DESPRE FUZZIFICARE

ing. Aurelia Cosmescu

Institutul de Cercetări în Informatică

Rezumat: În lucrare sunt prezentate două metode de fuzzificare a funcțiilor continue, folosind principiul extensiei și rețelele neuronale fuzzy.

Cuvinte cheie: fuzzificare, regulator fuzzy, control fuzzy, principiul extensiei, rețele neuronale fuzzy.

1. Introducere

Scopul controlului este acela de a influența comportarea unui sistem prin modificarea intrărilor acestuia, în conformitate cu o regulă sau cu un set de reguli care modelează felul în care operează sistemul.

Teoria clasică a controlului folosește modelul matematic, cel mai cunoscut exemplu de model de control fiind controlerul PID. El compară ieșirea sistemului cu starea dorită a acestuia și, pe baza diferenței notată cu *e*, dintre aceste două mărimi, ajustează valoarea de intrare notată cu *u*, conform relației:

$$u = A * e + B * INT(e) + C * de/dt$$

unde A, B, C sunt constante, INT(e) reprezintă integrala erorii, iar de/dt reprezintă derivata erorii.

Dezavantajul este că, în acest caz, se presupune că sistemul ce trebuie modelat este liniar sau cel puțin se comportă ca o funcție monotonă. De asemenea, dacă crește complexitatea sistemului, atunci este dificil de determinat modelul matematic.

Soluția ar putea fi controlul fuzzy, în care modelul fuzzy (reguli) ia locul modelului matematic. Modelarea unei strategii de control cu reguli fuzzy este foarte avantajoasă în cazul sistemelor neliniare deoarece, spre deosebire de modelele matematice, modelele fuzzy nu devin mai complexe din cauza neliniarității.

În ultimii ani, controlul fuzzy a devenit una din cele mai active arii de interes ale cercetării pentru aplicarea teoriei mulțimilor fuzzy. Ideea de bază a controlului fuzzy este simularea comportării unui expert uman, care se află în situația de a specifica cele mai importante proprietăți ale procesului pe care vrem să-l controlăm. Specificarea se face prin reguli lingvistice, ca de exemplu:

IF a este mediu și b este mic THEN c este mare.

Termenii "mediu", "mic" și "mare" sunt reprezentați ca mulțimi fuzzy și constituie descrieri ale variabilelor de intrare, respectiv ale variabilei de ieșire. Aceste reguli lingvistice de control sunt interpretate prin relații fuzzy și fiecare dintre ele specifică o legătură între valori vagi de intrare și valori vagi de ieșire.

Există câteva probleme și în cazul sistemelor fuzzy, și anume:

- este dificil să proiectăm un sistem fuzzy optimal; această problemă apare în special în cazul în care sistemul fuzzy se aplică proceselor complexe;
- este destul de dificil să determinăm mulțimile fuzzy ce descriu variabilele; în acest caz o rețea neuronală fuzzy poate constitui o metodă pentru rezolvarea acestei probleme;
- expresia lingvistică a sistemelor fuzzy face dificilă garantarea stabilității și a robusteții sistemului.

2. Regulatoare fuzzy

În general, un regulator este o funcție $f : X \rightarrow Y$. În cazul unui regulator fuzzy, transformările care au loc sunt următoarele:

$$X \xrightarrow{\text{fuzzificare}} \mathbf{F}(X) \xrightarrow{\text{inferență}} \mathbf{F}(Y) \xrightarrow{\text{defuzzificare}} Y$$

Problema este să extindem funcția f la o funcție

$$\tilde{f} : \mathbf{F}(X) \rightarrow \mathbf{F}(Y)$$

(prin această extindere se înțelege **fuzzificarea funcției f**).

Așadar, părțile componente ale unui regulator fuzzy sunt următoarele (vezi figura 1):

- fuzzificare
- inferență
- defuzzificare

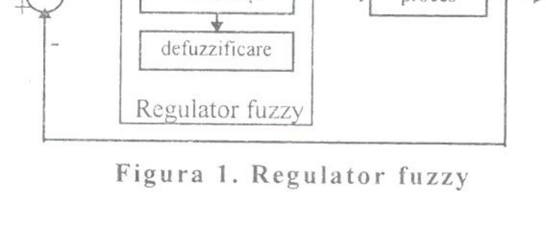


Figura 1. Regulator fuzzy

Fuzzificarea reprezintă translatarea variabilelor numerice de intrare în variabile lingvistice și determinarea gradelor de apartenență corespunzătoare.

Fiind dat un univers U, vom considera mulțimea $FP(U)$ a tuturor mulțimilor fuzzy pe U. Fuzzificare înseamnă definirea unei transformări φ care asignează unui număr real $x \in U$, o mulțime fuzzy $\varphi(x)$ pe U, numită fuzzificare a lui x.

Inferența definește o funcție Φ, care transformă mulțimea fuzzy de intrare F, într-o mulțime fuzzy de ieșire $G = \Phi(F)$. În cadrul acestui bloc se utilizează o bază de reguli IF-THEN și un mecanism de inferență.

Metode de inferență:

- MAX-PROD
- MAX-MIN

Defuzzificarea reprezintă calcularea unei valori numerice ca ieșire a controlerului, bazată pe reguli simbolice.

Defuzzificarea definește o transformare δ în felul următor $\delta : FP(U) \to U$.

Metode de defuzzificare:

- metoda centrului de greutate
- metoda maximului

3. Proceduri de fuzzificare

Definiția 1: O mulțime fuzzy μ a lui X este o funcție care transformă X în intervalul unitate, adică $\mu : X \to [0,1]$.

Valoarea $\mu(x)$ reprezintă *gradul de apartenență* al lui x la mulțimea fuzzy μ.

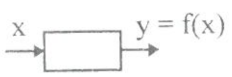
Definiția 2: O valoare numerică x_0 se transformă într-o mulțime fuzzy $\mu(x_0)$ (numită *singleton*) astfel

$\mu(x_0) = 1$, dacă $x = x_0$

$\mu(x_0) = 0$, în rest.

Se consideră o funcție $f : X \to Y$. În acest caz, există două interpretări:

1. "cutie neagră"



2. Fiind dat y, se pune întrebarea dacă există x, astfel încât $f(x) = y$.

În continuare, se pune problema extensiei funcției $f : X \to Y$ la o funcție $\tilde{f} : \boldsymbol{F}(X) \to \boldsymbol{F}(Y)$. Se va încerca rezolvarea acestei probleme prin două metode:

- folosind principiul extensiei
- folosind rețelele neuronale fuzzy (FNN)

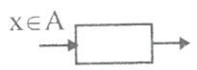
3.1 Principiul extensiei

Se consideră, mai întâi, un exemplu clasic și anume extinderea funcției $f : X \to Y$ la o funcție $\bar{f} : \boldsymbol{P}(X) \to \boldsymbol{P}(Y)$.

Se spune că o mulțime $A \in \boldsymbol{P}(X) \Leftrightarrow A \subset X$.

Atunci, $\bar{f}(A) = \{f(x);\ x \in A\} \subset Y$.

Acest exemplu se interpretează ca un exemplu de nedeterminare a intrării. Nu se cunoaște x (dar se știe că $x \in A$) și atunci avem o nedeterminare care se reflectă într-o nedeterminare asupra ieșirii.



Funcția $\bar{f}$ de mai sus are următoarele proprietăți:

- este monotonă,

adică dacă $A \subset B \Rightarrow \bar{f}(A) \subset \bar{f}(B)$;

- comută cu reuniunea,

adică $\bar{f}(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} \bar{f}(A_i)$.

Așadar, $\bar{f}$ extinde f la mulțimea părților unei mulțimi, în sensul că face că diagrama de mai jos să fie comutativă :

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow & & \downarrow \\ \mathbf{P}(X) & \xrightarrow{\bar{f}} & \mathbf{P}(Y) \end{array}$$

Precizare: Funcția $\bar{f}$ este unic determinată de f.

Să luăm acum un alt exemplu și anume:

$$X \xrightarrow{g} \mathbf{P}(Y)$$

Aceasta înseamnă o relație $R \subset X \times Y$.

Problema: Cum extindem $\mathbf{P}(X) \xrightarrow{\bar{g}} \mathbf{P}(Y)$?

$\overline{g}(A) = \bigcup_{x \in A} g(x)$ (funcția $\overline{g}$ comută cu reuniunea).

Revenind la problema în discuție, se va extinde prin analogie

$$X \xrightarrow{f} Y$$

la

$$\mathbf{F}(X) \xrightarrow{\tilde{f}} \mathbf{F}(Y)$$

În acest caz, extensia se va face astfel:

$$\tilde{f}(A)(y) = \sup_{x \in f^{-1}(y)} A(x), \text{ dacă } f^{-1}(y) \neq 0. \tag{1}$$

$$\tilde{f}(A)(y) = 0, \text{ în rest.}$$

unde $A : X \to [0, 1]$.

Funcția $\tilde{f}$ comută cu sup-ul (reuniunea în cazul fuzzy), dar această extensie nu este singura care comută cu sup-ul.

Cele de mai sus se pot formula după cum urmează.

O teorie fuzzy este T = (T, e, $(\_)^{\#}$) [5] unde:

- T atașează fiecărei mulțimi X o mulțime T(X);
- e atașează fiecărei mulțimi X o funcție $e_X : X \to T(X)$;
- $(\_)^{\#}$ atașează fiecărei funcții $\alpha : X \to T(Y)$ o funcție $\alpha^{\#} : T(X) \to T(Y)$.

Avem următoarele trei axiome:

1. axioma extensiei $\alpha^{\#} e_X = \alpha$
2. axioma post-identitate $(e_X)^{\#} = id_{T(X)}$
axioma de asociativitate $(\beta^{\#}\alpha)^{\#} = \beta^{\#}\alpha^{\#}$.

Se vor ilustra grafic cele trei axiome:

1.

$$\begin{tikzcd} X \arrow[r, "e_X"] \arrow[dr, "\alpha"'] & T(X) \arrow[d, "\alpha^{\#}"] \\ & T(Y) \end{tikzcd}$$

2.

$$X \xrightarrow{e_X} T(X)$$

$$T(X) \xrightarrow{e_X^{\#}} T(X)$$

3.

$$X \xrightarrow{\alpha} T(Y) \xrightarrow{\beta^{\#}} T(Z)$$

$$T(X) \xrightarrow{\alpha^{\#}} T(Y) \xrightarrow{\beta^{\#}} T(Z)$$

Acesta reprezintă un cadru matematic pentru condițiile necesare pe care trebuie să le îndeplinească o extensie. În cazul nostru, T este **F**, este singleton-ul din definiția 2, iar # se calculează cu formula (1). Se observă că, în acest, caz axiomele sunt îndeplinite.

3.2 Rețele neuronale fuzzy (FNN)

Se va schița tipul de rețea neuronală de baza de care avem nevoie pentru a defini o rețea neuronală fuzzy. Rețeaua este stratificată, feedforward, cu un neuron de intrare, *m* neuroni ascunși și un neuron de ieșire (3 straturi). Toate semnalele, ponderile și pragurile sunt numere reale. Fie x semnalul de intrare în neuronul de intrare. Neuronul de intrare nu schimbă acest semnal, dar îl distribuie tuturor neuronilor din stratul ascuns. Ponderea de la neuronul de intrare la neuronul i din stratul ascuns este w_i, iar $w_i x$ este intrarea în neuronul i al stratului ascuns. Toți neuronii din stratul ascuns au aceeași funcție de transfer $g(x) = 1/(1+e^{-x})$, astfel că ieșirea neuronului i va fi $g(w_i x + \theta_i)$ unde θ_i reprezintă pragul.

Ieșirea rețelei va fi:

$$o = \sum_{i=1}^{m} v_i g (w_i x + \theta_i) \tag{2}$$

unde v_i este ponderea conexiunii de la neuronul i din stratul ascuns, la neuronul de ieșire.

Se poate scrie o = NN(x) ca fiind forma ieșirii rețelei neuronale. Se știe că acest tip de rețea neuronală este aproximatorul universal.

Aceasta înseamnă că, fiind dată o funcție continuă h: [-M, M] → reale(M>0) și $\varepsilon > 0$, există o rețea neuronală NN(m *și* w_i, θ_i *și* v_i) astfel încât:

$$\max\{ | h(x) - NN(x) | : -M \leq x \leq M \} < \varepsilon \tag{3}$$

O rețea neuronală fuzzy (FNN) are aceeași structură de bază ca și rețeaua neuronală descrisă mai sus, excepție facând semnalele și/sau ponderile și pragurile care pot fi fuzzy.

Se disting trei tipuri de bază de FNN [2]:

- FNN1 are intrarea=număr real, dar ponderile şi/sau pragurile pot fi fuzzy;
- FNN2 are intrarea = fuzzy, dar ponderile şi pragurile sunt reale;
- FNN3 are intrarea = fuzzy şi ponderile şi pragurile tot fuzzy.

În continuare, se va considera o FNN3 cu semnalul de intrare $\overline{X}$, ponderile $\overline{W}_i$, $\overline{V}_i$ şi pragurile $\overline{\theta}_i$.

Arhitectura unei astfel de reţele este prezentată în figura 2.

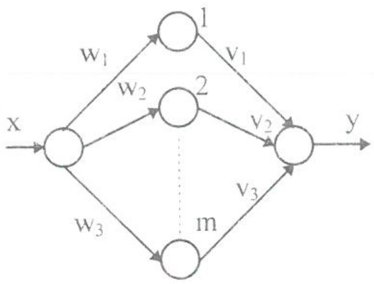


Figura 2. Arhitectura FNN3

Ecuaţia (2) devine:

$$\overline{O} = \sum_{i=1}^{m} \overline{V}_i g(\overline{W}_i \overline{X} + \overline{\Theta}_i) \qquad (4)$$

Dacă $\overline{W}_i$, $\overline{X}$, $\overline{\Theta}_i$ şi $\overline{V}_i$ sunt toate numere fuzzy, atunci $\overline{O}$ va fi număr fuzzy.

Pentru simplitate, se folosesc numere fuzzy triunghiulare simetrice. Un număr fuzzy triunghiular simetric $\overline{N}$ este specificat prin două numere reale a şi b, a < b, unde funcţia de transfer pentru $\overline{N}$ este un triunghi cu baza pe [a,b] şi vertex-ul (gradul de apartenenţă = 1) la 1/2 (a+b). Prin $\overline{N}$ vom indica (a,b).

Şi acum, se va prezenta metoda de fuzzificare a unei funcţii continue y=f(x), pentru $x \in [-M,M]$, analizând următoarele trei cazuri [3]:

1. x este fuzzy;
2. f este fuzzy;
3. f şi x sunt fuzzy.

Cazul 1

Deoarece NN este aproximatorul universal, fiind dat $\varepsilon=10^{-10}$, există NN*(m, ponderi, prag)* astfel încât:

$$\max\{ | f(x) - NN(x) | : -M \leq x \leq M \} < \varepsilon \qquad (5)$$

pentru M > 0 dat.

Intrarea $\overline{X}=(x_1,x_2)$ pe [-M,M] în această reţea produce o FNN2. Ieşirea $\overline{O}$ diferă de ieşirea obţinută cu ajutorul principiului extensiei.

Cazul 2

Se foloseşte aceeaşi NN ca şi în cazul 1, care aproximează y = f(x) uniform pe [-M,M]. De data aceasta însă, se vor fuzzifica ponderile şi pragurile într-o reţea neuronală pentru a obţine FNN1. Se presupune că există *m* neuroni ascunşi, ponderile sunt w_i, v_i, pentru $1 \leq i \leq m$, iar pragurile sunt θ_i, pentru $1 \leq i \leq m$. Pentru $\lambda > 0$ definim:

(1) $\overline{W}_i = (w_i - \lambda, w_i + \lambda)$ pentru $1 \leq i \leq m$;

(2) $\overline{V}_i = (v_i - \lambda, v_i + \lambda)$ pentru $1 \leq i \leq m$;

(3) $\overline{\Theta}_i = (\theta_i - \lambda, \theta_i + \lambda)$ pentru $1 \leq i \leq m$;

Avem acum o FNN1 cu ponderile $\overline{W}_i$, $\overline{V}_i$, pentru $1 \leq i \leq m$ şi pragurile $\overline{\Theta}_i$, $1 \leq i \leq m$. Dacă intrarea este x din [-M,M], atunci ieşirea va fi $\overline{O}$.

Cazul 3

Se va folosi aceeaşi reţea neuronală ca şi în cazul 1 şi se vor fuzzifica ponderile şi pragurile la fel ca în cazul 2. Intrarea reţelei va fi $\overline{X} = (x_1, x_2)$ din [-M,M], care va produce FNN3, iar ieşirea va fi $\overline{O}$.

4. Exemplu

Metodele de fuzzificare prezentate mai sus vor fi ilustrate printr-un exemplu, şi anume $y = x^2$.

Se va arăta că toate metodele de calcul pentru $\overline{Y}=(\overline{X})^2$ cu $\overline{X}$ = număr fuzzy triunghiular simetric în [-1,1] vor produce rezultate diferite.

Mai întâi, se consideră *principiul extensiei* ca metoda de a obţine $\overline{Y}$, atunci când $\overline{X} = (x_1,x_2)$, pentru $-1 \leq x_1 \leq 0 \leq x_2 \leq 1$ ($\overline{X}$ va fi număr fuzzy triunghiular simetric, care îl are pe 0 în suportul său). Se vor obţine:

$y_1 = 0$

$y_2 = \max\{x_1^2, x_2^2\}$,

unde $[y_1, y_2]$ este suportul lui Y.

În continuare, se consideră o *reţea neuronală* folosită pentru a aproxima $y = x^2$ uniform pe intervalul [-1,1]. Vom avea ca intrare $\overline{X} = (x_1, x_2)$, iar ca ieşire $\overline{O} = (o_1, o_2)$. Pentru FNN2 avem:

$$o_1 = G(x_1;\Gamma_1) + H(x_2; \Gamma_2) \qquad (a1)$$

$$o_2 = G(x_2;\Gamma_1) + H(x_1; \Gamma_2) \qquad (a2)$$

unde

$$\Gamma_1 = \{i |\ w_i \geq 0,\ v_i \geq 0 \text{ sau } w_i < 0,\ v_i < 0\} \qquad (a3)$$

$$\Gamma_2 = \{i |\ w_i < 0,\ v_i \geq 0 \text{ sau } w_i \geq 0,\ v_i < 0 \qquad (a4)$$

$G(x;\Gamma)=H(x;\Gamma)=\{\sum v_i g(w_i x+\theta_i) | i\in\Gamma\}$ (a5)

Să vedem acum dacă putem obține $[o_1, o_2]$= suportul lui $\bar{Y}$ rezultat prin principiul extensiei. În acest caz trebuie să avem:

$$0 = G(x_1;\Gamma_1) + H(x_2; \Gamma_2) \quad (a6)$$

$$\max\{x_1^2, x_2^2\} = G(x_2;\Gamma_1) + H(x_1; \Gamma_2 \quad (a7)$$

Avem două cazuri:

$-x_1 \le x_2$

$-x_1 > x_2$

Se va considera cazul $-x_1 > x_2$ și se va arăta că nu este posibilă construirea unei NN astfel încât ecuațiile (a6) și (a7) să fie adevarate. Acest lucru demonstrează că $\bar{O} \neq \bar{Y}$ din principiul extensiei.

Dacă $-x_1 > x_2 \Rightarrow$ ecuația (a7) devine

$$x_1^2 = G(x_2;\Gamma_1) + H(x_1; \Gamma_2) \quad (a8)$$

Din (a8) rezultă că $G(x_2;\Gamma_1)$ trebuie să fie o funcție constantă. Să luăm $G(x_2; \Gamma_1) = c$, pentru $0 \le x_2 \le 1$ și $x_2 < -x_1$. Singura soluție ca $G(x_2; \Gamma_1)$ să fie o constantă este ca :

$$\Gamma_1 = \{ i | w_i = 0, v_i = 0 \text{ sau } w_i = 0, v_i > 0 \text{ sau } w_i > 0, v_i = 0 \} \quad (a9)$$

Dacă Γ_1 e dat de (a9) $\Rightarrow G(x_1;\Gamma_1) = c$ pentru $-1 \le x_1 \le 0, -x_1 > x_2$. Din (a6) $\Rightarrow H(x_2; \Gamma_2) = -c$ pentru $0 \le x_2 \le 1, -x_1 > x_2$. Singura soluție pentru ca $H(x_2; \Gamma_2)$ să fie o constantă este ca Γ_2 să fie dat de ecuația (a10):

$$\Gamma_2 = \{ i | w_i < 0, v_i = 0 \text{ sau } w_i = 0, v_i < 0\} \quad (a10)$$

Dacă este așa, atunci rezultă $H(x_1; \Gamma_2) = -c$ pentru $-1 \le x_1 \le 0$, $x_2 < -x_1$. Ecuația (a8) nu poate fi adevărată pentru $-1 \le x_1 \le 0 \le x_2 \le 1$ și $-x_1 > x_2$.

În concluzie, $\bar{O} \neq \bar{Y} = (\bar{X})^2$, obținut prin aritmetica intervalului sau folosind principiul extensiei, pentru toate rețelele neuronale FNN2, unde $\bar{X}=(x_1,x_2)$, $-1 \le x_1 \le 0 \le x_2 \le 1$.

5. Concluzii

S-au prezentat două metode de extindere a unei funcții reale de una sau mai multe variabile reale

$g: R^m \rightarrow R^n$, la o funcție $\tilde{g}$: $\mathbf{F}(X)\rightarrow\mathbf{F}(Y)$ (procedura de fuzzificare), folosind principiul extensiei și rețelele neuronale fuzzy.

Bibliografie

1. **BUCKLEY J.J., HAYASHI, Y.:** Fuzzy neural networks: A survey. În: Fuzzy Sets and Systems 66, 1994.
2. **BUCKLEY, J.J., HAYASHI, Y.:** Neural nets for fuzzy systems. În: Fuzzy Sets and Systems 71, 1995.
3. **BUCKLEY, J.J.:** Fuzzify. În: Fuzzy Sets and Systems 73, 1995.
4. **NAUCK, D., KLAWONN, F., KRUSE, R.:** Fuzzy Sets, Fuzzy Controllers, and Neural Networks. În: Scientific Journal of Humboldt - University of Berlin, Series Medicine 41, No.4, 1992.
5. **MANES, E.G.:** Recenzie la "Fuzzy Sets and Systems: Theory and Applications" de D. Dubois & H. Prade. În: Bulletin of the American Mathematical Society, Vol.7, No.3, November, 1982.
6. **NEGOIȚĂ, C., RALESCU, D.A.:** Mulțimi vagi și aplicațiile lor, București, Editura Tehnică, 1974.
7. * * * : Lucrările seminarului "Mulțimi fuzzy: teorie și aplicații", ICI, 1996.