Danuta Bogocz
Zakład Statystyki Matematycznej AR w Krakowie

Analiza zastosowań teorii zbiorów rozmytych w badaniach ekonomicznych na wybranych przykładach

Wstęp

Pojęcia matematyczne, z jednej strony będąc w pełni autonomicznymi tworami abstrakcyjnymi, stanowią jednocześnie podstawę do rozwoju teorii znajdujących zastosowanie w różnych dziedzinach, niekoniecznie bezpośrednio związanych z ich wyjściowym pochodzeniem. Kontekst ten niewątpliwie wpływa na kształt rozwijanej teorii, koncentrując uwagę na tych aspektach bądź cechach danego pojęcia, które są tu szczególnie przydatne. Rezultatem tego rodzaju analiz są często nowe teorie, powstałe w wyniku modyfikacji definicji, odrzucenia aksjomatu itp.

Geneza pojęcia zbioru rozmytego sięga 1965 roku. Punktem wyjścia do rozwoju teorii jest definicja zbioru za pomocą funkcji charakterystycznej.

Definicja 1. Funkcja $\mu A : X \to [0, 1]$ jest funkcją charakterystyczną zbioru A wtedy i tylko wtedy, gdy dla wszystkich x

$$\mu_A(x) = \begin{cases} 1, & \text{dla } x \in A, \\ 0, & \text{dla } x \notin A. \end{cases}$$

Teoria zbiorów rozmytych stanowi rozszerzenie tak ujętej teorii zbiorów. Zdefiniowany wyżej zbiór jest tam nazywany zbiorem ostrym, gdyż dla dowolnego elementu przestrzeni X mamy: $x \in A$ albo $x \notin A$. Obok teorii zbiorów rozmytych, będącej najbardziej upowszechnionym uogólnieniem własności przynależności (tj. funkcji charakterystycznej), znane są inne próby jej modyfikacji, takie jak wstępna teoria zbiorów czy teoria Dempstera-Shafera.

Definicja zbioru rozmytego za pomocą funkcji przynależności jest następująca:

Definicja 2.1. Funkcja przynależności μ_A zbioru rozmytego A określonego w przestrzeni U jest funkcją postaci

$$\mu_F : U \rightarrow [0, 1].$$

Gwoli ścisłości trzeba tu dodać, że w ujęciu czysto formalnym teoria ta jest właściwie teorią podzbiorów rozmytych. Niemniej biorąc pod uwagę fakt, że każdy podzbiór z definicji sam w sobie jest zbiorem, możemy dla uproszczenia stosować te określenia zamiennie. Dalej podajemy szczegółową definicję podzbioru rozmytego przytaczaną przez C. Noworola za jednym z głównych prekursorów pojęcia, L.A. Zadehem (zob. [8]).

Definicja 2.2. Niech U będzie zbiorem skończonym lub nieskończonym. Podzbiór rozmyty A zbioru U określa formuła:

$$A=\{(x\colon \mu_A(x)\in \mathrm{M})\text{ , dla każdego } x\in U\},$$

gdzie:
$\mu_A(x)$ oznacza stopień przynależności elementu x do zbioru rozmytego A,
μ_A oznacza funkcję przynależności, która każdemu elementowi x przyporządkowuje pewną wartość ze zbioru M.
Zbiór M nazywa się zbiorem przynależności. Zakłada się, że jest to zbiór uporządkowany. W większości przypadków przyjmuje się przedział [0, 1] (zob. definicja 2.1.).

W niektórych pozycjach literatury, zwłaszcza z zakresu zastosowań teorii zbiorów rozmytych w badaniach operacyjnych i praktyce gospodarczej, można napotkać nieco inną definicję zbioru rozmytego jako liczby rozmytej.

Definicja 2.3. Liczbą rozmytą nazywamy uporządkowaną czwórkę liczb rzeczywistych dowolnego znaku $\left(\underline{a},\underline{b},\overline{b},\overline{a}\right)$, gdzie $\underline{a} \le \underline{b} \le \overline{b} \le \overline{a}$.

Definicja 2.4. Liczbę rozmytą $\left(\underline{a},\overline{a},\alpha,\beta\right)_{L\text{-}R}$ nazywamy liczbą typu *L-R*, jeśli jej funkcja uczestnictwa ma postać:

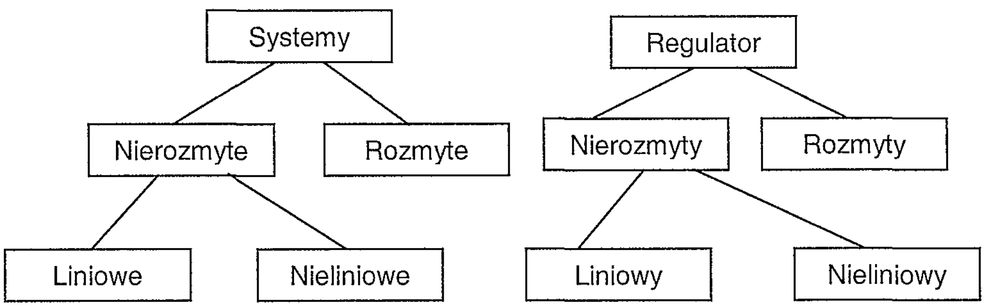
$$\mu_A(x)=\begin{cases} L((\underline{a}-x)/\alpha), & \text{dla } x\in [-\infty,\underline{a}] \\ 1, & \text{dla } x\in [\underline{a},\overline{a}], \\ R((x-a)/\beta) & \text{dla } x\in [\,\overline{a},\infty] \end{cases}$$

gdzie L i R są funkcjami kształtu, zaś α i β – nieujemnymi parametrami.

Definicja 2.3 stanowi pewne zawężenie rodziny liczb rozmytych do tzw. *trapezowych liczb rozmytych*, niemniej na potrzeby niniejszej pracy zakres ten jest w zupełności wystarczający.

Zbiory rozmyte w teorii sterowania

Pojęcie zbioru rozmytego najwcześniej zostało zaadaptowane na gruncie teorii systemów. Jedną z gałęzi tej szerokiej i prężnie rozwijającej się dziedziny jest teoria sterowania. Tu również w ostatnich czasach coraz większą popularność zyskuje tzw. sterowanie rozmyte. Przyczyn tego zjawiska jest wiele. Jedną z nich jest nieliniowość obiektów sterowania, gdyż procesy rzeczywiste wymagające automatycznego sterowania są z reguły procesami nieliniowymi. W zależności od rodzaju nieliniowości stosowane są różne typy regulatorów rozmytych opartych na wiedzy, nazywanych w skrócie regulatorami FKBC (ang. *Fuzzy Knowledge Based Controllers*). Na rysunku 1 schematycznie przedstawiono systemy i regulatory.



Rysunek 1
Uogólniony schemat systemów i regulatorów

Podstawowe elementy tworzące strukturę regulatora FKBC to:

1. Moduł rozmytości. Tu parametrem projektowania jest wybór strategii fuzyfikacji.
2. Baza wiedzy, na którą składają się:
   - baza danych – nośnik niezbędnych informacji zawierających zbiory rozmyte odwzorowujące znaczenie wartości lingwistycznych stanu

procesu i zmiennych sterujących oraz fizyczne dziedziny i ich znormalizowane odpowiedniki,
- baza reguł – strukturalne przedstawienie strategii sterowania.

1. Maszyna wnioskująca.
2. Moduł defuzyfikacji. Tu parametrem projektowania jest wybór operatorów defuzyfikacji.

Wśród regulatorów FKBC na szczególną uwagę zasługują regulatory adaptacyjne. Regulatory te przestrajane są automatycznie w zależności od bieżących charakterystyk procesu, co wobec zmienności parametrów procesu wraz ze zmianą punktu pracy stanowi istotny atut w porównaniu z regulatorami konwencjonalnymi, każdorazowo dostrajanymi w danym punkcie pracy lub w ograniczonym przedziale czasu.

Elementami różniącymi regulatory adaptacyjne od regulatorów konwencjonalnych są monitor procesu oraz mechanizm adaptacyjny. Funkcję monitora może spełniać miara jakości szacująca jakość sterowania lub estymator parametru aktualizujący model procesu. Z kolei mechanizm adaptacyjny wykorzystuje informację dochodzącą z monitora procesu do aktualizacji parametrów regulatora, adaptując tym samym regulator do zmieniających się charakterystyk procesu. Parametrami, które mogą być zmieniane w celu modyfikacji działania regulatora są:

- czynniki skalujące dla każdej zmiennej,
- zbiór rozmyty odwzorowujący znaczenie wartości lingwistycznych,
- reguły JEŻELI-TO.

Regulatory adaptacyjny zastosowano m.in. do symulacji komputerowej zbiornika masy, pieca cementowego oraz wyparki z pojedynczą wymuszoną cyrkulacją.

Zbiory rozmyte w badaniach operacyjnych

Badania operacyjne w praktyce gospodarczej

Pewne ciekawe propozycje rozmytych rozwiązań problemów oceny i wyboru inwestycji można znaleźć w pracy Kuchty [6]. Autorka przedstawia rozszerzenia metod klasycznych na przypadek nieprecyzyjnych danych, w szczególności nieprecyzyjnej znajomości momentów, w których będą następowały przepływy gotówkowe oraz wzrostu niepewności w ocenie stopy oprocentowania z upływem czasu, koncentrując uwagę na metodzie teraźniejszej wartości netto (ang. *Net Present Value*), nazywanej w skrócie NPV. Wskaźnik wzrostu wartości firmy dzięki realizacji danego projektu inwestycyjnego w metodzie NPV jest postaci:

$$NPV = CF_0 + \sum_{j=1}^{n} \frac{CF_j}{(1+r)^j}, \qquad (*)$$

gdzie:
n – przewidywany czas trwania projektu,
CF_j (j = 0, ..., n) – przepływy gotówkowe (nakłady lub wpływy) związane z danym projektem inwestycyjnym, mające miejsce na końcu j-tego roku,
r – stopa dyskontowa, wyznaczana na podstawie kosztu kapitału potrzebnego do pozyskania środków na projekt inwestycyjny i/lub stóp zwrotu innych dostępnych dla firm możliwości inwestycyjnych.
W celu eliminacji wad metody NPV wynikających z nieprecyzyjności informacji autorka wprowadza jej wersje rozmyte poprzez rozmycie we wzorze (*) następujących parametrów:

a) n, w zawężeniu do realizacji liczb całkowitych bądź dowolnie,
b) CF_j (j = 0, ..., n), w zawężeniu do liczb nieujemnych lub niedodatnich bądź dowolnie,
c) r.

Badania operacyjne w analizie danych

Przykładem zastosowań metod rozmytych w zadaniach programowania liniowego jest praca Chanasa [4]. Jako współczynniki w funkcji celu autor przyjmuje liczby rozmyte. Wówczas wartość funkcji celu dla każdego rozwiązania dopuszczalnego, będąca oceną jego użyteczności, jest liczbą rozmytą. Formalnie zadanie to można sformułować w sposób następujący:

$$F(x) = \sum C_j x_j \rightarrow \max,$$

$$x \in : X : \begin{cases} \sum a_{ij} x_j \leq b_i & i = 1, \ldots, m, \\ x_j \geq 0 & j = 1, \ldots, n, \end{cases}$$

gdzie $C_j = (\underline{c}_j, \overline{c}_j, \alpha_j, \beta_j)$, j = 1, ..., n są liczbami rozmytymi typu L-R.

Rozmyta postać funkcji celu jest konsekwencją faktu, że klasa liczb rozmytych typu L-R jest zamknięta ze względu na operację dodawania i mnożenia przez skalar.

Rozmyte modele taksonometryczne

Hierarchiczne skupianie zbiorów rozmytych

O dobrym dopasowaniu modeli rozmytych do procesów i zjawisk występujących w naukach pokrewnych ekonomii przekonuje praca Noworola [8]. Autor wykorzystuje hierarchiczną analizę skupień do konstrukcji testów psychometrycznych. Podobnie jak w przypadku klasycznej analizy hierarchicznej, podstawą skupiania obiektów bądź zmiennych jest tu macierz odległości, inaczej – macierz podobieństwa między nimi. Problem sprowadza się zatem do wyboru metryki właściwej danej sytuacji badawczej. W odniesieniu do konstruowania testów psychometrycznych autor zaleca stosowanie uogólnionej odległości euklidesowej oraz uogólnionej odległości Hamminga:

Uogólniona odległość Hamminga:

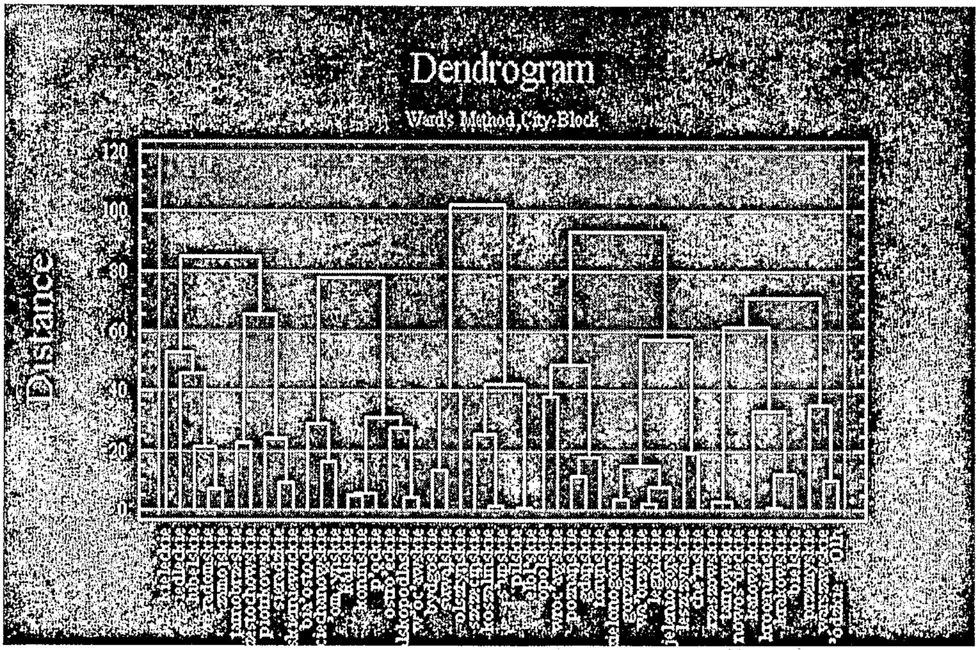
$$d_H(A,B)=\sum_{i=1}^{n}\left(MAX\left(\mu_A(x_i),\mu_B(x_i)\right)-MIN\left(\mu_A(x_i),\mu_B(x_i)\right)\right).$$

Uogólniona odległość euklidesowa:

$$d_H(A,B)=\left(\sum_{i=1}^{n}\left(MAX\left(\mu_A(x_i),\mu_B(x_i)\right)-MIN\left(\mu_A(x_i),\mu_B(x_i)\right)\right)^2\right)^{\frac{1}{2}}.$$

Algorytmy klasyfikacji obiektów wynikające z analizy hierarchicznej można znaleźć w wersjach elektronicznych większości pakietów statystycznych. Użytkownik może wybrać spośród wielu dostępnych tam metod oraz odległości. Oprócz macierzy podobieństwa oraz specyfikacji skupień, można otrzymać dendrogram przedstawiający połączenia obiektów na różnych poziomach podobieństwa.

Na rysunku 2 przedstawiono przykładowy dendrogram otrzymany w wyniku zastosowania metody Warda z odległością miejską do klasyfikacji województw pod względem podobieństwa ich struktury agrarnej. Maksymalna liczba skupień – w analizowanym przypadku 5 – jest parametrem, który zadaje użytkownik.



Rysunek 1
Struktura powiązań województw na różnych poziomach podobieństwa ich struktury agrarnej

Relacja podobieństwa rozmytego

Nieco inną koncepcję rozmytych rozwiązań problemów taksonometrycznych prezentuje Ostasiewicz [9]. Tu pojęcie rozmytości zostało wprowadzone do formuły relacji podobieństwa, na bazie której dokonuje się klasyfikacji rodziny zbiorów. W odróżnieniu od klasycznej relacji podobieństwa, relacja podobieństwa rozmytego nie musi spełniać postulatów odległości. Wymaga się, by relacja ta spełniała warunek zwrotności oraz symetryczności, natomiast warunek przechodniości, właściwy każdej relacji równoważności, występuje w wersji zmodyfikowanej.

Formalny zapis funkcji definiującej relację podobieństwa rozmytego R oraz pierwszych dwóch warunków przedstawia się następująco:

$$\mu_R : \chi \times \chi \rightarrow [0, 1]; \text{ przy czym:}$$
$$\mu_R(a, a) = 1 \quad \text{oraz} \quad \mu_R(a, b) = \mu_R(b, a).$$

Złożenie R°R relacji rozmytej R określa się analogicznie do sposobu definiowania złożenia relacji klasycznej, przyjmując uogólnienia działań, oznaczone tutaj symbolami # oraz *:

$$\#: [0, 1]^2 \rightarrow [0, 1];$$
$$*: [0, 1]^2 \rightarrow [0, 1].$$

Zakłada się, że są to działania przemienne, łączne oraz niemalejące. Wymaga się ponadto, aby spełniały następujące warunki:

$$\min (a, b) \leq a \# b \leq \max (a, b);$$
$$0*0 = 0 \,;\, 0*1 = 0 \,;\; 1*1 = 1.$$

Relację R°R = $[r_{ij}]$ określa następujący wzór:

$$r_{ij} = \bigcup_{k=1}^{n} p_{ik} * q_{kj},$$

gdzie:

$$\bigcup_{k=1}^{n} a_i = a_1 \# a_2 \# ... \# a_n .$$

Wtedy warunek przechodniości relacji rozmytej R można zapisać następująco:

$$R \supseteq R°R.$$

W zależności od przyjętych definicji działań # oraz * otrzymuje się różne rodzaje przechodniości rozmytej.

Działania # i * , określone następująco: #: „+”; *: min, zastosowano do grupowania województw o podobnej strukturze wiekowej grubizny drzewostanów leśnych [2]. Mając macierz A:

$$A = \begin{bmatrix} a_{11};\; a_{12};...a_{1n} \\ a_{21};a_{22};...a_{2n} \\ .... \\ a_{m1};a_{m2};..a_{mn} \end{bmatrix};$$

przedstawiającą udziały n zmiennych charakteryzujących m obiektów, relację podobieństwa rozmytego określa macierz R = A°A^T, gdzie T oznacza transpozycję macierzy. Zatem $r_{ij} = \sum_{t=1}^{n} \min(a_{it}, a_{tj})$, dla $i, j = 1, 2, m$.

Rozmyte metody wartościowania liniowego

Jedną z popularnych form przekazu informacji ekonomicznej są rankingi. Ranking, czyli uporządkowany pod względem jakości analizowanej cechy zbiór obiektów, służy z reguły do wyodrębnienia podzbioru obiektów najlepszych. Tu również można wprowadzić element rozmytości, definiując funkcję przynależności postaci (Cerioli, Zani):

$$f(o_i) = \begin{cases} 1, & \text{gdy } o_i \le c_1, \\ \dfrac{c_2 - o_i}{c_2 - c_1}, & \text{gdy } c_1 < o_i < c_2, \\ 0, & \text{gdy } o_i \ge c_2, \end{cases}$$

gdzie c_1 oznacza rangę progową, poniżej której przynależność do wyróżnionego zbioru jest gwarantowana, zaś c_2 – rangę progową, powyżej której przynależność ta jest niemożliwa. Formułę tę zastosowano w badaniach nad zróżnicowaniem skali inwestowania indywidualnych gospodarstw rolnych. Wyniki przedstawiono w formie tabelarycznej (tab. 1), zamieszczając zarówno wartości zmiennej syntetycznej, która stanowiła podstawę rankingu, jak i wartości funkcji przynależności.

Tak sformułowaną funkcję przynależności stosuje się również do ustalenia poziomu ryzyka ubóstwa. Adekwatność teorii zbiorów rozmytych w odniesieniu do pomiaru ubóstwa wynika z faktu, że przejście od stanu gorszej do lepszej sytuacji materialnej dokonuje się w sposób płynny. Dlatego też skonstruowano szereg charakterystyk rozmytych. Przykładowo, rozmyty indeks ubóstwa, FIP (ang. *Fuzzy Index of Poverty*):

$$FIP = \frac{1}{n} \sum_{i=1}^{n} \mu_P(i).$$

FIP interpretuje się jako procent populacji należącej w sensie rozmytym do podzbioru ubogich.

Tabela 1

Ranking województw ze względu na wielkość skali inwestowania indywidualnych gospodarstw rolnych

Województwo	Wartość zmiennej syntetycznej	Ranga	Wartość funkcji przynależności
kieleckie	8,873206	1	1
lubelskie	7,267246	2	1
zamojskie	6,241985	3	1
siedleckie	5,772944	4	0,956522
opolskie	5,618727	5	0,913043
kaliskie	5,174093	6	0,869565
tarnowskie	5,148303	7	0,826087
bydgoskie	5,126678	8	0,782609
tarnobrzeskie	5,084906	9	0,73913
rzeszowskie	5,003752	10	0,695652
poznańskie	4,942331	11	0,652174
radomskie	4,762446	12	0,608696
katowickie	4,724017	13	0,565217
białostockie	4,711953	14	0,521739
nowosądeckie	4,693512	15	0,478261
krakowskie	4,247282	16	0,434783
sieradzkie	4,018087	17	0,391304
konińskie	3,689062	18	0,347826
częstochowskie	3,609154	19	0,304348
piotrkowskie	3,58138	20	0,26087
ciechanowskie	3,579268	21	0,217391
toruńskie	3,465059	22	0,173913
łomżyńskie	3,462771	23	0,130435
bialskopodlaskie	3,358784	24	0,086957
płockie	3,356486	25	0,043478
gdańskie	3,158093	26	0
krośnieńskie	3,155732	27	0
olsztyńskie	3,134488	28	0
wrocławskie	3,10145	29	0
przemyskie	3,01644	30	0

cd. tab. 1

Województwo	Wartość zmiennej syntetycznej	Ranga	Wartość funkcji przynależności
bielskie	2,969265	31	0
leszczyńskie	2,796559	32	0
ostrołęckie	2,753406	33	0
skierniewickie	2,658571	34	0
pilskie	2,639397	35	0
włocławskie	2,549282	36	0
suwalskie	2,396454	37	0
chełmskie	2,200372	38	0
wałbrzyskie	1,857784	39	0
elbląskie	1,855929	40	0
zielonogórskie	1,766153	41	0
szczecińskie	1,703678	42	0
legnickie	1,519522	43	0
gorzowskie	1,359575	44	0
warszawskie	1,146369	45	0
koszalińskie	1,066384	46	0
słupskie	0,932102	47	0
jeleniogórskie	0,887614	48	0
łódzkie	0,387867	49	0

Literatura

1. BOGOCZ D., Regionalne zróżnicowanie struktury wiekowej grubizny krajowych drzewostanów leśnych, XXV Colloquium Biometryczne, AR w Lublinie, t. 27, 1997.
2. BOGOCZ D., *O rozmytych rozwiązaniach pewnych problemów taksonometrycznych*, Materiały z konferencji: Metody i zastosowania badań operacyjnych, Wrocław 1998.
3. BOGOCZ D., *Miejsce Polski w zbiorowości krajów Europy Zachodniej pod względem poziomu usług teleinformatycznych*, Zeszyty Naukowe AR w Krakowie, seria: Ekonomika, z. 27, 1999.
4. CHANAS S., *Zadanie programowania liniowego z rozmytymi współczynnikami w funkcji celu*, [w:] Metody i zastosowania badań operacyjnych, Praca zbiorowa pod red. T. Trzaskalika, Katowice 1998.

5. DRIANKOV D., HELLENDOORN H., REINFRANK M., *Wprowadzenie do sterowania rozmytego*, Wydawnictwa Naukowo-Techniczne, Warszawa 1996.
6. KUCHTA D., *Rozmyte metody wyboru i oceny inwestycji*, [w:] Metody i zastosowania badań operacyjnych, Praca zbiorowa pod red. T. Trzaskalika, Katowice 1998.
7. KUKUŁA K., *Ranking miejscowości turystycznych Ziemi Sądeckiej ze względu na stan bazy noclegowej*. Materiały z II Międzynarodowej Konferencji nt.: „Szanse i bariery rozwoju przedsiębiorczości w regionie Podkarpacia", 1998.
8. NOWOROL C., *Analiza skupień w badaniach empirycznych – Rozmyte modele hierarchiczne*, PWN, Warszawa 1989.
9. OSTASIEWICZ W., *Zastosowanie zbiorów rozmytych w ekonomii*, PWN, Warszawa 1986.

Some Applications of Fuzzy Set TheoryIn Economic Research

Abstract

The paper presents some applications of fuzzy set theory based on examples from literature and own investigations.

The introduction contains some theoretical background with special attention given to different aspects of the definition and the characteristics of its properties.

In choosing empirical illustrations different fields of economic research are considered such as operation research, data analysis, taxonomy and linear valuation.

As the history of the theory starts with the theory of systems fuzzy controllers and their applications are shortly discussed.