Modele rozmyte[1]

Cel tworzenia nowych modeli: dążenie do uzyskania coraz większej dokładności, wymiarowości lub uproszczenia struktury.

Model Mamdaniego

Mamdani (1974) zastosował ideę rozmytego modelu do sterowania.

- zbiór reguł, z których każda definiuje jeden rozmyty punkt w przestrzeni iloczynu kartezjańskiego XxY
- zbiór punktów tworzy rozmyty wykres

Model Mamdaniego jest reprezentowany przez zbiór reguł o postaci:

R1: JEŻELI (x jest A1) to (y jest B1)
R2: JEŻELI (x jest A2) to (y jest B2)
…

Modelowanie za pomocą charakterystycznych „ważnych” punktów modelu

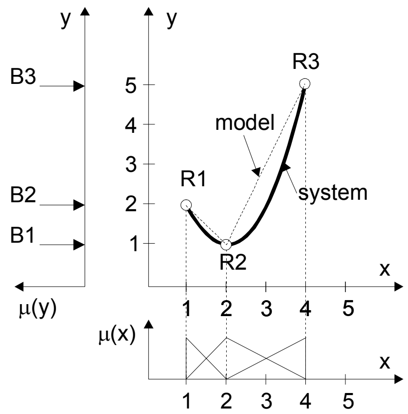
- usytuowane na charakterystyce modelu
- inne usytuowanie

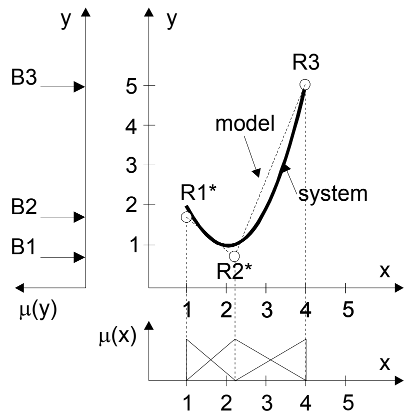
[1] Wykład opracowano na podstawie podręcznika Andrzeja Piegata „Modelowanie i sterowanie rozmyte”, Akademicka Oficyna Wydawnicza EXIT, Warszawa 1999.

Np.: y=(x-2)^2 +1

R1: JEŻELI (x jest A1) to (y jest B1)
R2: JEŻELI (x jest A2) to (y jest B2)
R3: JEŻELI (x jest A3) to (y jest B3)

Gdzie
A1 = około 1, A2 = około 2, A3 = około 4
B1 = około 2, B2 = około 1, B3 = około 5
x należy do przedziału (1,4)

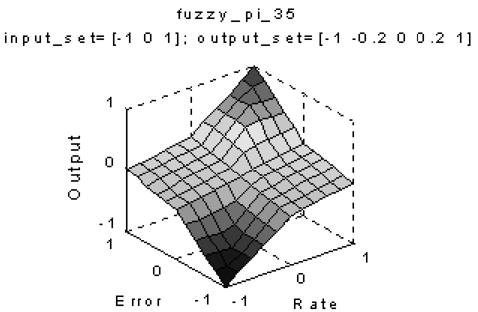


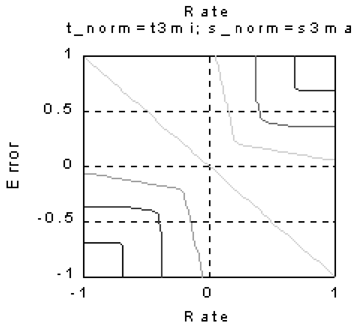


Lingwistyczne i nielingwistyczne modele Mamadaniego

- lingwistyczne gdy mała ilość etykiet lingwistycznych np. mały, sredni, duży.
- nielingwistyczne, gdy większa liczba np. 0-9, wtedy około 0, itd.

Powierzchnia systemu





Model Takagi-Sugeno

Różnii się od modeli Mamdaniego postacią reguł.

JEŻELI (x jest A1) to (y=f(x))

Czyli konkluzja zawiera nie zbiór rozmyty, a funkcję (liniową lub nieliniową).

Baza reguł ma postać

R1: JEŻELI (x jest A1) to (y = f1(x))
R2: JEŻELI (x jest A2) to (y = f2(x))

(…)

Wyjście modelu oblicza się na podstawie stopnia aktywacji poszczególnych konkluzji

$$y = \frac{\sum_{i=1}^{m} \mu_A(x) f_i(x)}{\sum_{i=1}^{m} \mu_A(x)}$$

Przykład:

Dany jest model TS z bazą reguł

R1: JEŻELI (x jest A1) TO (y = -x+3)
R2: JEŻELI (x jest A2) TO (y = (4x-10)/3)
R3: JEŻELI (x jest A3) TO (y = (-x+24)/3)

$$w_1 = \begin{cases} 1 & 0 \le x \le 2 \\ 0 & \end{cases} \qquad w_4 = \begin{cases} 1 & 7 \le x \le 9 \\ 0 & \end{cases}$$

$$w_2 = \begin{cases} 1 & 2 \le x \le 4 \\ 0 & \end{cases} \qquad w_5 = \begin{cases} 1 & 9 \le x \le 12 \\ 0 & \end{cases}$$

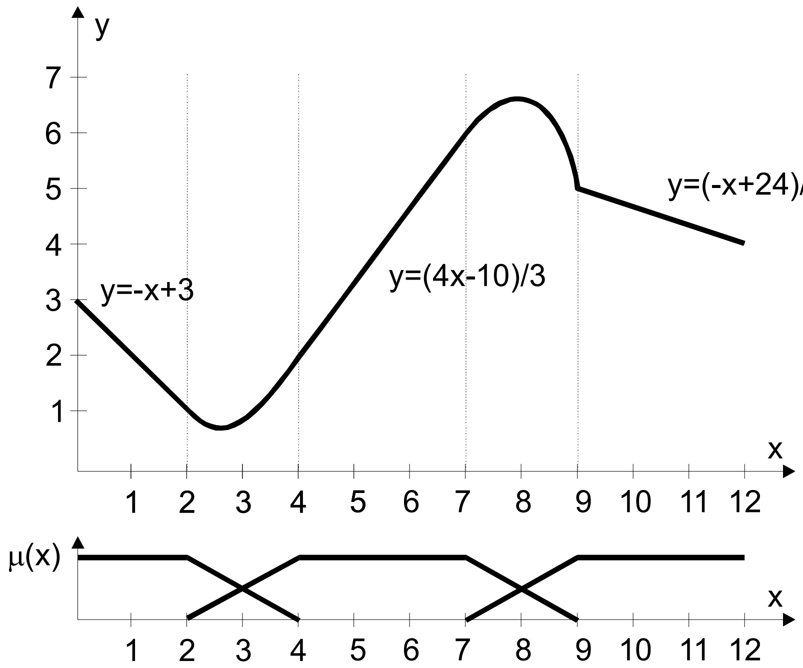
$$w_3 = \begin{cases} 1 & 4 \le x \le 7 \\ 0 & \end{cases}$$

Funkcje przynależności mają postać

$$\mu_{A_1}(x) = w_1 - 0.5(x-4)w_2$$

$$\mu_{A_2}(x) = 0.5(x-2)w_2 + w_3 - 0.5(x-9)w_4$$

$$\mu_{A_3}(x) = 0.5(x-7)w_4 + w_5$$



- powierzchnia systemu odpowiada dokładnie konkluzjom reguł tylko wtedy gdy, $\mu_A(x) = 1$, w pozostałych przypadkach jest przejście od jednej formy liniowej do innej
- trapezowe funkcje przynależności

Kryteria podobieństwa systemu i jego modelu

1. Kryterium podobnego odwzorowania wektora wejść X_S na wyjście y

Przy identycznych wektorach wejściowych systemu i modelu wyjście modelu y_M powinno być możliwie bliskie wyjścia systemu y_S dla całej dziedziny

2. Kryterium podobnego odwzorowania zmiany ΔX_S wektora wejściowego na zmianę wyjścia Δy

Model i system są podobne wówczas, gdy zmiana ΔX_S zostaje odwzorowana na podobne sobie zmiany Δy_S i Δy_M

Stosowanie operatorów MIN do agregacji przesłanek składowych lub przeprowadzanie defuzyfikacji metodą największego maksimum wprowadza do modelu strefy nieczułości, w których nie reaguje on na zmiany wejść (kryteria (2) i (1)).

Metody modelowania rozmytego

1. Tworzenie modeli rozmytych na bazie wiedzy eksperta
2. Tworzenie samonastrajających się modeli rozmytych na bazie danych pomiarowych we/wy systemu
3. Tworzenie samoorganizujących i samonastrajających się modeli rozmytych na bazie danych pomiarowych

ad.1. (na bazie wiedzy eksperta)

- pierwszy rodzaj modelowania rozmytego zastosowany w praktyce
- bazuje na wiedzy i doświadczeniu eksperta systemu (doskonale znającego system i wszystkie jego zachowania)

„Jeżeli pedał gazu ***mocno*** wciśnięty to ***duże*** przyspieszenie”

wiedza eksperta – co to znaczy ***mocno*** i ***duże***

Problemy:

- taki model nie zawiera niejawnej wiedzy o systemie (wyczucie systemu, intuicja)
- ekspert nie jest w stanie przekazać wiedzy o mechanizmach wnioskowania zachodzących w jego umyśle, o rodzaju funkcji przynależności, o rodzaju operatorów logicznych

Uwagi:

- modele tego samego systemu dla różnych ekspertów mogą się różnić
- model werbalny może być pierwszym etapem, dalej strojenie
- precyzyjne modele werbalne daje się skonstruować w przypadku systemów **mechanicznych** i **elektrycznych**, mniej dla **termicznych** i **chemicznych,** a najmniej dla **ekonomicznych** i **socjologicznych**
- model werbalny na bazie wiedzy eksperta umożliwia tworzenie tylko modelu Mamdaniego

Modele samonastrajające się na bazie danych pomiarowych

Struktura systemu znana (stała baza reguł i zbiorów rozmytych) i nie podlega zmianom

Strojenie modelu

– określenie takich parametrów f. przynależności, które minimalizują błąd modelu wzgl. modelowanego systemu
– zmiana rodzaju operatorów agregacji, metod wnioskowania i metody defuzyfikacji, rodzaju f. przynależności (odcinkowo-liniowe, gaussowskie)

Metody strojenia (optymalizacji):

1. rozmyte sieci neuronowe – przekształcenie systemu rozmytego w sieć neuronową
2. algorytmy genetyczne – poszukiwanie optymalnych parametrów modelu rozmytego
3. metody oparte na klasteryzacji
4. metody heurystyczne

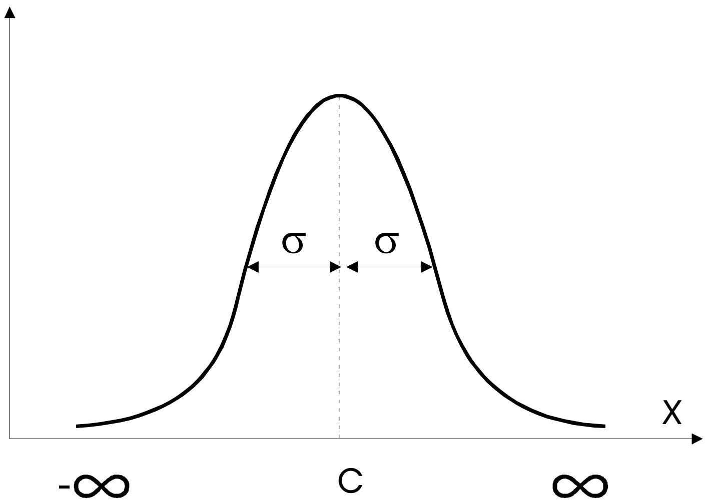
Rozmyte sieci neuronowe
strojenie parametrów systemu rozmytego

Sieci RBF – neurony z radialnymi funkcjami bazowymi aktywacji

$$y = f(x) = f(\|x - c\|)$$

c – wektor współrzędnych centrum funkcji

Gaussowska funkcja RBF – dobrze się nadaje do modelowania reguł rozmytych



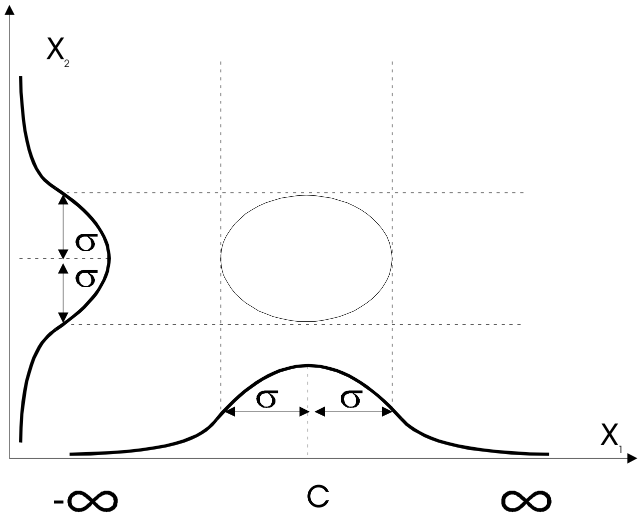
JEŻELI (x jest blisko c) TO (y jest blisko y_0)

$$y = f(x) = y_0 e^{-\frac{(x-c)^2}{2\sigma^2}}$$

Dla układów wielowejściowych:

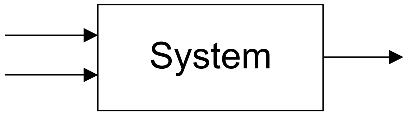
JEŻELI (x_1 jest blisko c_1) I (x_2 jest blisko c_2) TO (y jest blisko y_0)

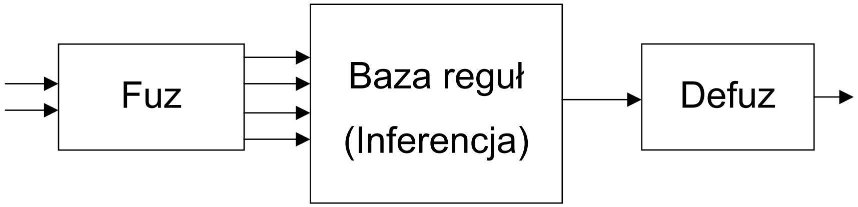
$$y = f(x) = y_0 e^{-\sum_{i=1}^{p} \frac{(x_i - c_i)^2}{2\sigma_i^2}}$$



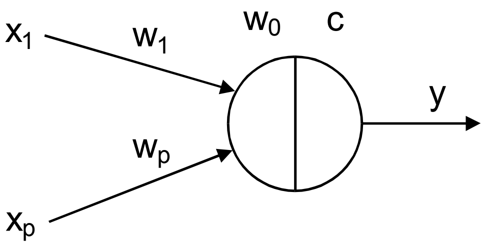
Czy można przedstawić system rozmyty w postaci sieci neuronowej?

Przekształcenie rozmytego modelu Mamdaniego w sieć neuronową





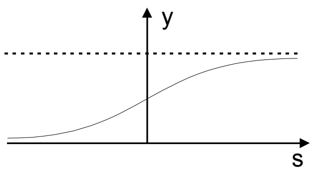
Schemat sztucznego neuronu



Funkcja propagacji

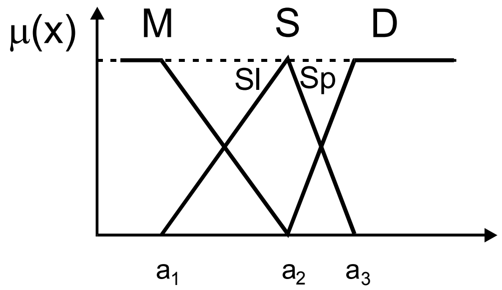
$$s = w_0 + \sum_{i=1}^{p} w_i x_i$$

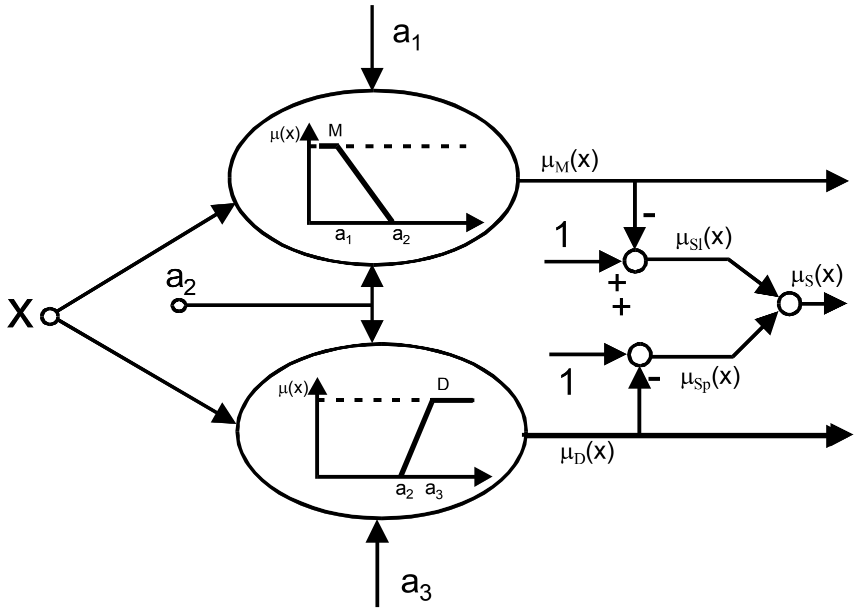
Funkcja aktywacji



Przekształcenie bloku FUZYFIKACJA

W trakcie nauczania sieci strojeniu podlegają parametry a_i funkcji przynależności. Stąd konieczne jest określenie pochodnych wyjść bloku względem tych parametrów.





$$\mu_M(x) = \begin{cases} 1 & \text{gdy } x < a_1 \\ \dfrac{a_2 - x}{a_2 - a_1} & \text{gdy } a_1 \le x < a_2 \\ 0 & \text{gdy } x \ge a_1 \end{cases}$$

$$\frac{\partial \mu_M(x)}{\partial a_1} = \begin{cases} \dfrac{a_2 - x}{(a_2 - a_1)^2} & \text{gdy } a_1 \le x < a_2 \\ 0 & \text{poza tym} \end{cases}$$

$$\frac{\partial \mu_M(x)}{\partial a_2} = \begin{cases} \dfrac{x - a_1}{(a_2 - a_1)^2} & \text{gdy } a_1 \le x < a_2 \\ 0 & \text{poza tym} \end{cases}$$

itd.

Sieci z neuronami GRBF są różniczkowalne w sposób ciągły, ale wprowadzają błąd transformacji (trójkąt to nie kapelusz).

Modele rozmyte wykorzystują często niesymetryczne funkcje trójkątne, trzeba by zatem wprowadzić niesymetryczne funkcje GRBF, a to komplikuje problem.

Przekształcenie elementów bloku Baza Reguł

Wyjście bloku fuzyfikacja - stopnie przynależności poszczególnych wejść do zbiorów rozmytych.

W bloku baza reguł oblicza się stopień spełnienia całej przesłanki złożonej, który aktywizuje funkcję przynależności w konkluzji.

Tu operacje I oraz LUB wykonywane są za pomocą t-norm i s-norm.

IF ($x_1 = A_{11}$) AND ($x_2 = A_{22}$) …
OR ($x_1 = A_{12}$) AND ($x_2 = A_{21}$) … THEN ($y = B_l$)

Kłopot z MAX i MIN oraz negacją (są nieróżniczkowalne?)

Wyjście neuronu MAX reaguje zmianą Δy na zmianę Δx_i wejścia x_i tylko wówczas, gdy $x_i = \max(x_1, \ldots, x_p)$

$$y=\min(x_1, \ldots, x_p)$$
$$y=\max(x_1, \ldots, x_p)$$

$$\frac{\partial y}{\partial x_i} = \begin{cases} 1 & y = x_i \\ 0 & y \neq x_i \end{cases}$$

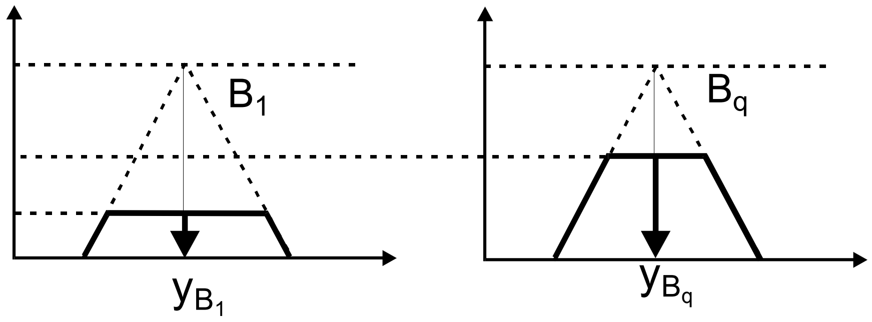
negacja

$$B = \overline{A}, \quad \mu_B = 1 - \mu_A, \quad \frac{\partial \mu_B}{\partial x_A} = -1$$

Przekształcenie bloku defuzyfikacja

np. defuzyfikacja singletonowa – zastąpienie zbiorów wyjścia wartością rzeczywistą (w najwyższym punkcie lub środku ciężkości)

$$y = \frac{\sum_{l=1}^{q} \mu_{Bl} y_{Bl}}{\sum_{l=1}^{q} \mu_{Bl}}$$



Metody strojenia

1. Strojenie wyłącznie parametrów warstwy defuzyfikacyjnej (szybkie)
2. Jednoczesne strojenie wszystkich parametrów całej sieci (wolniejsze, najlepsze rezultaty)
3. Przemienne strojenie parametrów (szybsze, do nauczania można stosować np. alg. gen.)

Zalety rozmytych sieci neuronowych

- umożliwia optymalizację (strojenie) parametrów modelu rozmytego na bazie danych pomiarowych we/wy systemu rzeczywistego
- umożliwia korektę niedokładnych modeli rozmytych sformułowanych przez ekspertów
- umożliwia uzupełnienie modeli rozmytych sformułowanych przez ekspertów w tych rejonach przestrzeni wejść, w których brak jest wiedzy eksperckiej
- posiada strukturę i parametry zrozumiałe dla człowieka (wyrażenia lingwistyczne i reguły). Umożliwia to uogólnienie wiedzy zawartej w zaszumionych danych pomiarowych i przedstawienia jej w zrozumiałej dla człowieka formie słownych reguł (ekstrakcja wiedzy)
- można łatwo wprowadzić wstępną lub częściową wiedzę o modelowanym systemie
- można łatwo trafnie określić strukturę sieci i ilość neuronów w każdej z warstw