten nuestra decisión. Un método importante de este problema es el uso de FUNCIONES DE DECISIÓN.

Una manera de introducir este concepto relativamente fácil se muestra en la figura 2.1, en donde se muestran clases de patrones hipotéticos. Como puede verse en esta figura existen 2 poblaciones que pueden ser separadas fácilmente por una línea.

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

La ecuación de la línea que separa las w's y (x_1, x_2) son las coordenadas de las variables. De la observación de la figura se ve que cualquier patrón X perteneciente a la clase w1 será una cantidad positiva cuando se sustituya en d(x). De igual modo, d(x) será negativa para cualquier patrón de w_2
Sin embargo, d(x) puede usarse como una FUNCION DE DECISIÓN (O DISCRIMINANTE), dado un patrón X de clasificación desconocida, podemos decir que

$$\mathbf{x} \in w_i \quad \text{si} \quad d(\mathbf{x}) < 0$$
$$\mathbf{x} \in w_2 \quad \text{si} \quad d(\mathbf{x}) > 0$$

Si X descansa en la frontera entre w_1 y w_2, diremos que **x** es indeterminado $d(\mathbf{x}) = 0$.
Como hemos visto estos conceptos no se limitan a 2 clases. Además ellos pueden fácilmente extenderse al caso más general de fronteras no lineales en cualquier espacio de dimensión finita del espacio Euclidiano.

El éxito del esquema de clasificación de patrones depende de dos factores:
1) la forma de la función $d(\mathbf{x})$ y
2) la habilidad para determinar sus coeficientes w_1, w_2, w_3.

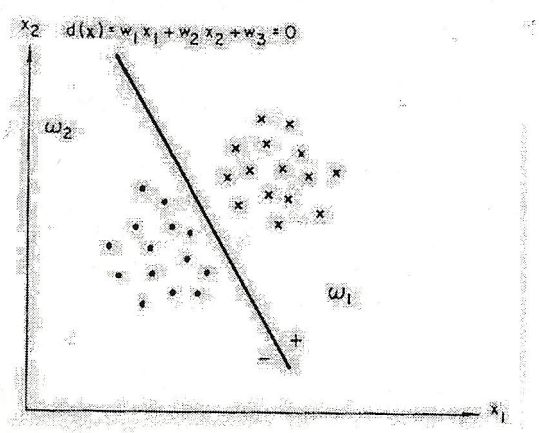


Figura 2.1 Una función de decisión simple para 2 clases [3c]

El primer problema está relacionado con las propiedades geométricas de las clases de patrones en consideración. No es difícil visualizar situaciones en las cuales existe dificultad para determinar las fronteras. Las fronteras se deciden a partir de la información a priori que tenemos del problema. Una vez que se seleccionaron las fronteras de separación entre las clases, el problema será la determinación de los coeficientes. Como veremos en las secciones que siguen existen esquemas que pueden aclarar el problema. Veremos como determinar los coeficientes.

2.2 FUNCIONES DE DECISION LINEALES

La manera mas simple de establecer fronteras entre número finito de clases es la ide que se dio anteriormente, extendiéndose al caso n-dimensional.

Ya que las funciones de decisión son de la forma

$$d(\mathbf{x}) = w_1 x_1 + w_2 x_2 + \ldots\ldots + w_n x_n + w_n + 1 = \mathbf{w_0}' \mathbf{x} + w_n + 1 \qquad (2.2\text{-}1)$$

donde

$$\mathbf{w_0} = (w_1, w_2, \ldots\ldots, w_n)'$$

Este vector se conoce como vector de pesos o VECTOR DE PARAMETROS
Por convención se ha aceptado ampliamente la notación

$$d(\mathbf{x}) = \mathbf{w}' \mathbf{x} \qquad (2.2\text{-}2)$$

en donde

$$\mathbf{x} = (x_1, x_2, \ldots\ldots, x_n, 1)' \quad y$$
$$\mathbf{w} = (w_1, \ldots\ldots, w_n, w_{n+1})'$$

Son llamados los patrones aumentados y vectores de peso, respectivamente. Ya que se aumenta la misma cantidad es igualmente aumentada a todos los patrones, las propiedades geométricas de las clases patrón no se alteran. Normalmente nos referiremos a **x** y **w** en la ecuación 2.2-2 simplemente como vectores patrón y de peso, respectivamente.

En un problema de dos clases la función de decisión d (**x**) suponemos que tiene la propiedad

$$d(\mathbf{x}) = \mathbf{w}' \mathbf{x} \quad \begin{cases} > 0 & si \quad \mathbf{x} \in \omega_1 \\ < 0 & si \quad \mathbf{x} \in \omega_2 \end{cases} \qquad (2.2\text{-}3)$$

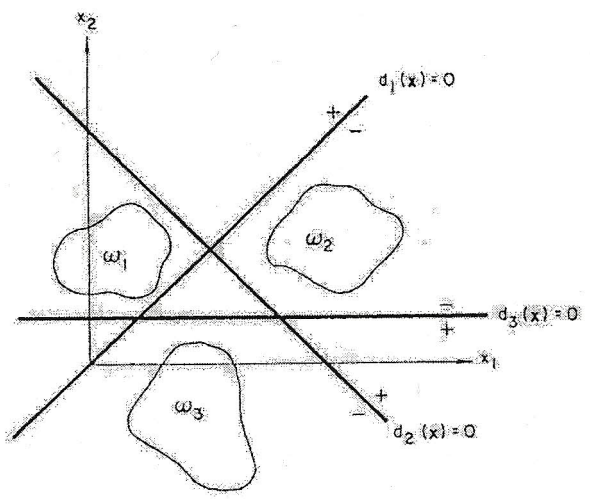
Cuando tenemos más de dos clases, las denotaremos por por $w_1, w_2, \ldots, w_M$ y vamos a considerar los siguientes tres casos multiclases.

Caso 1:

Cada patrón es separable de los otros patrones por una única superficie de decisión. En este caso necesitaremos M funciones de decisión con la propiedad

$$d_i(\mathbf{x}) = \mathbf{w}_i' \mathbf{x} = \left\{ \begin{matrix} > 0 & si \quad \mathbf{x} \in w_i \\ < 0 & de\ otra\ manera \end{matrix} \right\}, \quad i = 1, 2, \ldots, M \tag{2.2}$$

en donde $\mathbf{w}_i = (w_{i1}, w_{i2} \ldots, w_{in}, w_{i,n+1})'$ es el vector de ponderaciones asociado con la función de decisión i-ésima



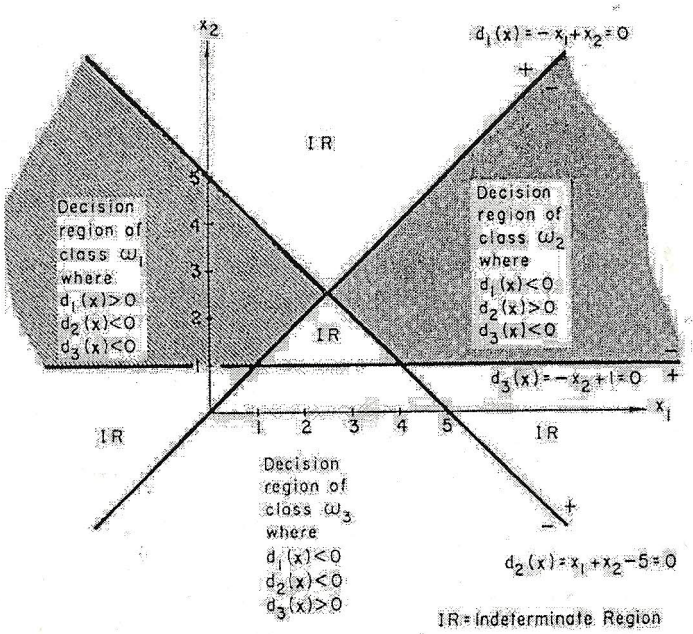


Figura 2.2 Caso 1 [36]

Ejemplo 2.1

Consideraremos un ejemplo sencillo del caso 1 multiclase que se muestra en la figura 2.2(a)

Se nota que cada clase es separable del resto por una única decisión de fronteras.
Por ejemplo, si un patrón específico $\mathbf{x} \in \mathbf{w_i}$, es claro, por la geometría de la figura 2.2.(a) que $d_1(\mathbf{x}) > 0$ mientras, $d_2(\mathbf{x}) < 0$ y $d_3(\mathbf{x}) < 0$.
La frontera entre la clase w_1 y las otras clases está dada por lo valores de **x** para los cuales $d_i(\mathbf{x}) = 0$.

Como ilustración numérica supongamos que las funciones de decisiones de la figura 2.2a son:

$$d_1(\mathbf{x}) = -x_1 + x_2$$
$$d_2(\mathbf{x}) = x_1 + x_2 - 5$$
$$d_3(\mathbf{x}) = -x_2 + 1$$

entonces, las tres funciones frontera son:

$$-x_1 + x_2 = 0$$
$$x_1 + x_2 - 5 = 0$$
$$-x_2 + 1 = 0$$

Entonces según las condiciones de la ecuación (2.2-4), cualquier patrón que satisface:

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

es automáticamente asignado a la clase w_1.

Y cualquier patrón que satisface:

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

es automáticamente asignado a la clase w_2.

Y cualquier patrón que satisface:

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

es automáticamente asignado a la clase w_3.

Sin embargo la región correspondiente a la clase w_1, está determinada por la línea positiva $-x_1 + x_2 = 0$ y el lado negativo de las líneas $x_1 + x_2 - 5 = 0$ y $-x_2 + 1 = 0$. Esta región se muestra en la figura 2.2 (b)

Nosotros vemos que aunque el caso de w_1 ocupa un área pequeña, la región de decisión real en la cual un patrón puede ser asignado a la clase w_1 se extiende al infinito. De igual manera resulta para las otras dos clases.
Es de notarse que en este ejemplo tenemos 4 regiones indeterminadas que se muestran con precisión en la figura 2.2b. Si algún punto por clasificar cae en alguna de estas regiones, no es posible clasificarlo.

Entonces para mostrar cómo funcionan las reglas de decisión supongamos por ejemplo que queremos clasificar el patrón $\mathbf{x} = (6,5)$.
Sustituyendo estos valores en las funciones de decisión tendremos:

$$d_1(\mathbf{x}) = -1$$
$$d_2(\mathbf{x}) = 6$$
$$d_3(\mathbf{x}) = -4$$

ya que y de acuerdo a las condiciones de la ecuación (2.2-4)

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

entonces el patrón es asignado a la clase w_2.

Caso 2

Cada patrón es separable de las otras clases por una superficie de decisión diferente esto es, las clases serán separables de 2 a 2. En este caso, existen M(M-1)/2 decisiones de superficie posibles (tomadas de 2 en 2). Las funciones de decisión aquí son de la forma

$$d_{ij}(\mathbf{x}) = \mathbf{w}'_{ij}\,\mathbf{x}$$

y tienen la propiedad de que

$$\text{si } \mathbf{x} \in w_i, \text{ entonces}$$
$$d_{ij}(\mathbf{x}) > 0 \quad \forall \quad j \neq i \qquad (2.2\text{-}5)$$

estas funciones tienen también la propiedad de que

$$d_{ij}(\mathbf{x}) = -d_{ji}(\mathbf{x})$$

entonces,

$$d_{12}(\mathbf{x}) = -d_{21}(\mathbf{x})$$
$$d_{23}(\mathbf{x}) = -d_{32}(\mathbf{x})$$
$$d_{13}(\mathbf{x}) = -d_{31}(\mathbf{x})$$

Ejemplo 2.2.

La figura 2.3 (a) ilustra tres clases de patrones separables bajo las condiciones del caso 2. Vemos que las clases no son separables por una superficie de decisión única. Cada frontera es capaz de separar en dos clases.

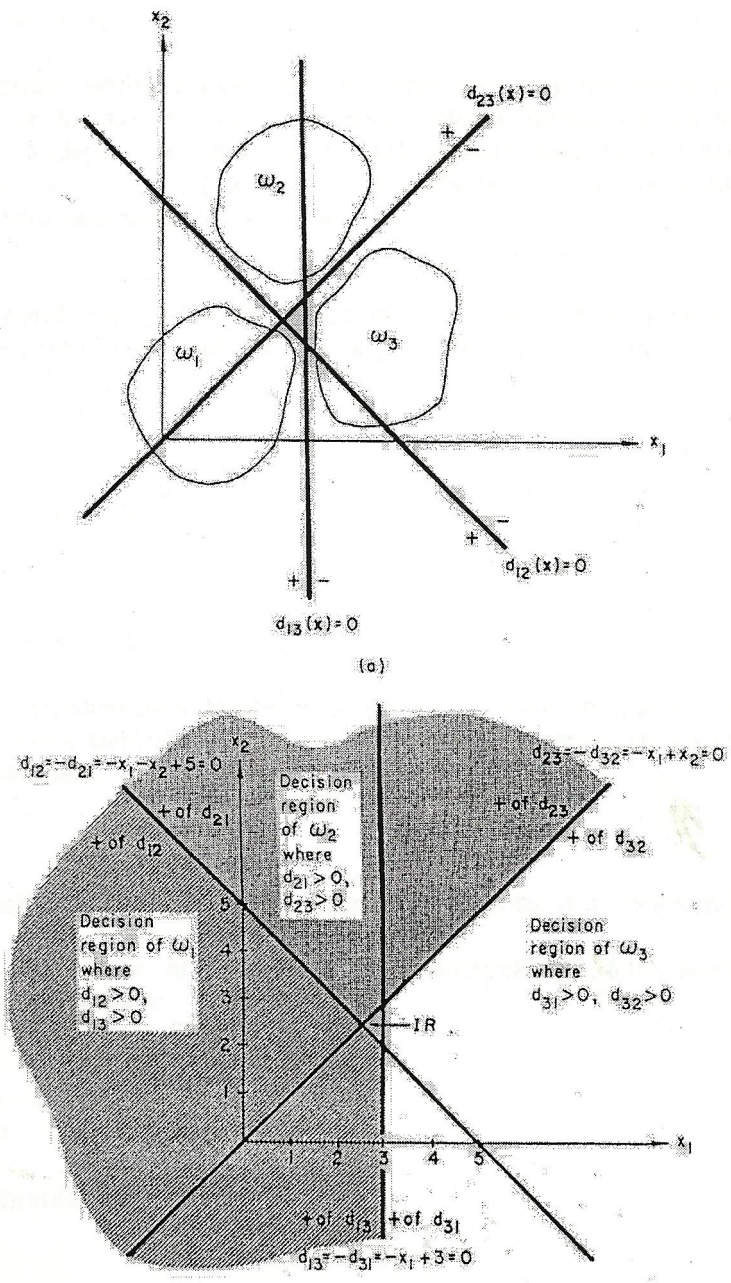


Figura 2.3 caso 2. [36]

Por ejemplo aunque la frontera $d_{12}(\mathbf{x}) = 0$ corta la región de la clase w_3, efectivamente sólo separa w_1 y w_2.

Para propósito de ilustración supongamos los siguientes valores numéricos:

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

Las fronteras de decisión están otra vez determinadas haciendo las funciones de decisión igual a cero. Las regiones de decisión sin embargo, están ahora dadas por los lados positivos de las fronteras de decisión múltiple. La región de la clase w_1 está determinada por los valores para los cuales $d_{12}(\mathbf{x}) > 0$ y $d_{13}(\mathbf{x}) > 0$. El valor de $d_{23}(\mathbf{x})$ en esta región es irrelevante ya que $d_{23}(\mathbf{x})$ no está relacionado con la clase w_1.

Las regiones correspondientes a las tres funciones de decisión mencionadas arriba se muestran en la figura 2.3 (b) en donde la condición $d_{ij}(\mathbf{x}) = -d_{ji}(\mathbf{x})$ ha sido usada para determinar las regiones de las clases.

Por ejemplo ya que

$$d_{12}(\mathbf{x}) = -x_1 - x_2 + 5$$

tenemos entonces que

$$d_{21}(\mathbf{x}) = x_1 + x_2 - 5$$

y se sigue que el lado positivo de la fronteras $d_{12}(\mathbf{x}) = 0$ está en el lado negativo de la frontera $d_{21}(\mathbf{x}) = 0$. Vemos que también existen regiones indeterminadas en las cuales las condiciones del caso 2 no se satisfacen.

Para mostrar el funcionamiento del clasificador, supongamos que deseamos clasificar el patrón $\mathbf{x} = (4,3)'$.
Sustituyendo estos valores en el patrón en las ecuaciones de éste ejemplo, nos lleva a las siguientes funciones de decisión:

$$d_{12}(\mathbf{x}) = -2$$
$$d_{13}(\mathbf{x}) = -1$$
$$d_{23}(\mathbf{x}) = -1$$

se sigue automáticamente que

$$d_{21}(\mathbf{x}) = 2$$
$$d_{31}(\mathbf{x}) = 1$$
$$d_{32}(\mathbf{x}) = 1$$

esto es,

$$d_{3j}(\mathbf{x}) > 0 \text{ para } j = 1,2$$

usando (2.2.5) asignamos este patrón a w_3.

Caso 3

Existen M funciones de decisión $d_k(\mathbf{x}) = \mathbf{w}'_k \mathbf{x}$ $k = 1, M$ con la propiedad de que

si $x \in$ a la clase w_i, entonces

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \text{ para todo } j \neq i \qquad (2.2\text{-}6)$$

Este es un caso especial del caso2 ya que podemos definir:

$$\begin{aligned} d_{ij}(\mathbf{x}) &= d_i(\mathbf{x}) - d_j(\mathbf{x}) \\ &= (\mathbf{w}_i - \mathbf{w}_j)' \mathbf{x} \\ &= \mathbf{w}'_{ij} \mathbf{x} \end{aligned}$$

En donde

$$\mathbf{w}_{ij} = \mathbf{w}_i - \mathbf{w}_j$$

Se puede verificar que si

$$d_i(\mathbf{x}) > d_i(\mathbf{x}) \quad \forall \quad j \neq i \Rightarrow d_{ij}(\mathbf{x}) > 0 \quad \forall \quad j \neq i$$

Esto es, si las clases son separables bajo las condiciones del caso 3 son automáticamente separables para el caso 2. En caso inverso no es válido.

Ejemplo 2.3

Antes de ilustrar el caso 3 hacemos notar primero que la frontera entre las clases w_i y w_j está dada por los valores de **x** para los cuales

$$d_i(\mathbf{x}) = d_j(\mathbf{x})$$

O dicho de otra manera:

$$d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0$$

Así vemos que los valores de d_i (**x**) y d_j (**x**) se combinan para dar una ecuación de la superficie que separa w_i de w_j

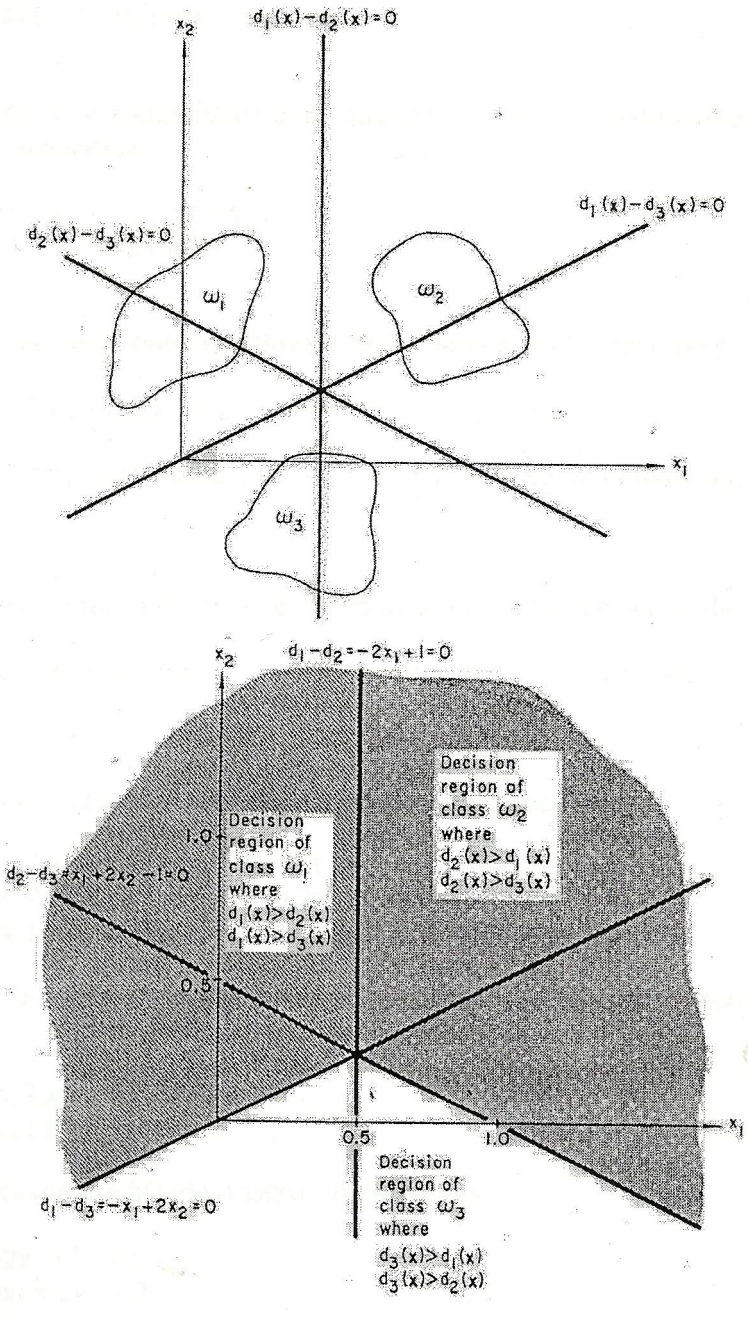


Figura 2.4 Caso 3 [36]

Un ejemplo simple del caso 3 se muestra en la figura 2.4 (a) para M = 3.
Para los patrones de la clase w_1 requerimos que

$$d_1(\mathbf{x}) > d_2(\mathbf{x}) \text{ y}$$
$$d_1(\mathbf{x}) > d_3(\mathbf{x})$$

Esto es equivalente al requerimiento de que los patrones de esta clase descansan sobre el lado positivo de las superficies

$$d_1(\mathbf{x}) - d_2(\mathbf{x}) = 0 \quad \text{y}$$
$$d_1(\mathbf{x}) - d_3(\mathbf{x}) = 0$$

en general, necesitamos que las patrones de la clase w_i estén del lado positivo de las superficies

$$d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, M, \quad j \neq i$$

Como antes el lado positivo de $d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0$ está sobre el lado negativo de $d_j(\mathbf{x}) - d_i(\mathbf{x}) = 0$.

Como ilustración numérica de éste caso, consideremos las funciones de decisión

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

Las fronteras entre las tres clases pueden determinarse como sigue:

$$d_1(\mathbf{x}) - d_2(\mathbf{x}) = -2x_1 + 1 = 0$$
$$d_1(\mathbf{x}) - d_3(\mathbf{x}) = -x_1 + 2x_2 = 0$$
$$d_2(\mathbf{x}) - d_3(\mathbf{x}) = x_1 + 2x_2 - 1 = 0$$

Para encontrar la región de decisión correspondiente a la clase w_1, encontramos la región para la cual

$$d_1(\mathbf{x}) > d_2(\mathbf{x}) \text{ y}$$
$$d_1(\mathbf{x}) > d_3(\mathbf{x})$$

Esta región corresponde al lado positivo de las líneas

$$-2x_1 + 1 = 0 \quad \text{y}$$
$$-x_1 + 2x_2 = 0$$

Como se muestra en la figura 2.4 (b)

La región para la clase w_2 está determinada por los lados positivos de las líneas

$$2x_1 - 1 = 0 \quad \text{y} \quad x_1 + 2x_2 - 1 = 0$$

Finalmente la región de w_3 está dado por dos lados positivos de las rectas

$x_1 - 2x_2 = 0$ y
$-x_1 - 2x_2 + 1 = 0$

Hay que hacer notar que en este caso 3 ya no existen regiones indeterminadas.

Como ejemplo del proceso que se sigue en la clasificación consideremos el patrón $\mathbf{x} = (1,1)$. Sustituyendo este patrón en las funciones de decisión definidas, obtendremos:

$$d_1(\mathbf{x}) = 0, \quad d_2(\mathbf{x}) = 1 \quad y \quad d_3(\mathbf{x}) = -1$$

Ya que y de acuerdo a las reglas de decisión (2.2-6)

$$d_2(\mathbf{x}) > d_j(\mathbf{x}) \quad \forall \quad j=1,3$$

Entonces, este patrón se asigna a la clase w_2.

2.2.1. EJERCICIO.

Ejemplo 2.4.

Supongamos un ejemplo de funciones de decisión para un problema multiclase en donde M=3: $\omega_1, \omega_2, \omega_3$ en donde

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

a) suponiendo que estas fuciones fueron determinadas bajo las hipótesis del caso1 para multiclases, dibujar las fronteras de decisión y las regiones para cada patrón clase.

b) Suponiendo las condiciones del caso 2 y haciendo

$d_{12}(\mathbf{x}) = d_1(x)$ $\quad$ $d_{13}(\mathbf{x}) = d_2(x)$ $\quad$ $d_{23}(\mathbf{x}) = d_3(x)$

dibujar las fronteras de decisión y las regiones para cada clase siguiendo las hipótesis del caso 2.

c) Suponiendo que $d_1(x), d_2(x), d_3(x)$ fueron determinadas bajo las condiciones del caso 3, dibujar las fronteras de decisión y las regiones para cada clase.

Solución:

a) La regla de asignación para las clases es:

+ para asignar **x** a la clase ω_1 es necesario que se cumplan las siguientes condiciones $d_1(\mathbf{x}) > 0$
$d_2(\mathbf{x}) < 0$
$d_3(\mathbf{x}) < 0$

+para asignar **x** a la clase ω_2 es necesario que se cumplan las siguientes condiciones $d_1(\mathbf{x}) < 0$
$d_2(\mathbf{x}) > 0$
$d_3(\mathbf{x}) < 0$

+ para asignar **x** a la clase ω_3 es necesario que se cumplan las siguientes condiciones $d_1(\mathbf{x}) < 0$
$d_2(\mathbf{x}) < 0$
$d_3(\mathbf{x}) > 0$

Figura 2.5 muestra las regiones y las fronteras de decisión:

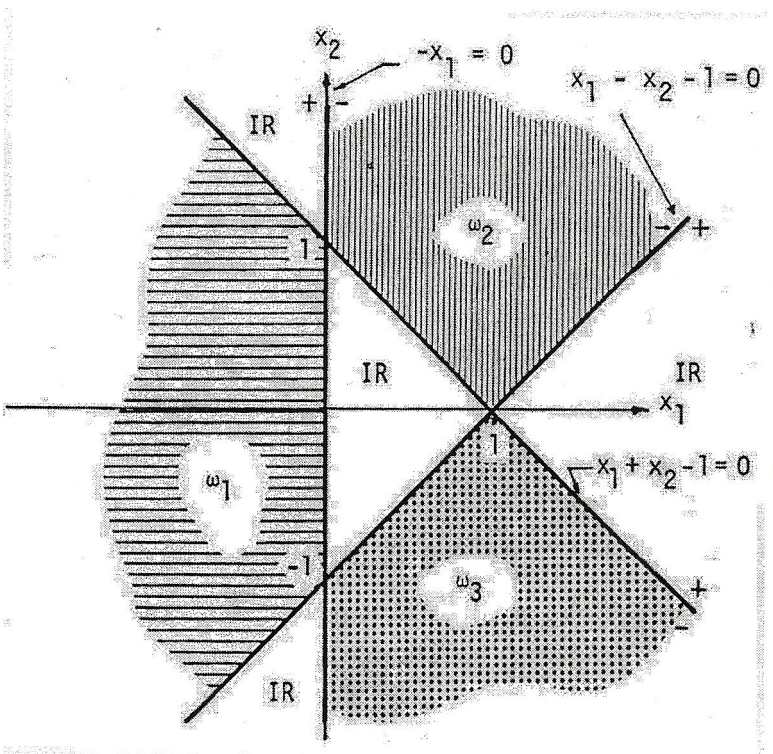


Figura 2.5 Regiones y frontera de decisión de (a). caso 1 [3.7]

b) Cada clase es separable de las otras clases por una superficie de decisión distinta, esto es, las clases son separables dos a dos, Si tenemos M=3 clases entonces tendremos

$\binom{M}{2}$ = 3 superficies de decisión.

$d_{12}(\mathbf{x}) = d_1(\mathbf{x}) = -x_1$

$d_{13}(\mathbf{x}) = d_2(\mathbf{x}) = x_1 + x_2 - 1$

$d_{23}(\mathbf{x}) = d_3(\mathbf{x}) = x_1 - x_2 - 1$

La regla de asignación para las clases es: (según 2.2-5)

+ para asignar **x** a la clase ω_1 es necesario que se cumplan las siguientes condiciones $d_{12}(\mathbf{x}) > 0$
$d_{13}(\mathbf{x}) < 0$

+para asignar **x** a la clase ω_2 es necesario que se cumplan las siguientes condiciones $d_{12}(\mathbf{x}) < 0$
$d_{13}(\mathbf{x}) > 0$
$d_{23}(\mathbf{x}) < 0$

+ para asignar **x** a la clase ω_3 es necesario que se cumplan las siguientes condiciones $d_{31}(\mathbf{x}) > 0$
$d_{32}(\mathbf{x}) > 0$

La figura 2.6 muestra las regiones y las fronteras de decisión:

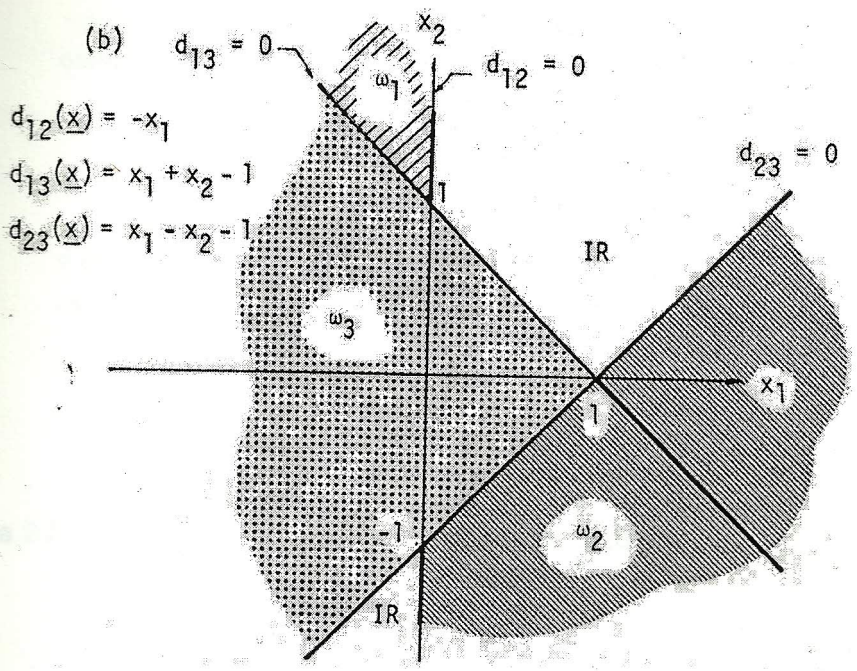


Figura 2.6 Regiones y frontera de (b). caso 2 [36]

c) Las funciones de decisión están dadas por:

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

$d_1(\mathbf{x}) - d_2(\mathbf{x}) = -2x_1 - x_2 + 1$
$d_1(\mathbf{x}) - d_3(\mathbf{x}) = -2x_1 + x_2 + 1$
$d_2(\mathbf{x}) - d_3(\mathbf{x}) = 2x_2$

La regla de asignación para las clases es: (según 2.2.-6)

+ para asignar **x** a la clase ω_1 es necesario que se cumplan las siguientes condiciones

$d_1(\mathbf{x}) - d_2(\mathbf{x}) > 0$ entonces $d_1(\mathbf{x}) > d_2(\mathbf{x})$
$d_1(\mathbf{x}) - d_3(\mathbf{x}) > 0$ entonces $d_1(\mathbf{x}) > d_3(\mathbf{x})$

$d_1(\mathbf{x}) - d_2(\mathbf{x}) = -2x_1 - x_2 + 1 = 0$
$d_1(\mathbf{x}) - d_3(\mathbf{x}) = -2x_1 + x_2 + 1 = 0$

+ para asignar **x** a la clase ω_2 es necesario que se cumplan las siguientes condiciones

$d_2(\mathbf{x}) - d_1(\mathbf{x}) > 0$ entonces $d_2(\mathbf{x}) > d_1(\mathbf{x})$

$d_2(\mathbf{x}) - d_3(\mathbf{x}) > 0$ entonces $d_2(\mathbf{x}) > d_3(\mathbf{x})$

+ para asignar **x** a la clase ω_3 es necesario que se cumplan las siguientes condiciones

$d_3(\mathbf{x}) - d_1(\mathbf{x}) > 0$ entonces $d_3(\mathbf{x}) > d_1(\mathbf{x})$

$d_3(\mathbf{x}) - d_2(\mathbf{x}) > 0$ entonces $d_3(\mathbf{x}) > d_2(\mathbf{x})$

La figura 2.7 muestra las regiones y las fronteras de decisión:

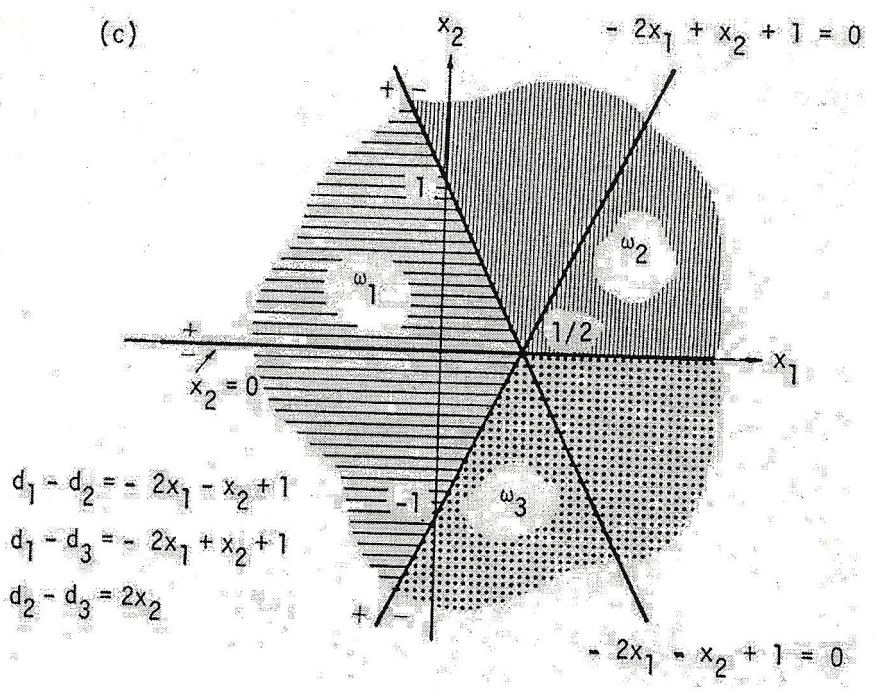


Figura 2.7 Regiones y fronteras de decisión para (c). caso 3 [36]

2.3. FUNCIONES DE DECISION GENERALIZADAS

Para el caso de funciones de decisión lineales teníamos en (2.2-2)

$$d(\mathbf{x}) = \mathbf{w}'\mathbf{x} = \sum_{i=1}^{k+1} w_i x_i$$

en donde $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \\ 1 \end{bmatrix}$

es posible generalizar este concepto al caso en que $\{x_i\}$ no es un real, sino una función, esto es $\{f_i(\mathbf{x})\}$ i=1,k son funciones reales de variable real, entonces,

$$d(\mathbf{x}) = w_1 f_1(\mathbf{x}) + w_2 f_2(\mathbf{x}) + \dots + w_k f_k(\mathbf{x}) + w_{k+1}$$

$$= \sum_{i=1}^{k+1} w_i f_i(\mathbf{x})$$

entonces, las ecuaciones (2.3-2) pueden expresarse como

$$d(\mathbf{x}) = \mathbf{w}'\mathbf{x}^*$$

donde $\mathbf{w} = (w_1, w_2, \dots, w_k, w_{k+1})$ y ser resueltas de la misma manera.

2.4. FUNCIONES DE VARIAS VARIABLES [5]

Las funciones multivariadas juegan un papel central en el estudio y diseño de sistemas de reconocimiento de patrones. Veremos en esta sección, cómo se construyen estas funciones para el caso de una variable y después los resultados pueden extenderse al caso multivariado.

2.4.1. Definiciones 2.4.1.[6]

Sean dos funciones f(x) y g(x) continuas sobre un intervalo cerrado [a,b] el *producto interior* se define como

$$(f\ ,\ g) = \int_a^b f(x)g(x)dx \qquad (2.4\text{-}1)$$

y el producto interior sobre la misma función f(x):

$$(f\ ,f) \quad = \int_a^b f^2(x)dx \qquad (2.4\text{-}2)$$

| f | se denomina *norma* de f(x) y si | f | = 1 se dice que f está *normalizada.*Una función puede normalizarse dividiéndola entre la raíz cuadrada de su norma.

Sean f(x) y g(x) se dice que son ortogonales con respecto a una función de ponderación u(x) en el intervalo cerrado [a,b], si

$$\int_a^b u(x)f(x)g(x)dx = 0\ .$$

Ejemplos de estas funciones son las funcion es polinomiales de Legendre, Laguerre, Hermite, Fourier.

---

[5] El contenido de esta sección representan los antecedentes teóricos para más tarde en el capítulo 5 poder obtener el clasificador de Bayes aproximando la función de densidad de probabilidad a partir de los datos de cada clase.

[6] Presentamos las definiciones de función ortogonal, ortonormal y sistema completo (base) y su fuente es el libro de Tou y González [ ].

Un sistema de funciones $\Phi_1(x), \Phi_2(x), \ldots$ son ortogonales sobre [a,b] si
Para cualquier par de funciones se cumple la *condición de ortogonalidad familiar:*

$$\int_a^b u(x)\Phi_i(x)\Phi_j(x)dx = A_{ij}\delta_{ij} \quad \text{en donde } \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \qquad (2.4\text{-}4)$$

y A_{ij} es un factor que depende de i y de j. Ya que la parte izquierda de la ecuación 2.4-4 es cero
excepto cuando i=j para abreviar la notación, denotaremos A_{ij} por A_i
Si A_i =1 entonces el sistema de ecuaciones se llama *sistema ortonormal* y la condición de ortonormalidad es:

$$\int_a^b u(x)\Phi_i(x)\Phi_j(x)dx = \delta_{ij} \qquad (2.4\text{-}6\ a)$$

Es común absorber ese término u(x) entonces:

$$\int_a^b \Phi_i(x)\Phi_j(x)dx = \delta_{ij} \qquad (2.4\text{-}6b)$$

Si un sistema de funciones $\Phi_i^*(x), i = 1,2,\ldots$ es ortogonal en el intervalo cerrado [a,b] es posible obtener un sistema ortonormal por medio de la relación

$$\Phi_i(x) = \sqrt{\frac{u(x)}{A_i}}\,\Phi_i^*(x) \qquad (2.4\text{-}7)$$

En donde A_i se obtiene de la ecuación 2.4-4 con i=j

Se puede demostrar que el conjunto de las $\{\Phi_i(x)\}$ son ortonormales.

Sea $\{f_1(x), f_2(x), \ldots, f_m(x)\}$ se dice que $\{f_i\}_{i=1}^m$ son linealmente independientes si

$\exists$ $c_1, c_2, \ldots, c_n$ tal que $c_i \neq 0 \forall i = 1,n$

y $\sum_{ii=1}^{m} c_i f_i(x) = 0$ es cierta $\forall x$[7]

[7] Es la terminología que Tou y González utilizan para definir base de "sistema completo"

2.4.2. CONSTRUCCIÓN DE FUNCIONES MULTIVARIADAS

Supongamos $\Phi_1(x), \Phi_2(x), \ldots,$ ortonormales sobre I = [a,b] se puede construir un sistema completo ortonormal de dos variables de la siguiente manera[8]

$$\varphi_1(x_1, x_2) = \Phi_1(x_1)\Phi_1(x_2)$$
$$\varphi_2(x_1, x_2) = \Phi_1(x_1)\Phi_2(x_2)$$
$$\varphi_3(x_1, x_2) = \Phi_2(x_1)\Phi_1(x_2)$$
$$\varphi_4(x_1, x_2) = \Phi_2(x_1)\Phi_2(x_2)$$

2.4.3. SISTEMA DE FUNCIONES ORTOGONALES Y ORTONORMALES

Aquí estudiaremos cómo se construyen estas funciones ortogonales, La motivación para su estudio dentro de la materia de Reconocimiento de Patrones es doble:

1°) son fáciles de generar
2°) estas funciones satisfacen el teorema de aproximación de Weiestrass:

Teorema:
Cualquier función f continua sobre un intervalo cerrado [a,b] puede aproximarse tanto como se quiera por un polinomio de grado n.

Existen varios métodos para generar polinomios: Legendre, Laguerre, Hermite.
Estos métodos proponen una relación de recurrencia y dos polinomios iniciales para generar los demás polinomios
Mencionaremos este teorema.

Teorema[9]

Sea { p_k (x), k=1,2,... } un conjunto de polinomios ortogonales sobre el intervalo[0,1] con relación a p(x) entonces existe una relación lineal entre tres ´polinomios consecutivos, i.e., $\exists$ A, B, C tales que

$$P_{k+1}(x) = (A_k x + B_k) P_k(x) - C_k P_{k-1}(x)$$

[8] Courant y Hilbert, Methods of Mathematical Physics, vol 1, 1955
[9] Aproximation Functions. Rice [ ]

Tenemos tres soluciones:

Laguerre:

$$A_k = \frac{-1}{k+1} \qquad B_k = \frac{2k+\alpha+1}{k+1} \qquad C_k = \frac{k+\alpha}{k+1}$$

Legendre:

$$A_k = \frac{2k+1}{k+1} \qquad B_k = 0 \qquad C_k = \frac{k}{k+1}$$

Hermite:

$$A_k = 2 \qquad B_k = 0 \qquad C_k = 2k \qquad \otimes$$

Entonces tendremos los siguientes polinomios:

a) Polinomios de Legendre

Sean $\{P_i(x), i=0,1,\ldots\}$ los polinomios que pueden derivarse por la siguiente función de recurrencia:

$$(k+1)P_{k+1}(x) - (2k+1)x\,P_k(x) + kP_{k-1}(x) = 0 \qquad k \geq 1 \qquad \text{(2.4-8).}$$

en donde

$$P_0(x)=1 \quad \text{y} \quad P_1(x) = x$$

estas funciones son ortogonales en el intervalo [-1,1].
Con estas condiciones podemos encontrar los cuatro primeros polinomios de Legendre usando la ecuación 2.4-8

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$$

$$P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x \qquad \text{(2.4-9)}$$

$$P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}$$

Estas funciones son ortogonales con respecto a la función de ponderación u(x)=1 entonces, para obtener el sistema ortonormal usaremos 2.7-8 como sigue

$$A_k = \int_{-1}^{1} P_k^2(x)dx$$

y puede demostrarse[10] que es igual a:

$$A_k = \frac{2}{2k+1}$$

Sustituyendo en este resultado en (2.7-7) con $\Phi_k^*(x) = P_k(x)$ los polinomios de Legendre quedan expresados por

$$\Phi_k(x) = \sqrt{\frac{u(x)}{A_k}}\,P_k(x) = \sqrt{\frac{2k+1}{2}}\,P_k(x) \qquad k=0,1,2,\ldots$$

b) Polinomios de Laguerre

Estos polinomios pueden generarse usando la relación recursiva :

$$L_{k+1}(x) - (2k + 1 - x)\,L_k(x) + k^2\,L_{k-1}(x) = 0 \quad k \geq 1 \qquad (2.4\text{-}10)$$

en donde

$L_0(x) = 1$
$L_1(x) = -x+1$

L_0 y L_1 son ortogonales en el intervalo $0 \leq x < \infty$ con respecto a la función de ponderación

$u(x) = e^{-x}$

Los primeros 5 polinomios de Laguerre son:

$L_0(x) = 1$
$L_1(x) = -x + 1$
$L_2(x) = x^2 - 4x + 2$
$L_3(x) = -x^3 + 9x^2 - 18x + 6$
$L_4(x) = x^4 - 16x^3 + 72x^2 - 96x + 24$

[10] Courant y Hilbert [9]

que están determinados por la ecuación 2.4-10. Usando 2.4-8 para determinar A_k con

$$\Phi_k^*(x) = L_k(x)$$

puede demostrarse que los polinomios ortogonales de Laguerre están dados por la relación

$$\Phi_R(x) = \frac{\exp(\frac{x}{2})L_k(x)}{k!} \qquad k = 0,1,2..... \tag{2.4-11}$$

b) Polinomios de Hermite

Relación de recurrencia:

$$H_{k+1}(x) - 2xH_k(x) + 2k\,H_{k-1}(x) = 0\ ,\ k >= 1 \tag{2.4-12}$$

y los polinomios iniciales:

$H_0(x)=1$ $H_1(x)=2x$ $H_2(x)=4x^2-2$ $H_3(x)=8x^3-12x$

$H(x)= 16x^4 -48x^2 + 12$

$$\Phi_k(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2^k k!}\sqrt{\Pi}} H_k(x) \qquad k = 0,1,2, \ldots \tag{2.4-13}$$

Ejemplo 2.5

La construcción de funciones multivariadas usando cualesquiera de estos polinomios, por ejemplo, construir los polinomios de Legendre para tres variables :

$$\varphi_1(x) = \Phi_1(x_1)\Phi_1(x_2)\Phi_1(x_3) = 1$$

$$\varphi_2(x) = \Phi_1(x_1)\Phi_1(x_2)\Phi_2(x_3) = x_3$$

$$\varphi_3(x) = \Phi_1(x_1)\Phi_2(x_2)\Phi_1(x_3) = x_2$$

$$\varphi_4(x) = \Phi_2(x_1)\Phi_1(x_2)\Phi_1(x_3) = x_1$$

$$\varphi_5(x) = \Phi_1(x_1)\Phi_2(x_2)\Phi_{31}(x_3) = x_2x_3$$

en donde $\Phi_1(x) = P_0(x) = 1$

y $\Phi_2(x) = P_1(x) = x$

Por supuesto que existe un número infinito de otras combinaciones con las cuales se pueden construir estas cinco funciones.

Y las funciones de decisión pueden construirse como:

Dado un conjunto $\{\Phi_i(x)\}_{i=1}^{m}$ ortogonal, ortonormal. Podemos expresar la función de decisión $d_1(x)$, $d_2(x)$, ..., $d_M(x)$ como un conjunto de combinaciones lineales de las $\varphi(x)$, esto es,

$$d_i(x) = \sum_{j=1}^{m} w_{ij}\varphi_j(x) \qquad (2.4\text{-}14)$$

$$d_i(x) = \frac{1}{\sqrt{u(x)}} \sum_{j=1}^{m} c_{ij}\varphi_j(x)$$

$$d_i(x) = \sum_{j=1}^{m} c_{ij}\varphi_j(x) \qquad (2.4\text{-}15)$$

2.4.4 EJEMPLO

Ejemplo 2.6

Ejemplos que muestran la obtención de los polinomios.

a) Supongamos que queremos construir los 5 primeros polinomios de Legendre con 3 variables.

Sea el conjunto de polinomios ortogonales:

$\{ P_k(x), k=1,2,... \}$ tal que $P_k \perp P_j \quad k \neq j$

entonces existen A_k B_k y C_k tales que cumplen la relación de recurrencia

$$P_{k+1}(x) = (A_k x + B_k)P_k(x) - C_k P_{k-1}$$

$$A_k = \frac{2k+1}{k+1}P_k(x) \qquad B_k = 0 \qquad y \qquad C_k = \frac{k}{k+1}$$

$$P_{k+1}(x) = \frac{2k+1}{k+1}P_k(x) - \frac{k}{k+1}P_{k-1}(x)$$

$$(k+1)P_{k+1}(x) = (2k+1)P_k(x) - kP_{k-1}(x) \qquad P_i \perp P_j \quad i \neq j \quad y \quad -1 \leq x \leq 1$$

Los polinomios ya fueron expresados en las ecuaciones (2.4-9)
Estas funciones son ortogonales con respecto a u(x)=1, esto es, es un sistema ortogonal

$$A_k = \int_{-1}^{1} P_k^2(x)dx$$

que puede simplificarse a

$$A_k = \frac{2}{2k+1}$$

de donde puede deducirse que

$$A_1 = \frac{1}{3} \qquad A_2 = \frac{2}{5} \qquad A_3 = \frac{2}{7} \qquad A_4 = \frac{2}{9} \qquad A_5 = \frac{2}{11}$$

Entonces, usando (2.4-7) con $\Phi_k^*(x) = P_k(x)$ los polinomios ortogonales de Legendre son:

$$\Phi_0 = \sqrt{\frac{1}{2}}\, P_0 = \sqrt{\frac{1}{2}}$$

$$\Phi_2 = \sqrt{\frac{5}{2}}\, P_2 = \sqrt{\frac{3}{2}}x$$

$$\Phi_3 = \sqrt{\frac{7}{2}}\ P_3 = \sqrt{\frac{7}{2}}\left(\frac{5}{2}x^3 - \frac{1}{2}\right)$$

$$\Phi_4 = \sqrt{\frac{9}{2}}\ P_4 = \sqrt{\frac{9}{2}}\left(\frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}\right)$$

y las

$$\varphi_1(x) = \Phi_1(x_1)\Phi_1(x_2)\Phi_1(x_3) = \left(\sqrt{\frac{3}{2}}\right)^3 x_1 x_2 x_3$$

$$\varphi_2(x) = \Phi_1(x_1)\Phi_1(x_2)\Phi_2(x_3) = \left(\sqrt{\frac{1}{2}}\right)^2\left(\sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2 - \frac{1}{2}\right)\right)$$