

Tema 4: Sistemas Neurodifusos

Universidad Pontificia de Salamanca

Manuel Martín-Merino

Contenido

- Motivación
- Sistemas de Inferencia Difusos Sugeno (TSK) y Tsukamoto
- Sistema Adaptativo de Inferencia Neurodifuso (ANFIS)
 - ▷ Arquitectura
 - ▷ Algoritmo de aprendizaje
 - ▷ ANFIS y redes RBF
 - ▷ Características: Generalización (Aproximador Universal)
- Tópicos avanzados
- Bibliografía

Motivación (I)

- Los sistemas difusos son fácilmente interpretables por los expertos humanos en términos de reglas
- * No permiten adaptarse a cambios en la distribución de probabilidad de los datos
- * La extracción de reglas a partir de expertos humanos es un proceso complejo, caro y, en ocasiones, inviable
- * La lógica difusa no proporciona un formalismo riguroso para la estimación de los parámetros óptimos de las reglas

Motivación (II)

Sistemas Neurodifusos

- Permiten obtener la base de reglas que controla el sistema a partir de una muestra finita de ejemplos de entrada-salida
- Mantienen la interpretabilidad de los sistemas difusos en forma de reglas. Permiten incorporar conocimiento a priori
- Utilizan el formalismo desarrollado en las redes neuronales para estimar los parámetros óptimos de las reglas difusas
- Permiten desarrollar sistemas con alta capacidad de generalización utilizando la teoría desarrollada para las redes neuronales clásicas

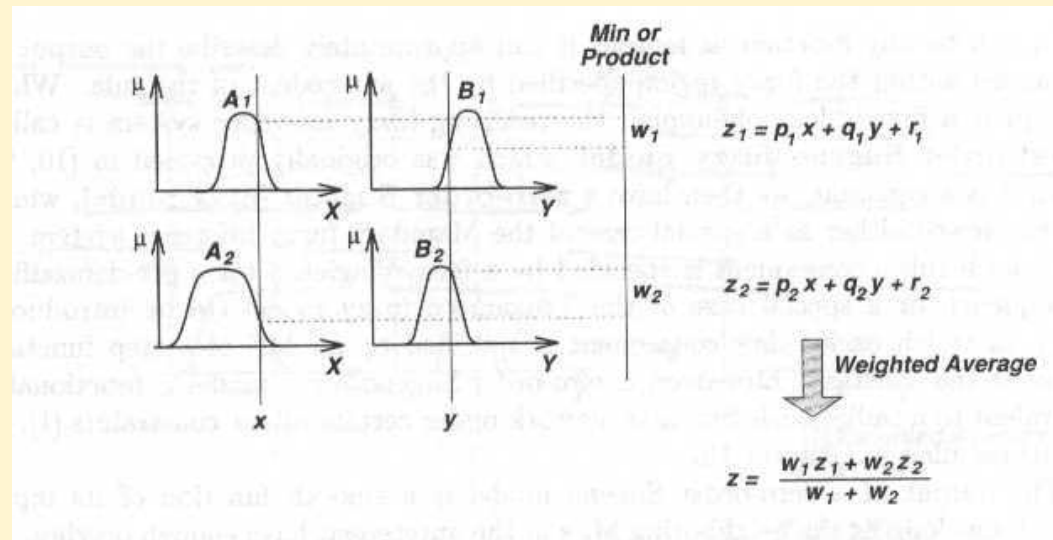
Sistemas de Inferencia Difusos Sugeno (TSK) (I)

- Permite obtener de manera sistemática reglas difusas a partir de datos de entrada salida.
El sistema inferencia Mamdani es excesivamente complejo
- Utiliza reglas del tipo:

$$\textbf{If } x \text{ is } A \text{ and } y \text{ is } B \textbf{ then } z = f(x, y)$$

- $f(x, y)$ especifica la salida en la región difusa determinada por los antecedentes.
 - ▷ Sistema difuso Sugeno primer orden: $f(x, y) = p * x + q * y + r$.
 - ▷ Sistema difuso Sugeno grado 0: $f(x, y) = r$.
Equivale a Mamdani con consecuentes singleton.

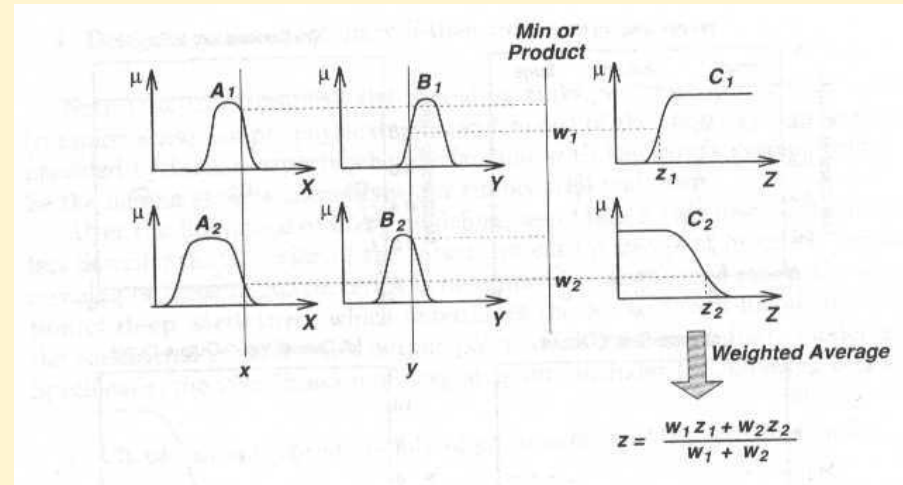
Sistema de Inferencia Difuso Sugeno (TSK) (II)



- Operador mínimo o producto para la t-norma. Defusificador promedio:

$$z = \frac{w_1 z_1 + w_2 z_2}{w_1 + w_2} \quad (1)$$

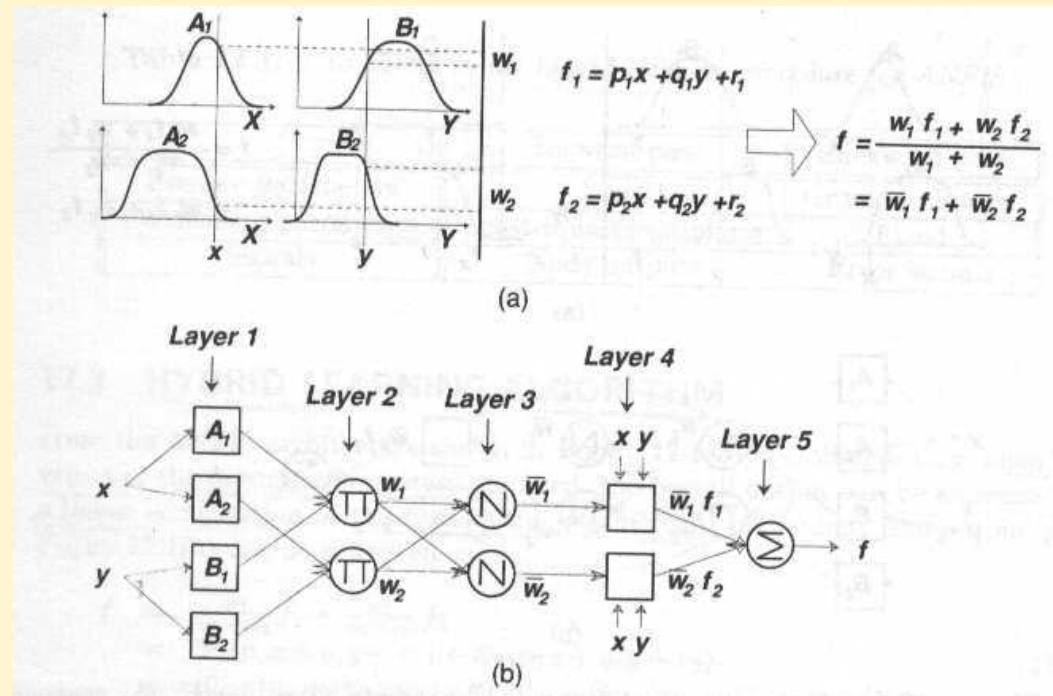
Sistemas de Inferencia Difusos Tsukamoto (I)



- El consecuente de cada regla se representa por una función de pertenencia monótona
- Operador mínimo-producto para la t-norma. Defusificador promedio.
- Menos intuitivo y más complejo

Sistema Adaptativo de Inferencia Neurodifuso (ANFIS)

Arquitectura (I)



- Implementa un sistema difuso Sugeno de grado 1 con reglas del tipo:

if x is A_1 and y is B_1 **then** $f_1 = p_1 * x + q_1 * y + r_1$

Arquitectura (II)

Capas ANFIS

O_{li} denota la salida del nodo i en la capa l

- **Capa 1:** Implementa funciones de pertenencia del tipo:

$$O_{1i} = \mu_{A_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad (2)$$

donde $\{a_i, b_i, c_i\}$ son los parámetros de las premisas.

- **Capa 2:** Realiza la t -norma de las funciones de pertenencia

$$O_{2i} = w_i = \mu_{A_i}(x) \mu_{B_i}(y) \quad \forall i = 1, 2 \quad (3)$$

Arquitectura (III)

- **Capa 3:** Normaliza el grado de activación de cada regla:

$$O_{3i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (4)$$

- **Capa 4:** Multiplica el consecuente de la regla por su grado de activación:

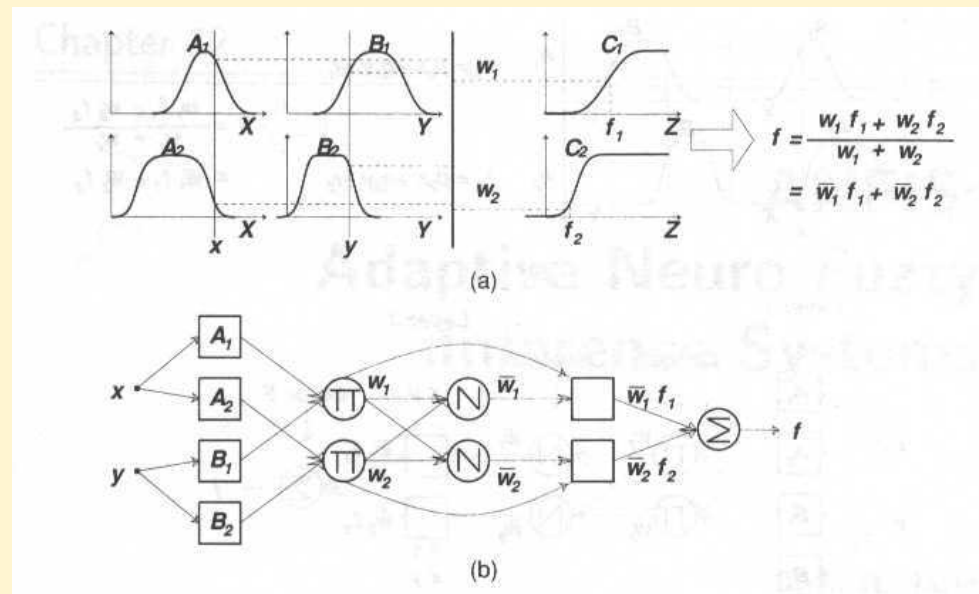
$$O_{4i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (5)$$

- **Capa 5:** Calcula la salida mediante un promedio pesado:

$$O_{51} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (6)$$

Arquitectura (IV)

- La extensión al modelo neurodifuso Tsukamoto ANFIS es inmediato como muestra la figura:



- El sistema de inferencia producto (Mamdani) con defusificador promedio de los centros se puede implementar fácilmente como una red neurodifusa

Algoritmo de aprendizaje (I)

- ANFIS es un sistema lineal en los parámetros del consecuente cuando se fijan los parámetros de las premisas

$$\begin{aligned} f &= \bar{w}_1(p_1x + q_1y + r_1) + \bar{w}_2(p_2x + q_2y + r_2) \\ &= (\bar{w}_1x)p_1 + (\bar{w}_1y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2x)p_2 + (\bar{w}_2y)q_2 + (\bar{w}_2)r_2 \quad (7) \end{aligned}$$

- Sea $S = S_1 S_2$ donde
 - ▷ S_1 es el conjunto de parámetros de las premisas
 - ▷ S_2 es el conjunto de parámetros del consecuente. El sistema es lineal en S_2

Algoritmo de aprendizaje (II)

El algoritmo de aprendizaje se puede implementar en dos pasos:

- Paso hacia adelante:

- ▶ Se obtienen las salidas de los nodos de la capa 4, \bar{w}_1, \bar{w}_2). El vector de parámetros del consecuente verifica:

$$\mathbf{A}\theta = \mathbf{f} \quad (8)$$

- ▶ La solución θ^* minimiza $\|\mathbf{A}\theta - \mathbf{f}\|^2$

$$\theta^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{f} \quad (9)$$

donde $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ es la pseudoinversa de \mathbf{A} .

Algoritmo de aprendizaje (III)

- Paso hacia atrás:
Los errores se propagan hacia atrás actualizando con un algoritmo iterativo basado en gradiente descendente los parámetros de las funciones de pertenencia del antecedente.
- La estrategia anterior converge mucho más rápido que gradiente descendente puro. El espacio de parámetros para el gradiente descendente se reduce disminuyendo la dimensión del espacio de búsqueda.

ANFIS y redes RBF

- Las redes RBF son equivalentes al modelo de Inferencia Difuso de Sugeno grado 0. (Se deja como ejercicio)
- El sistema híbrido de aprendizaje de ANFIS puede ser utilizado para estimar los parámetros de la red RBF
- El sistema de optimización de las redes RBF puede ser empleado para estimar los parámetros de ANFIS.