



UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
Instituto de Geociências e Ciências Exatas
Campus de Rio Claro

Uma Aplicação da Lógica Fuzzy

Cristina Helena Bovo Batista Dias

Dissertação apresentada ao Programa de
Pós-Graduação – Mestrado Profissional em
Matemática Universitária do Departamento
de Matemática como requisito parcial para a
obtenção do grau de Mestre

Orientador
Prof. Dr. Henrique Lazari

2010

511.3 Dias, Cristina Helena Bovo Batista
D541a Uma Aplicação da Lógica Fuzzy/ Cristina Helena Bovo Batista
Dias- Rio Claro: [s.n.], 2010.
83 f., il., figs., gráfs., tab.

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas.
Orientador: Henrique Lazari

1. Lógica Simbólica e Matemática. 2. Entropia Fuzzy. 3. Fuzziness. 4. Geometria dos Conjuntos Fuzzy. 5. Hipercubo Fuzzy. I.
Título

TERMO DE APROVAÇÃO

Cristina Helena Bovo Batista Dias
UMA APLICAÇÃO DA LÓGICA FUZZY

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação Mestrado Profissional em Matemática Universitária do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, pela seguinte banca examinadora:

Prof. Dr. Henrique Lazari
Orientador

Prof. Dr. Adilson José Vieira Brandão
Universidade Federal de São Carlos - Campus Sorocaba

Prof. Dr. Wlademir Seixas
Universidade Federal de São Carlos - Campus Sorocaba

Rio Claro, 14 de Outubro de 2010

*Ao meu Deus, Jeová, aos meus pais,
ao meu querido esposo e aos meus filhos e irmãos*

Agradecimentos

Primeiramente, agradeço a Deus pela vida. A seguir, ao meu orientador, aos membros da banca e, de modo geral, a todos os docentes e funcionários do Departamento de Matemática do IGCE, Rio Claro, por sua valiosa colaboração e apoio, sem os quais este trabalho não teria sido realizado.

Agradeço, ainda, aos funcionários da UNESP, em especial, os que trabalham na Biblioteca, no Restaurante Universitário e na Cantina que, dos bastidores, dão suporte para que trabalhos como este sejam produzidos. Finalmente, agradeço a meus pais, ao meu esposo e aos meus filhos, e a todos meus familiares, amigos e irmãos.

“Pense que não há coisa como chance no mundo, nossa ignorância é a causa real de qualquer evento ter alguma influência sobre o entendimento e requerer alguma espécie de crença.”

David Hume

Resumo

Desde cedo entramos em contato com as implicações lógicas. O binômio verdadeiro-falso está sempre presente em nossas vidas e nós nos acostumamos a aceitar que as coisas ou são verdadeiras ou são falsas. Divertimo-nos quando alguém nos conta histórias interessantes envolvendo lógica e que terminam em contradições, tais como, por exemplo, a do barbeiro que pode e não pode barbear a si mesmo, ou como a do advogado que consegue ganhar e perder a mesma causa. Apreciamos tais “paradoxos” sem nos apercebermos que por trás deles existe toda uma teoria matemática, a chamada lógica fuzzy. Essa dissertação tem por objetivo apresentar um resumo desta teoria, mostrando como ela trata a existência de tais paradoxos e dar detalhes sobre uma visão compacta dos conjuntos fuzzy, a saber, utilizando uma representação geométrica. A análise de alguns resultados sobre tais conjuntos usando esta representação leva a uma justificativa para o estudo da lógica fuzzy, a saber, a diferença entre “fuzziness” e probabilidade, incluindo uma demonstração de que “fuzziness”, de fato, existe.

Palavras-chave: Lógica Simbólica e Matemática, Entropia Fuzzy, Fuzziness, Geometria dos Conjuntos Fuzzy, Hipercubo Fuzzy.

Abstract

Early on we got in touch with the logical implications. The binomial true-false is always present in our lives and we have come to accept that things are either true or false. We have fun when somebody tells interesting stories involving logic and ending with contradictions, such as, for example, the barber who can and cannot shave himself, or as the lawyer who can win and lose the same cause. We appreciate such “paradoxes” without realizing that behind them there is a whole mathematical theory, called fuzzy logic. This dissertation aims to present a summary of this theory, showing how it deals with the existence of such paradoxes and give details about a compact view of fuzzy sets, namely, using a geometrical representation. The analysis of some results on such sets using this representation leads to a justification for the study of fuzzy logic, namely the difference between “fuzziness” and probability, including a demonstration that “fuzziness”, in fact, exists.

Keywords: Symbolic Logic and Mathematics, Fuzzy Entropy, Fuzziness, Fuzzy Sets’ Geometry, Fuzzy Hypercube.

Lista de Figuras

1.1	Exemplos de representações gráficas de Conjuntos cujos elementos podem ser considerados próximos de 6	15
1.2	Gráfico do Conjunto Fuzzy definido pela Função de Pertinência do Exemplo 1.3	18
1.3	Conjunto Fuzzy A representado como Ponto	21
1.4	Conjunto Potência $F(2^B)$ como Hiperretângulo no Hiper cubo Fuzzy . .	24
1.5	Partição do Quadrado Fuzzy	25
1.6	Exemplo no Quadrado Fuzzy particionado	26
1.7	Partição do Cubo Fuzzy	27
2.1	Gráfico do Conjunto Fuzzy definido pela Função de Pertinência do Exemplo 2.1	31
2.2	Gráfico do Conjunto Fuzzy do Exemplo 2.4	37
3.1	Número Fuzzy triangular	46
3.2	Número Fuzzy trapezoidal	46
3.3	Exemplo de um Número Fuzzy em forma de sino	47
4.1	Simetria em relação ao Ponto Médio	58
4.2	Simetria em relação ao Ponto Médio com $\varphi(x_i) < \frac{1}{2}, \forall i$	59
4.3	Medida M do Conjunto Fuzzy A	61
5.1	Interpretação Geométrica da Medida de Entropia Fuzzy	64
5.2	Geometria do Teorema da Entropia Fuzzy	65
5.3	Simplexo S^2 no Quadrado I^2	69
5.4	Simplexo S^3 no Cubo I^3	69
5.5	Dependência da Inclusão em relação ao Contador $M(A)$	74
5.6	Ilustração do Teorema da Inclusão	77
5.7	Ilustração do Teorema da Entropia-Inclusão em duas dimensões	80

Lista de Tabelas

1.1	Comparação entre Redes Neurais e Sistemas Fuzzy	20
2.1	Conectivos da Lógica Clássica	41

Sumário

I	Resumo da Teoria	12
1	Conceitos Preliminares	13
1.1	Introdução	13
1.1.1	Organização da Dissertação	13
1.2	Um pouco de história	14
1.3	Subconjuntos Fuzzy	16
1.4	Representações de Conjuntos Fuzzy	16
1.4.1	Função de Pertinência	16
1.4.2	Conjunto Clássico de Pares Ordenados	18
1.4.3	Notação $\frac{\varphi_A(x_i)}{x_i}$	19
1.4.4	Notação Hipercubo	19
1.5	Aplicação em Redes Neurais	19
1.5.1	Conjuntos como Pontos	21
1.5.2	Grau de Inclusão	22
1.6	Entropia e Informação	27
1.6.1	Teoria da Informação	27
1.6.2	Distribuição de Probabilidade	28
1.6.3	Entropia	28
1.7	Simplexos	30
2	Aritmética e Lógica Fuzzy	31
2.1	Suporte de um Conjunto Fuzzy	31
2.2	Operações com Conjuntos Fuzzy	32
2.3	O conceito de α -nível	37
2.4	Lógica Fuzzy	40
2.4.1	Conectivos da Lógica Clássica	40
2.4.2	Conectivos Básicos da Lógica Fuzzy	40
2.4.3	Implicações Fuzzy	43
3	Números Fuzzy	45
3.1	Introdução	45

3.2	Operações com Números Fuzzy	47
3.2.1	Operações com Intervalos	47
3.2.2	Operações com Números Fuzzy	49
 II Aplicação:		
Representação Geométrica dos Conjuntos Fuzzy		52
 4	“Fuzziness” e Probabilidade	53
4.1	Conjuntos Fuzzy e Aleatoriedade	53
4.1.1	“Fuzziness” e Probabilidade	53
4.1.2	A Diferença entre SE e QUANTO	54
4.2	Retomando a Visão de Conjuntos como Pontos	56
4.2.1	Operando com a Representação Geométrica	56
4.3	O Ponto Médio do Quadrado Fuzzy	57
4.4	O Contador Sigma	61
 5	Teoremas sobre Conjuntos Fuzzy	63
5.1	Teorema da Entropia Fuzzy	63
5.1.1	Entropia e Conjuntos Fuzzy	63
5.1.2	Consequências	67
5.1.3	Aplicação	68
5.1.4	Aplicação em uma Distribuição de Probabilidades	69
5.2	O Teorema da Inclusão	71
5.2.1	Dedução Algébrica da Medida do Grau de Inclusão	71
5.2.2	Dedução Geométrica da Medida do Grau de Inclusão	73
5.3	O Teorema da Entropia-Inclusão	79
5.3.1	Principal Consequência	79
 6	Conclusão	81
 Referências		82

Parte I

Resumo da Teoria

1 Conceitos Preliminares

1.1 Introdução

Nessa dissertação será apresentado um resumo da Lógica Fuzzy, como base para a compreensão de um tema muito interessante, o centro deste trabalho, a saber, a visão geométrica dos conjuntos fuzzy. Adicionalmente será analisada a questão: Será que “fuzziness” existe? Ou será que a probabilidade e a lógica bivalente conseguem lidar com todos os problemas ligados à incerteza? Mas, inicialmente, o que significam “fuzzy” e “fuzziness”?

“*Fuzzy*” é uma palavra da língua inglesa, que significa nebuloso, difuso, incerto, impreciso, subjetivo, etc.... No entanto, como a palavra original tem um significado muito mais amplo do que o destas traduções, na maioria dos casos prefere-se utilizá-la ao invés de traduzí-la. Quando traduzido o termo geralmente usado é *nebuloso*.

A melhor maneira de se ter uma ideia geral do que seja “fuzziness” é compará-la com a teoria da probabilidade. O objetivo de “fuzziness” é medir o QUANTO um evento ocorre enquanto que a teoria da probabilidade descreve a aleatoriedade de um evento. SE ele ocorre ou não. Se a ocorrência de um evento envolver componentes INCERTOS, envolverá a probabilidade de um evento “fuzzy”. Portanto, “fuzziness”, assim como probabilidade, é uma forma de medir a incerteza.

1.1.1 Organização da Dissertação

Conceitos Preliminares

Esse capítulo apresenta uma breve introdução histórica seguida de alguns conceitos preliminares. Entre estes merecem destaque o de subconjunto fuzzy, o de função de pertinência e o de grau de inclusão. Na conclusão, introduz a representação geométrica dos conjuntos fuzzy finitos por meio de um hipercubo, base da dissertação.

Aritmética e Lógica Fuzzy

O objetivo do Capítulo 2 é dar uma ideia de como operar com os conjuntos fuzzy através de sua função de pertinência. São apresentados maiores detalhes sobre o suporte

de um conjunto fuzzy e sobre os chamados α -níveis.

Números Fuzzy

Para $X = \mathbb{R}$ existe uma família de conjuntos fuzzy bastante utilizada em aplicações. Os conjuntos fuzzy que fazem parte dessa família são chamados números fuzzy. Costuma-se representar os números fuzzy através do gráfico de sua função de pertinência. Maiores detalhes sobre eles serão analisados no Capítulo 3.

“Fuzziness” e Probabilidade

A partir do Capítulo 4 inicia-se a discussão da questão principal apresentando algumas diferenças básicas entre “fuzziness” e probabilidade, retomando e ampliando a visão geométrica dos conjuntos fuzzy e destacando o ponto médio do hipercubo fuzzy. O capítulo termina com a definição do contador sigma de Zadeh, que será utilizado a seguir na medição da “fuzziness” de um conjunto.

Teoremas sobre Conjuntos Fuzzy

Finalmente, o Capítulo 5, apresenta alguns resultados importantes utilizados para responder à questão sobre se “fuzziness” é de fato uma medida de incerteza diferente da probabilidade.

1.2 Um pouco de história

Em 1920, Jan Lukasiewicz escreveu o artigo *On Three-valued Logic*¹, onde apresentou pela primeira vez, o conceito de lógicas multivalentes, que estende os valores verdade das proposições a conjuntos mais complexos que $\{0, 1\}$.

Em 1965, Lotfi Asker Zadeh [4] apresentou a Teoria dos Conjuntos Fuzzy como maneira de tratar de forma matemática termos linguísticos subjetivos como “*em torno de*” e “*aproximadamente*”. A ideia era conseguir tratar de forma computacional conceitos vagos, da mesma forma que um ser humano consegue fazer. A fim de formalizar este conceito, Zadeh estendeu o conceito de *função característica* de um conjunto clássico, ou seja, de um conjunto tradicional (também chamado *crisp* na lógica fuzzy), a saber:

Definição 1.1. *Sejam X um conjunto e A um subconjunto de X . A função $\chi_A : X \rightarrow \{0, 1\}$ dada por:*

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases},$$

¹O artigo original foi escrito em polonês, [2] e traduzido para o inglês em 1970, [3] no entanto, o acesso à tradução bem como ao original é difícil.

é chamada **função característica** de A .

Observação 1.1. Sejam $\wp(X)$ o conjunto das partes de X e $\mathfrak{S}(X, \{0, 1\})$ o conjunto das funções de X em $\{0, 1\}$. Fixado um conjunto $X \neq \emptyset$, existe uma bijeção:

$$\chi : \wp(X) \rightarrow \mathfrak{S}(X, \{0, 1\})$$

dada por $A \mapsto \chi_A$, que faz corresponder a cada subconjunto A de X a sua função característica. Pode-se dizer então, por um abuso de linguagem, conforme a conveniência, que $\chi_A : X \rightarrow \{0, 1\}$ é um subconjunto em X .

Observação 1.2. A função da Definição 1.1 descreve plenamente o conjunto estudado assumindo o valor 1, se o elemento x pertence ao conjunto e 0, se ele não pertence. Quando esta descrição simples não corresponde à ideia que se quer expressar matematicamente procura-se definir funções que façam isto de maneira mais apropriada.

Exemplo 1.1. Exemplificando o que foi dito na observação anterior considere o conjunto $A = \{x \in \mathbb{R} : "x \text{ está próximo de } 6"\} \subseteq \mathbb{R}$. Que elementos pertencem a A ?

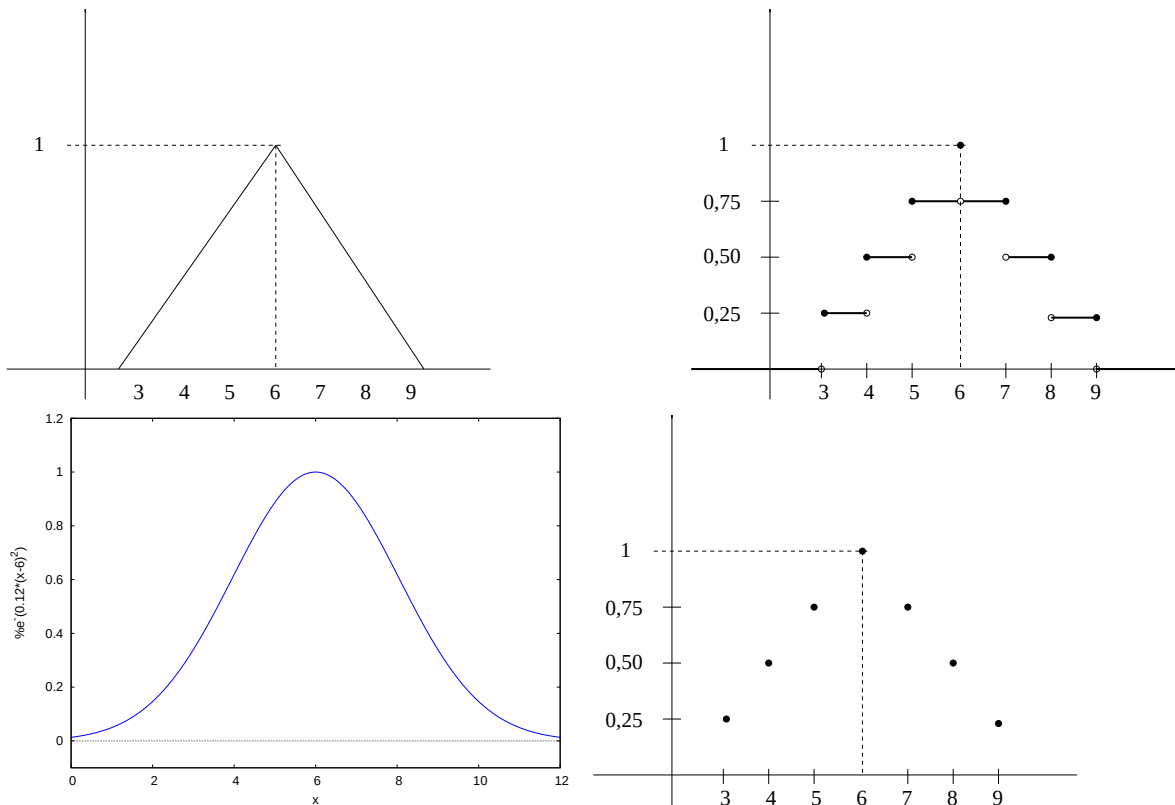


Figura 1.1: Exemplos de representações gráficas de Conjuntos cujos elementos podem ser considerados próximos de 6

Note que a resposta a esta pergunta depende do contexto no qual ela foi levantada. Quando estão envolvidas distâncias numa rodovia o km 5 pode ser tomado como próximo do km 6. Em um laboratório de Química, por outro lado, 0,5 mm pode ser

distante demais de 0,6 mm. Em outros contextos como, por exemplo, se nosso conjunto universo for o conjunto $\mathbb{Z}$ ($x \in \mathbb{Z}$), estar próximo de 6 pode significar ser igual a 6.

A Figura 1.1 apresenta o gráfico de diversas funções que podem ser escolhidas para representar o conjunto dos pontos x próximos de 6, assumindo valores maiores do que zero.

1.3 Subconjuntos Fuzzy

Para tratar casos como o do Exemplo 1.1, Zadeh [4] propôs uma teoria que incluiu a definição de **subconjunto fuzzy** F como sendo “um subconjunto F em X caracterizado por uma função

$$\varphi_F : X \longrightarrow [0, 1],$$

pré-fixada, chamada de **função de pertinência** do subconjunto F .”

O valor $\varphi_F(x) \in [0, 1]$ indica o grau com que o elemento x de X pertence ao subconjunto fuzzy F . $\varphi_F(x) = 0$ e $\varphi_F(x) = 1$ indicam, respectivamente, a não pertinência e a pertinência total de x a F .

Observação 1.3. Nem sempre o conjunto universo X , no qual um subconjunto fuzzy é referenciado, será mencionado. De acordo com isso, ao se citar um subconjunto fuzzy em X , tanto a expressão “subconjunto fuzzy” como a expressão “conjunto fuzzy” serão utilizadas, dependendo de se o conjunto universo X está explicitamente citado ou não.

1.4 Representações de Conjuntos Fuzzy

Observe que a maneira de representar um conjunto fuzzy está relacionada ao contexto. Neste tópico, serão apresentadas algumas das representações mais utilizadas para conjuntos fuzzy.

1.4.1 Função de Pertinência

A função de pertinência de um conjunto fuzzy é obtida ampliando-se o contradomínio da função característica $\{0, 1\}$ para o intervalo $[0, 1]$. Portanto, um conjunto clássico é um caso particular de conjunto fuzzy, cuja função de pertinência φ_F é tal que $Im \varphi_F = \{0, 1\}$.

Desta forma, da mesma maneira que um subconjunto clássico A de X é identificado com sua função característica; fixado X , toma-se $\wp(X)$, o conjunto das partes de X e $\mathfrak{S}(X, [0, 1])$, o conjunto das funções de X em $[0, 1]$, e identifica-se um subconjunto fuzzy F em X com a função de pertinência que o referencia, de acordo com uma bijeção dada por:

$$\begin{array}{ccc} \varphi : \wp(X) & \rightarrow & \mathfrak{S}(X, [0, 1]) \\ F & \mapsto & \varphi_F \end{array}.$$

Tendo isso em mente pode-se definir, formalmente, subconjunto fuzzy em X como abaixo.

Definição 1.2. Um **subconjunto fuzzy** φ_F em X [5] é uma função do conjunto de referência X no intervalo unitário, ou seja,

$$\varphi_F : X \rightarrow [0, 1].$$

Ainda, levando em conta que a função de pertinência é uma extensão da função característica e que, para o complementar F^c de um conjunto $F \subseteq X$ tem-se que $\chi_{F^c} = 1 - \chi_F$, define-se:

Definição 1.3. O **complementar** F^c de um **subconjunto fuzzy** F em X é a função

$$\varphi_{F^c} = 1 - \varphi_F : X \rightarrow [0, 1].$$

Exemplo 1.2. Seja C o conjunto clássico dos números naturais ímpares cuja função característica é dada por

$$\chi_C(n) = \begin{cases} 1 & \text{se } n \text{ for ímpar} \\ 0 & \text{se } n \text{ for par} \end{cases}.$$

Observa-se então que, tomando $\varphi_C(n) = \chi_C(n)$, podemos considerar todo conjunto clássico como um conjunto fuzzy, definindo-se a função de pertinência da mesma maneira que a função característica.

Exemplo 1.3. Seja F o conjunto fuzzy dos números reais próximos de 0, ou seja,

$$F = \{x \in \mathbb{R} : x \text{ é próximo de } 0\}.$$

Pode-se, então, escolher de maneira coerente com o conceito usual de “próximo” uma função de pertinência para este conjunto. Por exemplo, ela pode ser definida como:

$$\varphi_F(x) = \begin{cases} \frac{1}{x^2} & \text{se } |x| \geq 1 \\ 1 & \text{se } |x| < 1 \end{cases},$$

cujo gráfico encontra-se na Figura 1.2.

Tem-se, então, que φ_F caracteriza um conjunto fuzzy F , um subconjunto fuzzy em $\mathbb{R}$. Calculando alguns valores obtém-se:

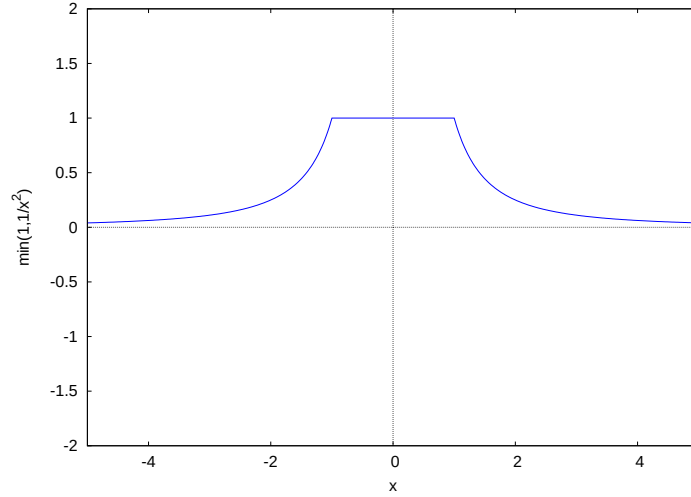


Figura 1.2: Gráfico do Conjunto Fuzzy definido pela Função de Pertinência do Exemplo 1.3

$$\begin{aligned}\varphi_F(1) &= 1, \\ \varphi_F(0.1) &= 1, \\ \varphi_F(3) &= \frac{1}{9}, \\ \varphi_F(7) &= \frac{1}{49}, \\ \varphi_F(999) &= \frac{1}{998001}.\end{aligned}$$

Note, porém, que a escolha de φ_F foi arbitrária, uma vez que outras funções também podem representar bem o conceito de “estar próximo”. Portanto, não existe uma bijeção entre a expressão “estar próximo” e a descrição do conjunto.

Estes são exemplos simples. Em geral, a definição da função de pertinência depende de um estudo detalhado do problema analisado. Os conjuntos fuzzy serão estudados em maiores detalhes no Capítulo 2

1.4.2 Conjunto Clássico de Pares Ordenados

Um subconjunto fuzzy F em X também pode ser representado por um conjunto clássico de pares ordenados, a saber, $F = \{(x, \varphi_F(x)) : x \in X\}$.

Exemplo 1.4. Seja $X = \{x_1, x_2, x_3, x_4, x_5\}$.

O conjunto F , dado por $F = \{(x_1; 0, 3), (x_2; 0, 7), (x_3; 1), (x_4; 0, 5), (x_5; 0)\}$ é um conjunto clássico de pares ordenados, que representa um subconjunto fuzzy F em X , onde o segundo elemento de cada par ordenado indica o grau de pertinência de x_i ao conjunto F .

1.4.3 Notação $\frac{\varphi_A(x_i)}{x_i}$

Uma das representações para um conjunto fuzzy A finito usa a seguinte simbologia:

$$\sum_{i=1}^n \frac{\varphi_A(x_i)}{x_i}, \forall x_i \in X.$$

Perceba que a notação $\frac{\varphi_A(x_i)}{x_i}$ não é utilizada no sentido usual, mas apenas como uma maneira de visualizar o elemento x_i e seu respectivo grau de pertinência $\varphi_A(x_i)$. O mesmo se aplica ao símbolo $+$ e ao símbolo $\sum$, que são usados apenas para ligar os elementos na notação.

Exemplo 1.5. Sejam $X = \{1, 3, 7, 8, 12\} \subset \mathbb{R}$ e A um subconjunto fuzzy em X representado por:

$$A = \sum_{i=1}^n \frac{\varphi_A(x_i)}{x_i} = \frac{0}{1} + \frac{0,2}{3} + \frac{0,3}{5} + \frac{0,25}{7} + \frac{0,7}{8} + \frac{1}{12}.$$

Então seu complementar pode ser representado por:

$$A^c = \sum_{i=1}^n \frac{\varphi_A(x_i)}{x_i} = \frac{1}{1} + \frac{0,8}{3} + \frac{0,7}{5} + \frac{0,75}{7} + \frac{0,3}{8} + \frac{0}{12}.$$

1.4.4 Notação Hipercubo

Outra maneira de representar conjuntos fuzzy finitos é considerá-los como pontos num hipercubo. Esta forma de representação é a base dessa dissertação. Sendo assim, serão destacadas, no próximo tópico, essa representação bem como uma de suas aplicações.

1.5 Aplicação em Redes Neurais

Em engenharia e ciência, sistemas físicos são geralmente descritos por modelos matemáticos. São exemplos as equações diferenciais, os processos probabilísticos e o modelo de relacionamento entre entidades na teoria dos sistemas de bases de dados. Uma dificuldade normalmente encontrada quando se tenta construir um modelo é que simplificações frequentemente são necessárias para dar a um problema um modelo matemático adequado. O avanço da tecnologia de computadores tornou possível a manipulação de sistemas cada vez mais complexos, é verdade, mas também aumentou a demanda por softwares mais simples. Um método para simplificar sistemas complexos é tolerar um grau razoável de imprecisão, indistinção e incerteza durante a fase da modelagem. Embora os sistemas resultantes não sejam perfeitos são, em grande parte, capazes de resolver, de maneira adequada, os problemas modelados. Esse princípio já

se mostrou satisfatório durante a construção de modelos, quando, na fase de desenvolvimento de sistemas baseados em conhecimento, a imprecisão dos sistemas especialistas precisa ser levada em conta.

Um sistema fuzzy é um tipo de modelo matemático para um problema no qual se utilizam conjuntos e relações fuzzy.² Neste tipo de modelo, o objetivo é tirar vantagem da menor complexidade que é conseguida pelo uso criterioso da informação imprecisa. É claro que não são desconsiderados aspectos relevantes do mundo real, pois senão o modelo construído poderia não corresponder à realidade. Por outro lado, a qualidade de um modelo não deve ser medida apenas pelo grau de precisão da informação, mas também pela consideração de critérios como correção, completude, adequabilidade, eficiência e conveniência de uso. Assim, não surpreende que um modelo “fuzzy”; que reduza grandemente a complexidade da descrição do sistema estudado, apenas pela inclusão de alguma informação imprecisa, mas que leve em conta os critérios de qualidade; possa ser considerado melhor do que um modelo que, apenas por exigir o uso de informações totalmente precisas, seja mais difícil de manusear. O uso de tais sistemas tem apresentado expressivo sucesso, especialmente no campo da engenharia de controle e da inteligência artificial.

Dentro da teoria dos sistemas, conhecer a assim chamada “fuzziness” de um conjunto, ou seja, o quanto ele é fuzzy, permite a construção de chips de computadores e de sistemas que, de maneira “inteligente” controlam trens metropolitanos, sistemas automotivos, numerosos aparelhos eletrônicos e uma série de outros instrumentos. Nesse nível o processamento fuzzy é muito similar ao comportamento de redes neurais, como mostrado na Tabela 1.1.

Redes Neurais	Sistemas Fuzzy
Processam Informações Imprecisas.	Processam Informações Imprecisas.
Reconhecem padrões mal-definidos sem precisar de um conjunto de regras.	Estimam funções e controlam sistemas com uma descrição parcial do comportamento deles.
Usam um conjunto de n neurônios.	Usam uma família de conjuntos fuzzy contínuos n -dimensionais.
Emitem sinais que vão de um valor mínimo a um valor máximo (digamos de 0 a 1).	Assumem valores reais entre 0 e 1.

Tabela 1.1: Comparação entre Redes Neurais e Sistemas Fuzzy

Observando-se a tabela, surge naturalmente a ideia de associar a cada sinal de saída de um neurônio da rede um número entre 0 e 1. Por analogia ao conceito de bit em Informática chama-se **fit** (fuzzy unit), o valor produzido por esse sinal de saída. Num conjunto de n neurônios são produzidos n sinais de saída que podem ser associados

²Para detalhes sobre sistemas fuzzy, veja *Fuzzy Control and Fuzzy Systems* [1]

a um vetor *fit*, um vetor-linha de n elementos, onde cada um representa o grau de pertinência de um elemento a um conjunto fuzzy n -dimensional. A ideia de como é feita esta associação será dada a seguir.

1.5.1 Conjuntos como Pontos

Como já visto, cada conjunto fuzzy pode ser associado a uma função de pertinência de um domínio X no conjunto $[0, 1]$. Quando se considera um conjunto como ponto, como na representação de conjuntos fuzzy finitos, utiliza-se o domínio $X = \{x_1, x_2, \dots, x_n\}$ e o contra-domínio $[0, 1]$ da função $\varphi_A : X \rightarrow [0, 1]$.

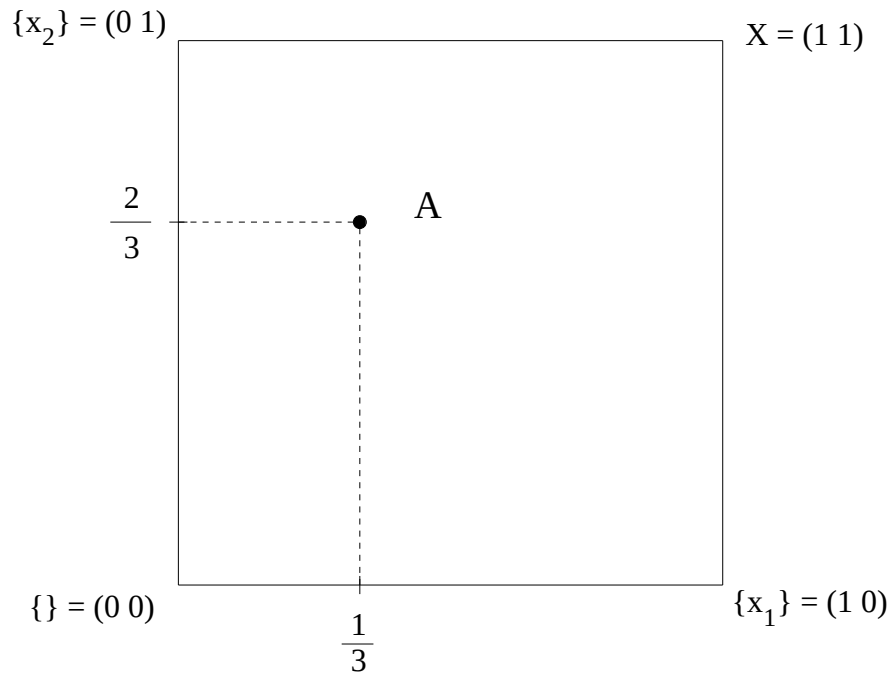


Figura 1.3: Conjunto Fuzzy A representado como Ponto

Tomando, por exemplo, o conjunto de dois elementos $X = \{x_1, x_2\}$, considera-se todos os subconjuntos A desse conjunto. Tem-se dessa maneira, o conjunto-potência clássico 2^X de X , que é dado por $\{\emptyset, X, \{x_1\}, \{x_2\}\}$. Tendo em mente que, nesta notação, a ordem em que os elementos do conjunto 2^X são tomados é importante, estes quatro conjuntos podem ser representados, respectivamente, pelos vetores-linha $(0\ 0)$, $(1\ 1)$, $(1\ 0)$ e $(0\ 1)$ correspondendo aos quatro vértices de um quadrado. Os 1s e os 0s indicam a presença ou a ausência do i -ésimo elemento x_i , respectivamente, no conjunto A considerado. De forma mais abstrata, associa-se unicamente cada conjunto A a uma das funções que tomam os valores $\{0, 1\}$. Tais conjuntos podem ser posicionados no reticulado B^2 (ou ainda em $\mathbb{Z}_2 \times \mathbb{Z}_2$) numa comparação do conjunto clássico 2^X com o quadrado booleano B^2 . Como preencher o interior do quadrado?

Tendo por objetivo encontrar a resposta a esta pergunta e lembrando que todo conjunto clássico pode ser visto como um conjunto fuzzy, define-se a seguir o que se

entende por conjunto fuzzy não-clássico.

Definição 1.4. Chamam-se **conjuntos fuzzy não-clássicos** ou **conjuntos propriamente fuzzy** aqueles subconjuntos fuzzy F em algum conjunto X tais que φ_F não é uma função característica, ou seja, que não podem ser caracterizados por subconjuntos (no sentido usual) de X .

Observação 1.4. Os subconjuntos fuzzy não-clássicos preenchem o interior do quadrado produzindo um quadrado sólido I^2 . Denota-se por $F(2^X)$ o conjunto cujos elementos são todos os subconjuntos fuzzy (clássicos e não clássicos) em X . $F(2^X)$ é denominado **conjunto potência fuzzy** de X e pode ser representado por um hiper-cubo sólido (no exemplo, um quadrado).

Segundo a definição clássica de Zadeh, um conjunto fuzzy A é associado a uma função de pertinência contínua e multivalorada dada por $\varphi_A : X \rightarrow [0, 1]$. Suponha, por exemplo, que um elemento x_1 pertença ao conjunto A com um grau de pertinência igual a $\frac{1}{3}$ e um elemento x_2 com um grau de pertinência $\frac{2}{3}$. Então, o vetor-linha $\left(\frac{1}{3} \ \frac{2}{3}\right)$ pode ser utilizado para representar o conjunto fuzzy A . A imagem $\varphi_A(x_i)$ é representada pela i -ésima coluna do vetor. A visão geométrica dos conjuntos como pontos representa o conjunto A como um ponto em I^2 , o quadrado unitário, como na Figura 1.3. Os valores $\frac{1}{3}$ e $\frac{2}{3}$ medem, dessa forma, a pertinência parcial de x_1 e x_2 a A , respectivamente. O conjunto potência fuzzy $F(2^X)$, que contém todos os subconjuntos fuzzy em X corresponde, como já mencionado, a um quadrado unitário.

1.5.2 Grau de Inclusão

Conjuntos contêm subconjuntos. A será um subconjunto clássico de B , denotado por $A \subset B$ se, e somente se, todo elemento de A é elemento de B . O conjunto potência 2^B contém todos os subconjuntos clássicos de B . Então,

$$A \subset B \text{ se, e somente se, } A \in 2^B.$$

Uma implicação, na lógica clássica, somente será falsa quando seu antecedente for verdadeiro e seu consequente falso, ou seja, quando algo verdadeiro implicar em algo falso.

O mesmo vale para subconjuntos. Pode-se representar a relação inclusão por funções do tipo $\varphi_I : X \rightarrow \{0, 1\}$. Assim, no caso clássico, A será um subconjunto de B se, e somente se, não existe elemento x pertencente a A , mas que não pertença a B , ou seja, onde $\varphi_A(x) = 1$ com $\varphi_B(x) = 0$. Essa definição pode ser reescrita, a fim de estender o contradomínio de φ_I para $[0, 1]$ da seguinte forma

$$A \subset B \text{ se, e somente se, para todo } x \in X \ \varphi_A(x) \leq \varphi_B(x).$$

Zadeh propôs que esta forma da relação acima fosse usada para conjuntos fuzzy. Por analogia, esta relação entre conjuntos fuzzy também foi chamada de inclusão. Pode-se então considerar, tendo essa inclusão em mente, a seguinte definição:

Definição 1.5. *Sejam A e B dois subconjuntos fuzzy em X , com funções de pertinência indicadas por φ_A e φ_B , respectivamente. Diz-se que A é **subconjunto fuzzy de B** , e indica-se $A \subset B$, se $\varphi_A(x) \leq \varphi_B(x)$, $\forall x \in X$.*

Observação 1.5. É importante ressaltar, contudo, que esta relação de inclusão não é fuzzy.

Observação 1.6. O conjunto vazio ($\emptyset$) tem função de pertinência $\varphi_{\emptyset}(x) = 0$, para todo $x \in X$, enquanto que o conjunto universo (X) tem função de pertinência $\varphi_X(x) = 1$, para todo $x \in X$. Portanto, $\emptyset \subset A$ e $A \subset X$, para todo subconjunto fuzzy A em X . Desta maneira, a Definição 1.5 estende a definição de subconjunto fuzzy em X .

Exemplo 1.6. Sejam $A = (0,5 \ 0,2 \ 0,4)$ e $B = (0,8 \ 0,7 \ 0,6)$, dois subconjuntos fuzzy em $X = \{x_1, x_2, x_3\}$. Então, usando a definição de inclusão de Zadeh, A será subconjunto fuzzy de B , mas B não será subconjunto fuzzy de A . Ou seja, de acordo com esta relação, um conjunto fuzzy A é, ou não é, um subconjunto do conjunto fuzzy B . Portanto, a relação de inclusão de Zadeh não é fuzzy, pois sua imagem é o conjunto $\{0,1\}$.

$F(2^B)$ é um conjunto clássico

Lembrando do ponto de vista conjuntos-como-pontos pode-se perguntar qual a representação geométrica dos subconjuntos fuzzy de B . Isto é, com o que o conjunto potência fuzzy de B se parece? A inclusão implica que $F(2^B)$, o conjunto de todos os subconjuntos fuzzy de B , define no hipercubo unitário, um hiperretângulo que inclui a origem e cujos lados têm comprimento de acordo com os valores $\varphi_B(x_i)$. A Figura 1.4 mostra o conjunto potência fuzzy do conjunto fuzzy $B = \left(\frac{1}{4} \ \frac{2}{3}\right)$.

Observação 1.7. $F(2^B)$ tem cardinalidade infinita quando B não é vazio. No entanto, quando se trata de hipercubos de dimensão finita, é possível medir o tamanho de $F(2^B)$ como o volume

$$V(B) = \prod_{i=1}^n \varphi_B(x_i).$$

Como mostra Figura 1.4, $F(2^B)$ é um conjunto clássico. Cada ponto A do hipercubo está ou não está no hiperretângulo $F(2^B)$. No entanto, diferentes pontos A fora do hiperretângulo $F(2^B)$ podem ser considerados como subconjuntos de B com diferentes graus de pertinência, ou seja, podem ser considerados elementos de $F(2^B)$ com um

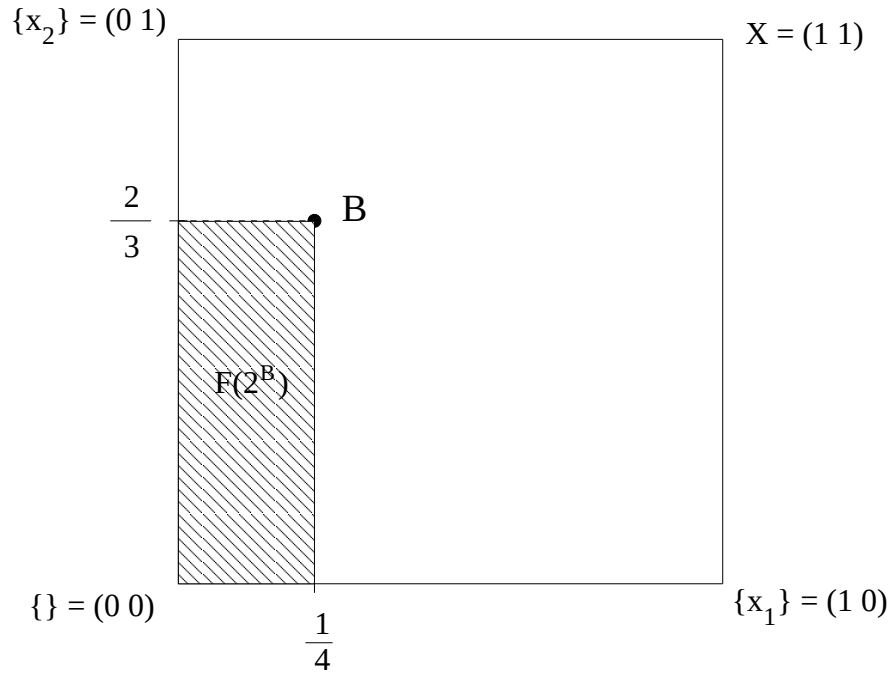


Figura 1.4: Conjunto Potência $F(2^B)$ como Hiperretângulo no Hipercubo Fuzzy

certo grau de pertinência. A definição bivalente de inclusão não leva este fato em conta.

A seguinte generalização é natural e define os subconjuntos fuzzy de B : Conjuntos A distintos pertencerão a $F(2^B)$ com diferentes graus de pertinência. Desta forma, a função de pertinência $\varphi_{F(2^B)}(A)$ pode assumir qualquer valor no intervalo $[0, 1]$. Isso define o chamado grau de inclusão, cuja definição formal é dada a seguir.

Definição 1.6. *O grau de pertinência com que A é um subconjunto de B pode ser definido como*

$$S(A, B) = \text{Grau}(A \subset B) = \varphi_{F(2^B)}(A) = F(2^B)(A),$$

onde $S(A, B)$, chamado de **grau de inclusão**, assume valores no intervalo $[0, 1]$.

Hiperretângulos Puros e Mistos

Sejam $X = \{x_1, x_2, \dots, x_n\}$ um conjunto finito, A e B subconjuntos fuzzy em X . Considere B^* tal que B^* seja o subconjunto de B mais próximo de A no hipercubo fuzzy gerado por $F(2^X)$. É fácil localizar tal conjunto no hipercubo geométrico. Se A é um subconjunto de B , isto é, está no hiperretângulo $F(2^B)$, então $B^* = A$. Supondo, então, que A não seja um subconjunto de B , ele se encontra fora do hiperretângulo $F(2^B)$.

Divide-se o hipercubo unitário I^n em 2^n hiperretângulos por estender os lados de $F(2^B)$ por hiperplanos (Figura 1.5). Os hiperplanos se intersectam perpendicularmente

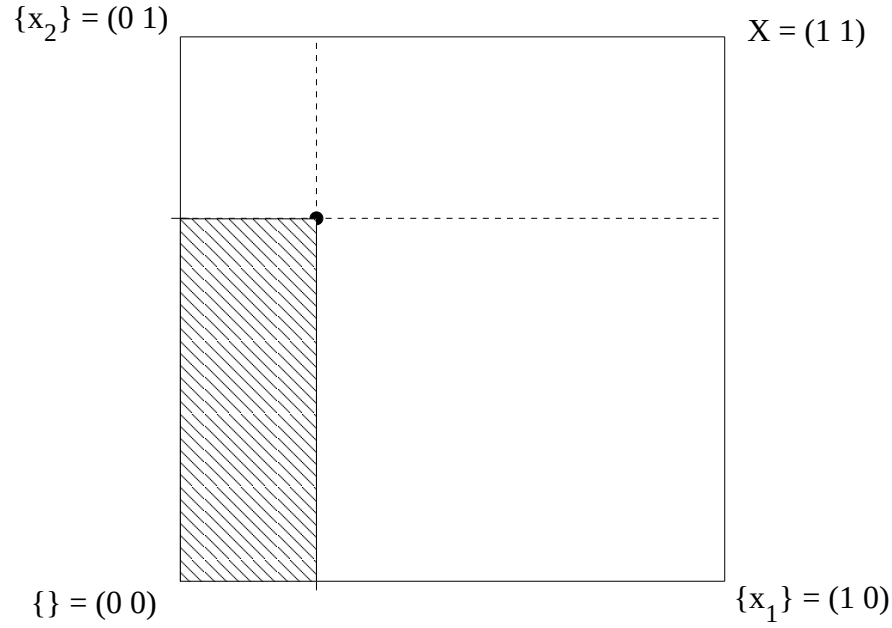


Figura 1.5: Partição do Quadrado Fuzzy

(ortogonalmente). Um dos hiperretângulos é o próprio $F(2^B)$. Para B fixo e A qualquer, os hiperretângulos interiores corresponderão aos 2^n casos se $\varphi_A(x_i) < \varphi_B(x_i)$ ou $\varphi_A(x_i) > \varphi_B(x_i)$. As intersecções dos hiperplanos correspondem ao lugar geométrico dos pontos onde $\varphi_A(x_i) = \varphi_B(x_i)$.

Os 2^n hiperretângulos classificam-se, em relação a inclusão, de acordo com o número de violações da inclusão. Dado A no interior de um hiperretângulo classifica-se o hiperretângulo como

- **puro** se

$$\varphi_A(x_i) < \varphi_B(x_i) \text{ ou } \varphi_A(x_i) > \varphi_B(x_i), \quad \forall x_i \in X;$$

- **misto** se

$$\varphi_A(x_i) < \varphi_B(x_i) \text{ para alguns } x_i \in X$$

e

$$\varphi_A(x_i) > \varphi_B(x_i) \text{ para os demais.}$$

Sendo assim, existem apenas duas inclusões de hiperretângulos puros, a saber, o hiperretângulo que corresponde ao conjunto dos subconjuntos próprios de $F(2^B)$, que inclui o conjunto vazio, e o hiperretângulo que corresponde ao conjunto dos superconjuntos próprios de $F(2^B)$, isto é, ao conjunto de todos os subconjuntos de X que contém propriamente $F(2^B)$ e que, evidentemente, inclui X .

Exemplo 1.7. A Figura 1.6 ilustra como o conjunto potência fuzzy $F(2^B)$ de $B = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ estendido linearmente particiona o quadrado unitário em 2^2 retângulos. Os conjuntos A_1 , A_2 e A_3 que não pertencem a $F(2^B)$ estão em quadrantes distintos. Os

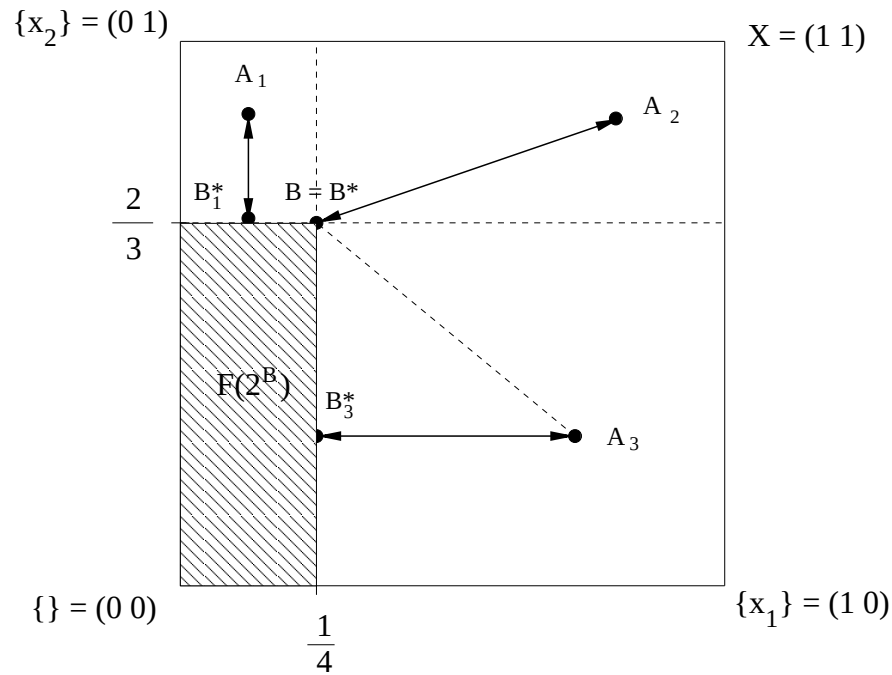


Figura 1.6: Exemplo no Quadrado Fuzzy particionado

quadrantes noroeste e sudeste definem retângulos de inclusões mistas e os quadrantes nordeste e sudoeste definem retângulos de inclusões puras.

Devido à convexidade de $F(2^B)$, no superconjunto hiperretângulo puro o conjunto B^* mais próximo de A_2 é o conjunto B . Para encontrar o conjunto mais próximo B^* no caso misto basta desenhar um segmento perpendicular (ortogonal) de A a $F(2^B)$. Na Figura 1.6 as retas perpendiculares partindo de A_1 e de A_3 intersectam o retângulo $F(2^B)$ nos segmentos gerados pelas coordenadas de B no hipercubo.

Particionando em Três Dimensões

Essa condição de “ortogonalidade” vale também para dimensões maiores. Veja, por exemplo, quando $n = 3$ na Figura 1.7. Para facilitar a visualização, imagine que a sala onde se encontra seja o cubo unitário. Encostando um dicionário num dos quatro cantos do piso, ao qual se faz corresponder à origem, o ponto B equivalerá ao ponto do dicionário mais distante da origem. Estendendo as três faces visíveis do dicionário a sala fica particionada em oito octantes. O dicionário ocupa um octante. Conecta-se pontos dos outros sete octantes aos pontos mais próximos do dicionário por desenhar linhas que intersectem perpendicularmente as três faces visíveis dele.

A visão geométrica dos conjuntos fuzzy e suas implicações serão analisadas em detalhes na segunda parte dessa dissertação.

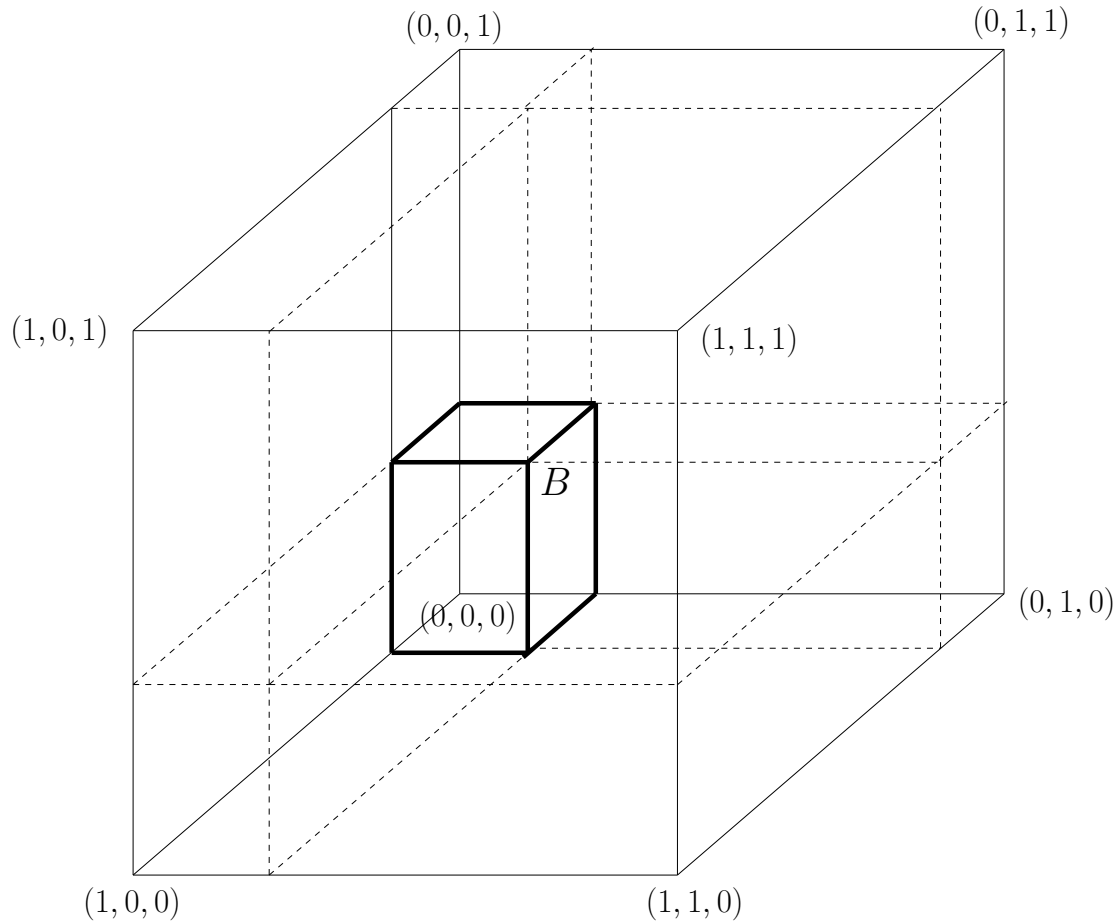


Figura 1.7: Partição do Cubo Fuzzy

1.6 Entropia e Informação

Outro conceito mencionado nesse trabalho é o de entropia. Por isso, essa seção apresentará uma pequena introdução à teoria da informação.

1.6.1 Teoria da Informação

Teoria da informação é um ramo da matemática vinculado à teoria das probabilidades e à estatística que lida com sistemas de comunicação, transmissão de dados, criptografia, codificação, teoria do ruído, correção de erros, compressão de dados, etc. Claude E. Shannon (1916-2001), conhecido como “o pai da teoria da informação”, foi o primeiro a considerar comunicação como um problema matemático rigorosamente embasado na estatística e a dar aos engenheiros da comunicação um modo de determinar a capacidade de um canal de comunicação em termos de *bits*, ou seja, de unidades de informação binária (**binary digits**, em inglês).

Informação é a remoção de uma incerteza. O termo passou a ser utilizado de forma mais ampla a partir da década de 1950. Informação pode ter vários significados, do uso cotidiano ao técnico. Em geral, está intimamente ligada às noções de restrição,

comunicação, controle, dados, forma, instrução, conhecimento, significado, estímulo, padrão, percepção e representação de conhecimento. A **quantidade de informação** de uma mensagem é entendida, na teoria da informação, como o menor número de *bits*, necessários para conter todos os valores ou significados desta mensagem.

1.6.2 Distribuição de Probabilidade

Observação 1.8. Em geral, utiliza-se a palavra vetor associada a uma tabela com uma única coluna e várias linhas. Nesse trabalho, com o objetivo de facilitar a notação, vetores $n \times 1$ serão representados como vetores-linha, ou seja, $1 \times n$.

Definição 1.7. Um **vetor de probabilidades** ou **vetor estocástico** é um vetor composto de entradas positivas, cuja soma é um. Ou seja,

$$p = (p_1 \ p_2 \ \dots \ p_n),$$

onde

$$\sum_{i=1}^n p_i = 1.$$

Cada vetor de probabilidade representa uma determinada **distribuição de probabilidade** registrada em um fenômeno que se deseja estudar.

Exemplo 1.8. São exemplos de vetores de probabilidades $p_1 = (0,20 \ 0,50 \ 0,30)$ e $p_2 = (0,40 \ 0,10 \ 0,30 \ 0,20)$.

As posições (índices) em um vetor de probabilidades representam os possíveis valores de uma variável qualquer, discreta e é uma maneira padrão de caracterizar uma *distribuição de probabilidades discreta*.

Exemplo 1.9. O exemplo clássico na teoria das probabilidades baseada em frequência relativa é uma moeda jogada para cima. Temos 50% de chance de sair “cara” e 50% de chance de sair “coroa”. Essa distribuição de probabilidade pode ser representada pelo vetor $(0,5 \ 0,5)$.

Definição 1.8. Uma distribuição de probabilidade é dita **uniforme** quando todos os valores de seu vetor de probabilidades são iguais.

1.6.3 Entropia

A visão da informação como remoção de uma incerteza surgiu com a publicação, em 1948, do artigo de Claude Shannon, *A Mathematical Theory of Communication* [6]. Este artigo apresenta os fundamentos da teoria da informação e mune a palavra informação não apenas de um significado técnico, mas também de medida. Shannon

foi o primeiro a definir de forma precisa os conceitos de entropia e de informação. Ele partiu do pressuposto de que uma mensagem transmitida e enviada ao receptor pode ser recebida com ruídos, isto é, quando uma informação passa por um canal de comunicação pode sofrer perturbações (ruídos).

Se o dispositivo emissor é igualmente capaz de enviar qualquer um dentre um conjunto de N mensagens, então a medida preferida da “informação produzida quando uma mensagem é escolhida do conjunto” é o logaritmo da base dois de N (esta medida é chamada entropia). Em seu artigo, Shannon afirma:

“A escolha de uma base logarítmica corresponde a escolha de uma unidade para medir a informação. Se a base 2 é usada, as unidades resultantes podem ser chamadas dígitos binários, ou mais resumidamente, bits, uma palavra sugerida por J. W. Tukey. Um dispositivo com duas posições estáveis, tais como um relé ou um circuito flip-flop, pode armazenar um bit de informação. N de tais dispositivos podem armazenar N bits...”

A informação média de uma mensagem é medida pelo que se chama de **entropia da mensagem**, a qual também mede a incerteza que essa mensagem remove. No processo de desenvolvimento de uma teoria da comunicação que pudesse ser aplicada por engenheiros eletricitas para projetar sistemas de telecomunicação, Shannon definiu a entropia de uma fonte de informação ou distribuição de probabilidade de X , para X finito, como

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x),$$

que determina a incerteza média que pode ser removida pelo conhecimento de X e é usada para definir o que se entende por capacidade do canal.

A entropia de Shannon passou a ser considerada como uma medida da informação contida numa mensagem. A entropia aumenta quando a certeza diminui e vice-versa. Ou seja, se tudo é conhecido a respeito de uma determinada mensagem a entropia dela é zero. Por outro lado, se pouco se sabe sobre ela sua entropia é máxima. A distribuição uniforme maximiza, de modo único, a entropia, ou seja, se $\sum_{i=1}^n p(x_i) = 1$

para $X = \{x_1, x_2, \dots, x_n\}$ e $p(x_i) = \frac{1}{n}$ então $H(X)$ é máxima. Na segunda parte dessa dissertação será analisada uma distribuição de probabilidade no hipercubo I^n relacionada à maximização da entropia fuzzy, e que mostra a associação entre entropia e a visão geométrica dos conjuntos fuzzy.

1.7 Simplexos

Outro conceito a ser mencionado nesse trabalho é o de simplexo, associado a uma particular distribuição de probabilidade. Em geometria, chama-se simplexo a uma generalização da noção de triângulo para uma dimensão arbitrária. Apresenta-se abaixo a definição formal.

Definição 1.9. *Um n -simplexo, indicado por S^n , é um poliedro n -dimensional com $n + 1$ vértices.*

Exemplo 1.10. Os simplexos mais conhecidos são:

- Triângulo: um 2-simplexo.
- Tetraedro: um 3-simplexo.
- Ponto: um 0-simplexo.
- Segmento de reta: um 1-simplexo.

Observação 1.9. De modo geral, chama-se de simplexo ao menor conjunto convexo contendo todos os vértices dados.

Observação 1.10. Note, ainda que nas dimensões menores os simplexos são degenerados, uma vez que um ponto e um segmento não somente podem ser comparados a um triângulo, mas a qualquer figura geométrica.

A seguir será apresentado um breve resumo da teoria dos conjuntos fuzzy com o objetivo de familiarizar o leitor com as operações sobre tais conjuntos.

2 Aritmética e Lógica Fuzzy

Este capítulo dará uma ideia de como operar com os conjuntos fuzzy através de sua função de pertinência. A fim de simplificar a linguagem fica estabelecido que, quando o contexto permitir um conjunto será mencionado simplesmente como “conjunto fuzzy”, seja ele clássico ou não.

2.1 Suporte de um Conjunto Fuzzy

Definição 2.1. *Seja F um conjunto fuzzy. Chama-se **suporte de F** , e indica-se por $\text{supp } F$, ao conjunto*

$$\text{supp } F = \{x \in X : \varphi_F(x) > 0\}.$$

Observação 2.1. O suporte sempre coincide com o conjunto no caso dos conjuntos clássicos. No entanto, isto não é verdade para os conjuntos fuzzy em geral.

Exemplo 2.1. Sejam $X = \mathbb{R}$ e $F = \{x \in X : x \text{ é distante de } 3\}$ dado pela função de pertinência (veja Figura 2.1).

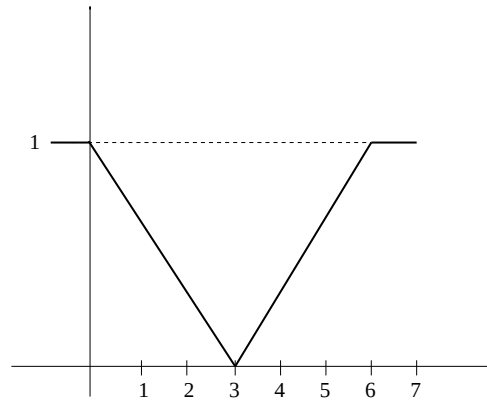


Figura 2.1: Gráfico do Conjunto Fuzzy definido pela Função de Pertinência do Exemplo 2.1

$$\varphi_F(x) = \begin{cases} \frac{|3-x|}{3} & \text{se } |3-x| \leq 3 \\ 1 & \text{se } |3-x| > 3 \end{cases}.$$

Então $\text{supp } F = \{x \in X : x \neq 3\}$.

2.2 Operações com Conjuntos Fuzzy

Definição 2.2. *Sejam A e B dois subconjuntos fuzzy em X . Destacam-se os seguintes subconjuntos fuzzy em X :*

1. A **união** entre A e B ($A \cup B$) cuja função de pertinência é dada por:

$$\varphi_{A \cup B}(x) = \max(\varphi_A(x), \varphi_B(x)), \text{ para todo } x \in X;$$

2. A **intersecção** entre A e B ($A \cap B$) cuja função de pertinência é dada por:

$$\varphi_{A \cap B}(x) = \min(\varphi_A(x), \varphi_B(x)), \text{ para todo } x \in X;$$

3. O **complementar** de A (A^c) cuja função de pertinência é dada por:

$$\varphi_{A^c}(x) = 1 - \varphi_A(x), \text{ para todo } x \in X.$$

Definição 2.3. *Os subconjuntos fuzzy A e B em X são **iguais** se suas funções de pertinência coincidem, isto é, se $\varphi_A(x) = \varphi_B(x)$, para todo $x \in X$.*

Exemplo 2.2. Seja $X = [0, 10]$. Considere as funções $\varphi_A(x)$ e $\varphi_B(x)$ dadas por

$$\varphi_A(x) = \begin{cases} 0,5 & \text{se } x \geq 1 \\ 0,3 & \text{se } x < 1 \end{cases} \quad e$$

$$\varphi_B(x) = \begin{cases} 0,4 & \text{se } x \geq 8 \\ 0,6 & \text{se } x < 8 \end{cases}.$$

Então:

$$\varphi_{A \cap B}(x) = \begin{cases} 0,4 & \text{se } x \geq 8 \\ 0,5 & \text{se } 1 \leq x < 8 \\ 0,3 & \text{se } x < 1 \end{cases},$$

$$\varphi_{A \cup B}(x) = \begin{cases} 0,5 & \text{se } x \geq 8 \\ 0,6 & \text{se } x < 8 \end{cases},$$

$$\varphi_{A^c}(x) = \begin{cases} 0,5 & \text{se } x \geq 1 \\ 0,7 & \text{se } x < 1 \end{cases} \quad e$$

$$\varphi_{B^c}(x) = \begin{cases} 0,6 & \text{se } x \geq 8 \\ 0,4 & \text{se } x < 8 \end{cases}.$$

Proposição 2.1. As operações entre conjuntos fuzzy satisfazem as seguintes propriedades usuais de operações de conjuntos:

- (a) $A \cup B = B \cup A$
- (b) $A \cap B = B \cap A$
- (c) $A \cup (B \cup C) = (A \cup B) \cup C$
- (d) $A \cap (B \cap C) = (A \cap B) \cap C$
- (e) $A \cup A = A$
- (f) $A \cap A = A$
- (g) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- (h) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- (i) $A \cap \emptyset = \emptyset$ e $A \cup \emptyset = A$
- (j) $A \cap X = A$ e $A \cup X = X$
- (k) $(A \cup B)^c = A^c \cap B^c$ e $(A \cap B)^c = A^c \cup B^c$ (*leis de De Morgan*)

Demonstração. Para todo $x \in X$ vale:

- (a) $\varphi_{A \cup B}(x) = \max(\varphi_A(x), \varphi_B(x)) = \max(\varphi_B(x), \varphi_A(x)) = \varphi_{B \cup A}(x)$ e, portanto, $\varphi_{A \cup B} = \varphi_{B \cup A}$.

Logo, $A \cup B = B \cup A$;

- (b) $\varphi_{A \cap B}(x) = \min(\varphi_A(x), \varphi_B(x)) = \min(\varphi_B(x), \varphi_A(x)) = \varphi_{B \cap A}(x)$ e, portanto, $\varphi_{A \cap B} = \varphi_{B \cap A}$.

Logo, $A \cap B = B \cap A$;

- (c) $\varphi_{A \cup (B \cup C)}(x) = \max(\varphi_A(x), \varphi_{B \cup C}(x)) = \max(\varphi_A(x), \max(\varphi_B(x), \varphi_C(x))) = \max(\varphi_A(x), \varphi_B(x), \varphi_C(x)) = \max(\max(\varphi_A(x), \varphi_B(x)), \varphi_C(x)) = \max(\varphi_{A \cup B}(x), \varphi_C(x)) = \varphi_{(A \cup B) \cup C}(x)$ e, portanto, $\varphi_{A \cup (B \cup C)}(x) = \varphi_{(A \cup B) \cup C}(x)$.

Logo, $A \cup (B \cup C) = (A \cup B) \cup C$;

- (d) $\varphi_{A \cap (B \cap C)}(x) = \min(\varphi_A(x), \varphi_{B \cap C}(x)) = \min(\varphi_A(x), \min(\varphi_B(x), \varphi_C(x))) = \min(\varphi_A(x), \varphi_B(x), \varphi_C(x)) = \min(\min(\varphi_A(x), \varphi_B(x)), \varphi_C(x)) = \min(\varphi_{A \cap B}(x), \varphi_C(x)) = \varphi_{(A \cap B) \cap C}(x)$ e, portanto, $\varphi_{A \cap (B \cap C)}(x) = \varphi_{(A \cap B) \cap C}(x)$.

Logo, $A \cap (B \cap C) = (A \cap B) \cap C$;

(e) $\varphi_{A \cup A}(x) = \max(\varphi_A(x), \varphi_A(x)) = \varphi_A(x)$ e, portanto,

$$\varphi_{A \cup A}(x) = \varphi_A(x).$$

Logo, $A \cup A = A$;

(f) $\varphi_{A \cap A}(x) = \min(\varphi_A(x), \varphi_A(x)) = \varphi_A(x)$, e, portanto, $\varphi_{A \cap A}(x) = \varphi_A(x)$.

Logo, $A \cap A = A$;

(g) $\varphi_{A \cap (B \cup C)}(x) = \min(\varphi_A(x), \varphi_{B \cup C}(x))$
 $= \min(\varphi_A(x), \max(\varphi_B(x), \varphi_C(x))) = \max(\min(\varphi_A(x), \varphi_B(x)), \min(\varphi_A(x), \varphi_C(x))) =$
 $\varphi_{(A \cap B) \cup (A \cap C)}(x)$ e, portanto, $\varphi_{A \cap (B \cup C)}(x) = \varphi_{(A \cap B) \cup (A \cap C)}(x)$.

Logo, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

(h) $\varphi_{A \cup (B \cap C)}(x) = \max(\varphi_A(x), \varphi_{B \cap C}(x))$
 $= \max(\varphi_A(x), \min(\varphi_B(x), \varphi_C(x))) = \min(\max(\varphi_A(x), \varphi_B(x)), \max(\varphi_A(x), \varphi_C(x))) =$
 $\varphi_{(A \cup B) \cap (A \cup C)}(x)$ e, portanto, $\varphi_{A \cup (B \cap C)}(x) = \varphi_{(A \cup B) \cap (A \cup C)}(x)$.

Logo, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

(i) $\varphi_{A \cap \emptyset}(x) = \min(\varphi_A(x), \varphi_{\emptyset}(x)) = \min(\varphi_A(x), 0) = 0 = \varphi_{\emptyset}(x)$ e, portanto, $\varphi_{A \cap \emptyset}(x) = \varphi_{\emptyset}(x)$.

Logo, $A \cap \emptyset = \emptyset$;

e

$\varphi_{A \cup \emptyset}(x) = \max(\varphi_A(x), \varphi_{\emptyset}(x)) = \max(\varphi_A(x), 0) = \varphi_A(x)$ e, portanto, $\varphi_{A \cup \emptyset}(x) = \varphi_A(x)$.

Logo, $A \cup \emptyset = A$;

(j) $\varphi_{A \cap X}(x) = \min(\varphi_A(x), \varphi_X(x)) = \min(\varphi_A(x), 1) = \varphi_A(x)$ e, portanto, $\varphi_{A \cap X}(x) = \varphi_A(x)$.

Logo, $A \cap X = X$;

e

$\varphi_{A \cup X}(x) = \max(\varphi_A(x), \varphi_X(x)) = \max(\varphi_A(x), 1) = \varphi_X(x)$ e, portanto, $\varphi_{A \cup X}(x) = \varphi_X(x)$.

Logo, $A \cup X = X$;

(k) $\varphi_{(A \cup B)^c}(x) = 1 - \varphi_{A \cup B}(x) = 1 - \max(\varphi_A(x), \varphi_B(x)) = (*)$.

Neste caso,

– se $\varphi_A(x) > \varphi_B(x)$ então $1 - \varphi_A(x) < 1 - \varphi_B(x)$. Daí, $(*) = 1 - \varphi_A(x) = \varphi_{A^c}(x)$;

- se $\varphi_A(x) < \varphi_B(x)$ então $1 - \varphi_A(x) > 1 - \varphi_B(x)$. Daí, $(*) = 1 - \varphi_B(x) = \varphi_{B^c}(x)$;
- se $\varphi_A(x) = \varphi_B(x)$ então $(*) = 1 - \varphi_A(x) = 1 - \varphi_B(x) = \varphi_{B^c}(x) = \varphi_{A^c}(x)$.

Em todos os casos $(*) = \min(\varphi_{A^c}(x), \varphi_{B^c}(x)) = \varphi_{A^c \cap B^c}(x)$.

Logo, $(A \cup B)^c = A^c \cap B^c$;

e

$$\varphi_{(A \cap B)^c}(x) = 1 - \varphi_{A \cap B}(x) = 1 - \min(\varphi_A(x), \varphi_B(x)) = (**).$$

Neste caso,

- se $\varphi_A(x) > \varphi_B(x)$ então $1 - \varphi_A(x) < 1 - \varphi_B(x)$. Daí, $(**) = 1 - \varphi_B(x) = \varphi_{B^c}(x)$;
- se $\varphi_A(x) < \varphi_B(x)$ então $1 - \varphi_A(x) > 1 - \varphi_B(x)$. Daí, $(**) = 1 - \varphi_A(x) = \varphi_{A^c}(x)$;
- se $\varphi_A(x) = \varphi_B(x)$ então $(**) = 1 - \varphi_A(x) = 1 - \varphi_B(x) = \varphi_{B^c}(x) = \varphi_{A^c}(x)$.

Em todos os casos $(**) = \max(\varphi_{A^c}(x), \varphi_{B^c}(x)) = \varphi_{A^c \cup B^c}(x)$.

Logo, $(A \cap B)^c = A^c \cup B^c$.

□

Observação 2.2. Nem todas as propriedades das operações usuais sobre conjuntos clássicos são satisfeitas pelos conjuntos fuzzy que não são clássicos. Por exemplo, nos conjuntos clássicos sempre vale que $A \cap A^c = \emptyset$. No entanto, nos conjuntos fuzzy não-clássicos pode ocorrer: $A \cap A^c \neq \emptyset$.

Exemplo 2.3. Considere $X = [0, 10]$. Sejam as funções dadas por

$$\begin{aligned}\varphi_A(x) &= \begin{cases} 0,5 & \text{se } x \geq 1 \\ 0,3 & \text{se } x < 1 \end{cases}, \\ \varphi_B(x) &= \begin{cases} 0,4 & \text{se } x \geq 8 \\ 0,6 & \text{se } x < 8 \end{cases}, \\ \varphi_C(x) &= \begin{cases} 0,2 & \text{se } x \geq 4 \\ 0,7 & \text{se } x < 4 \end{cases} \quad e \\ \varphi_D(x) &= \begin{cases} 0,1 & \text{se } x \geq 6 \\ 0,8 & \text{se } x < 6 \end{cases}.\end{aligned}$$

Ilustrando algumas das propriedades apresentadas na Proposição 2.1, obtém-se, por exemplo:

$$1. A \cup (B \cup C) = (A \cup B) \cup C$$

$$\begin{aligned}\varphi_{A \cup B}(x) &= \begin{cases} 0,5 & \text{se } x \geq 8 \\ 0,6 & \text{se } x < 8 \end{cases} e \\ \varphi_{B \cup C}(x) &= \begin{cases} 0,4 & \text{se } x \geq 8 \\ 0,6 & \text{se } 4 \leq x < 8 \\ 0,7 & \text{se } x < 4 \end{cases} .\end{aligned}$$

Então:

$$\begin{aligned}\varphi_{A \cup (B \cup C)}(x) &= \begin{cases} 0,5 & \text{se } x \geq 8 \\ 0,6 & \text{se } 4 \leq x < 8 \\ 0,7 & \text{se } x < 4 \end{cases} e \\ \varphi_{(A \cup B) \cup C}(x) &= \begin{cases} 0,5 & \text{se } x \geq 8 \\ 0,6 & \text{se } 4 \leq x < 8 \\ 0,7 & \text{se } x < 4 \end{cases} .\end{aligned}$$

$$2. A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$\begin{aligned}\varphi_{B \cap C}(x) &= \begin{cases} 0,2 & \text{se } x \geq 4 \\ 0,6 & \text{se } x < 4 \end{cases} , \\ \varphi_{A \cup B}(x) &= \begin{cases} 0,5 & \text{se } x \geq 8 \\ 0,6 & \text{se } x < 8 \end{cases} e \\ \varphi_{A \cup C}(x) &= \begin{cases} 0,5 & \text{se } x \geq 4 \\ 0,7 & \text{se } x < 4 \end{cases} .\end{aligned}$$

Então:

$$\begin{aligned}\varphi_{A \cup (B \cap C)}(x) &= \begin{cases} 0,5 & \text{se } x \geq 4 \\ 0,6 & \text{se } x < 4 \end{cases} e \\ \varphi_{(A \cup B) \cap (A \cup C)}(x) &= \begin{cases} 0,5 & \text{se } x \geq 4 \\ 0,6 & \text{se } x < 4 \end{cases} .\end{aligned}$$

Note ainda que:

$$A \cap A^c \neq \emptyset,$$

pois, tem-se que

$$\varphi_{A^c}(x) = \begin{cases} 0,5 & \text{se } x \geq 1 \\ 0,7 & \text{se } x < 1 \end{cases} .$$

Daí,

$$\varphi_{A \cap A^c}(x) = \begin{cases} 0,5 & \text{se } x \geq 1 \\ 0,3 & \text{se } x < 1 \end{cases},$$

que é diferente de:

$$\varphi_{\emptyset}(x) = 0, \forall x \in X.$$

2.3 O conceito de α -nível

Definição 2.4. *Sejam A um subconjunto fuzzy em X e $\alpha \in [0, 1]$. O α -nível de A é o subconjunto clássico de X definido por:*

$$[A]^\alpha = \{x \in X : \varphi_A(x) \geq \alpha\} \text{ para } 0 < \alpha \leq 1.$$

Observação 2.3. Para subconjuntos fuzzy A em $\mathbb{R}$ define-se também o nível zero de A como sendo o MENOR subconjunto clássico de $\mathbb{R}$, fechado (com a topologia usual da reta), que contém o conjunto suporte de A . Diz-se que $[A]^0$ é o fecho de $\text{supp } A$ e indica-se $\overline{\text{supp } A}$. Vale destacar que o conjunto $\{x \in X : \varphi_A(x) \geq 0\} = X$ não é necessariamente igual a $[A]^0 = \overline{\text{supp } A}$.

Exemplo 2.4. Seja A um subconjunto fuzzy em $\mathbb{R}$, com a seguinte função de pertinência:

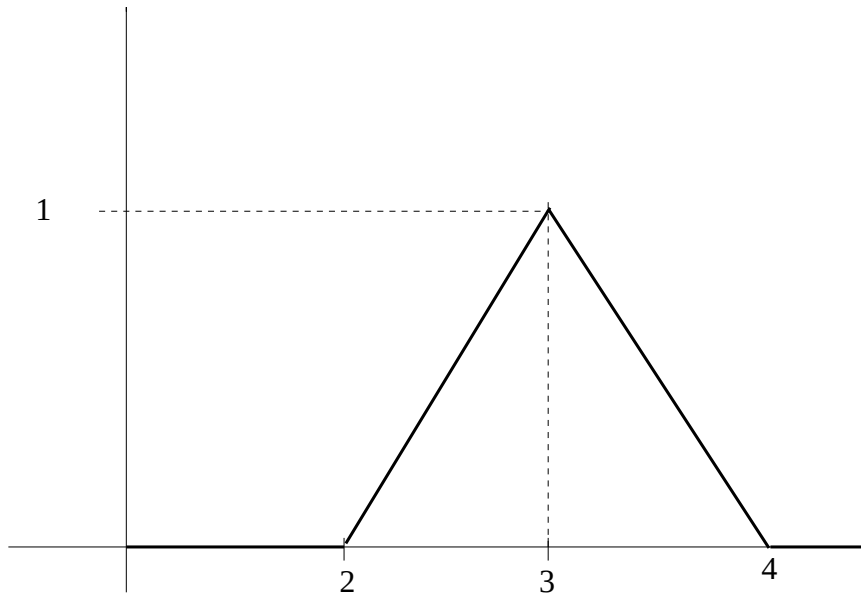


Figura 2.2: Gráfico do Conjunto Fuzzy do Exemplo 2.4

$$\varphi_A(x) = \begin{cases} x - 2 & \text{se } 2 \leq x \leq 3 \\ 4 - x & \text{se } 3 < x < 4 \\ 0 & \text{se } x \notin [2, 4) \end{cases}.$$

O objetivo, pela definição de $[A]^\alpha$ é encontrar os valores de x tais que $\varphi_A(x) \geq \alpha$, ou seja,

$$x - 2 \geq \alpha \iff x \geq \alpha + 2 \text{ e}$$

$$4 - x \geq \alpha \iff -x \geq \alpha - 4 \iff x \leq 4 - \alpha.$$

Então,

$$[A]^\alpha = [\alpha + 2, 4 - \alpha] \text{ para } 0 < \alpha \leq 1 \text{ e } [A]^0 = \overline{]2, 4[} = [2, 4] \subsetneq \mathbb{R}.$$

Exemplo 2.5. Sejam $X = [1, 3] \subset \mathbb{R}$ e A o subconjunto fuzzy em X cuja função de pertinência é $\varphi_A(x) = -x^2 + 4x - 3$.

Pela definição de $[A]^\alpha$ busca-se os valores de x tais que:

$$\alpha \leq \varphi_A(x) = -x^2 + 4x - 3,$$

ou seja, as soluções da desigualdade

$$-x^2 + 4x - 3 - \alpha \geq 0.$$

As raízes de $-x^2 + 4x - 3 = 0$ (*) são:

$$x = 2 \pm \sqrt{1 - \alpha}.$$

Como o coeficiente de x^2 é negativo $\alpha \leq \varphi_A(x) = -x^2 + 4x - 3$ vale entre as raízes da equação (*). Portanto:

$$[A]^\alpha = [2 - \sqrt{1 - \alpha}, 2 + \sqrt{1 - \alpha}] \text{ para todo } \alpha \in [0, 1].$$

Observação 2.4. Percebe-se destes exemplos que, se x é um elemento de $[A]^\alpha$, então x pertence ao conjunto fuzzy A com, no mínimo, grau α . Tem-se, então, os resultados abaixo.

Proposição 2.2. Seja A um subconjunto fuzzy em X . Se $\alpha < \beta$ então $[A]^\beta \subset [A]^\alpha$.

Demonstração. $x \in [A]^\beta$ se, e somente se, $\{x \in X : \varphi_A(x) \geq \beta\}$ o que implica, por hipótese, que $\{x \in X : \varphi_A(x) \geq \alpha\}$ se, e somente se, $x \in [A]^\alpha$. $\square$

Teorema 2.1. Sejam A e B subconjuntos fuzzy em X . Então,

$$A = B \text{ se, e somente se, } [A]^\alpha = [B]^\alpha, \forall \alpha \in [0, 1].$$

Demonstração. Se $A = B$ a demonstração é imediata a partir da definição de $[A]^\alpha$. Suponha, para a recíproca, que $[A]^\alpha = [B]^\alpha, \forall \alpha \in [0, 1]$. Se $A \neq B$ então existe $x \in X$ tal que $\varphi_A(x) \neq \varphi_B(x)$. Daí, $\varphi_A(x) < \varphi_B(x)$ ou $\varphi_A(x) > \varphi_B(x)$. Supondo $\varphi_A(x) > \varphi_B(x)$, conclui-se que $x \in [A]^{\varphi_A(x)}$ e $x \notin [B]^{\varphi_A(x)}$ e, portanto, $[A]^{\varphi_A(x)} \neq [B]^{\varphi_A(x)}$ o que contradiz a hipótese $[A]^\alpha = [B]^\alpha, \forall \alpha \in [0, 1]$. Analogamente, chega-se a uma contradição se $\varphi_A(x) < \varphi_B(x)$. $\square$

Corolário 2.1. *A função de pertinência φ_A de um conjunto fuzzy A pode ser expressa por meio das funções características de seus α -níveis como abaixo.*

$$\varphi_A(x) = \sup_{\alpha} \min(\alpha, \chi_{[A]^\alpha}(x)), \text{ onde } \chi_{[A]^\alpha}(x) = \begin{cases} 1 & \text{se } x \in [A]^\alpha \\ 0 & \text{se } x \notin [A]^\alpha \end{cases},$$

com $\alpha \in [0, 1]$.

Demonstração. Considere $x \in X$, fixo. Para simplificar a notação sejam $m(\alpha) = \min(\alpha, \chi_{[A]^\alpha}(x))$ e $\varphi_A(x) = \beta$.

Dado $\alpha \in [0, 1]$ tem-se que, se $x \notin [A]^\alpha$ então $\chi_{[A]^\alpha}(x) = 0$ e $m(\alpha) = 0$. Adicionalmente, se $x \in [A]^\alpha$ então $\chi_{[A]^\alpha}(x) = 1$ e $m(\alpha) = \alpha$, ou seja,

$$m(\alpha) = \begin{cases} \alpha & \text{se } x \in [A]^\alpha \\ 0 & \text{se } x \notin [A]^\alpha \end{cases}.$$

Observe que, se $\beta < \alpha$ então $m(\alpha) = 0$ e, se $\beta \geq 0$, então $m(\alpha) = \min(\alpha, \chi_{[A]^\alpha}(x)) = \alpha$. Logo, para todo $\alpha \in [0, 1]$ tem-se que

$$\sup_{\alpha} \min(\alpha, \chi_{[A]^\alpha}(x)) = \sup_{\alpha} m(\alpha) = \sup_{\alpha} \{\alpha\} = \sup_{\beta \geq \alpha} \{\alpha\} = \beta = \varphi_A(x).$$

$\square$

Proposição 2.3.

$$a) [A \cup B]^\alpha = [A]^\alpha \cup [B]^\alpha,$$

$$b) [A \cap B]^\alpha = [A]^\alpha \cap [B]^\alpha.$$

Demonstração. Para o item (a) basta notar que a função $\varphi_{A \cup B}(x)$ pode ser escrita como

$$\varphi_{A \cup B}(x) = \begin{cases} \varphi_A(x) & \text{se } \varphi_A(x) \geq \varphi_B(x) \\ \varphi_B(x) & \text{se } \varphi_A(x) < \varphi_B(x) \end{cases}.$$

Então

$$x \in [A \cup B]^\alpha \text{ se, e somente se, } \varphi_{A \cup B}(x) \geq \alpha \text{ se, e somente se, } \varphi_A(x) \geq \alpha \vee \varphi_B(x) \geq \alpha \text{ se, e somente se, } x \in [A]^\alpha \cup [B]^\alpha.$$

Para o item b escreve-se a função $\varphi_{A \cap B}(x)$ como

$$\varphi_{A \cap B}(x) = \begin{cases} \varphi_A(x) & \text{se } \varphi_A(x) \leq \varphi_B(x) \\ \varphi_B(x) & \text{se } \varphi_A(x) > \varphi_B(x) \end{cases}.$$

Então

$x \in [A \cap B]^\alpha$ se, e somente se, $\varphi_{A \cap B}(x) \geq \alpha$ se, e somente se,
 $\varphi_A(x) \geq \alpha \wedge \varphi_B(x) \geq \alpha$ se, e somente se, $x \in [A]^\alpha \cap [B]^\alpha$.

□

2.4 Lógica Fuzzy

Neste trabalho serão mencionadas as operações t-normas e t-conormas. Estas são extensões das operações “e” e “ou” da lógica clássica. Nesta seção apresenta-se as definições formais destas operações e uma pequena introdução à lógica fuzzy.¹

2.4.1 Conectivos da Lógica Clássica

Na lógica matemática sentencial clássica são estudados os conectivos fundamentais “e”, “ou”, “não” e “implicação”. O valor lógico destes conectivos são calculados através de tabelas-verdade. Na lógica clássica uma sentença tem valor verdadeiro (1) ou falso (0). Para a extensão para o caso fuzzy são usadas notações similares as da lógica clássica, a saber, $\wedge$ (mínimo) para a conjunção e, $\vee$ (máximo) para a disjunção “ou”, $\sim$ para a negação e $\longrightarrow$ para a implicação. A Tabela 2.1 apresenta os valores destes conectivos na lógica clássica.²

Formalmente, diz-se que cada uma destas tabela representa uma função de

$$\{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}.$$

2.4.2 Conectivos Básicos da Lógica Fuzzy

Na lógica fuzzy os conectivos não assumirão apenas os valores 0 e 1, mas qualquer um dos valores do intervalo $[0, 1]$. Atribui-se, assim, um valor entre 0 e 1 a uma proposição, como por exemplo, “ x pertence a A ”. Portanto, para efetuar a avaliação lógica dos conectivos fuzzy estende-se os conectivos da lógica clássica. Esta seção mostra como fazer isto.

¹Para maiores detalhes sobre lógica fuzzy veja *A First Course in Fuzzy Logic*. [7]

²Para um estudo mais detalhado da lógica clássica veja *Lógica Uma introdução voltada para as ciências* [8], *Lógica e linguagem cotidiana* [9] e *A Mathematical Introduction to Logic*. [10]

p	q	$p \wedge q$
V	V	V
V	F	F
F	V	F
F	F	F

p	q	$p \vee q$
V	V	V
V	F	V
F	V	V
F	F	F

p	$\sim p$
V	F
F	V

p	q	$p \longrightarrow q$
V	V	V
V	F	F
F	V	V
F	F	V

Tabela 2.1: Conectivos da Lógica Clássica

t-norma

Definição 2.5. O operador $\Delta : [0, 1] \times [0, 1] \rightarrow [0, 1]$, $\Delta(x, y) = x \Delta y$, é uma **t-norma**, se satisfaz as seguintes condições:

- $t_1)$ elemento neutro: $\Delta(1, x) = 1 \Delta x = x$;
- $t_2)$ comutativa: $\Delta(x, y) = x \Delta y = y \Delta x = \Delta(y, x)$;
- $t_3)$ associativa: $x \Delta (y \Delta z) = (x \Delta y) \Delta z$;
- $t_4)$ monotonicidade: se $x \leq u$ e $y \leq v$, então $x \Delta y \leq u \Delta v$.

Observação 2.5. O operador **t-norma** estende o operador $\wedge$. Note que $\Delta \mid \{0, 1\} \times \{0, 1\} = \wedge$ é a conjunção clássica “e”.

Exemplo 2.6. São exemplos de **t-normas** ³

1. $x \Delta_0 y = \begin{cases} x \wedge y & \text{se } x \vee y = 1 \\ 0 & \text{caso contrário} \end{cases}$,
2. $x \Delta_1 y = \max(0, x + y - 1) = 0 \vee (x + y - 1)$,
3. $x \Delta_2 y = \frac{xy}{2 - (x + y - xy)}$,
4. $x \Delta_3 y = xy$,
5. $x \Delta_4 y = \frac{xy}{x + y - xy}$,
6. $x \Delta_5 y = \min(x, y) = x \wedge y$.

³Para detalhes veja *A First Course in Logic Fuzzy*. [7]

Proposição 2.4. Se Δ é uma t-norma, então para $x, y \in [0, 1]$,

$$x \Delta_0 y \leq x \Delta y \leq x \Delta_5 y.^4$$

Observação 2.6. A t-norma $\wedge$ é a única **idempotente**, ou seja, é a única t-norma tal que $x \Delta x = x$. Para uma t-norma Δ qualquer tem-se que $x \Delta x$ nunca é maior do que x , isto é, $x \leq x \Delta x$. Para qualquer t-norma onde $\Delta \neq \wedge$ teremos $x < x \Delta x$.

Definição 2.6. Uma t-norma Δ é dita **convexa** se toda vez que $x \Delta y \leq c \leq x_1 \Delta y_1$, então existe um r entre x e x_1 e um s entre y e y_1 tal que $c = r \Delta s$.

Corolário 2.2. Se uma t-norma Δ é convexa e $a < b$, então há um $c \in [a, 1]$ tal que $a = b \Delta c$.

Demonstração. $a \Delta b \leq 1 \Delta a = a < 1 \Delta b$, então, por convexidade, existe tal c . $\square$

Negação

O conjunto fuzzy A^c , complementar do conjunto fuzzy A foi definido como $A^c(x) = 1 - A(x)$. Isso equivale a rastrear A pela função $[0, 1] \rightarrow [0, 1] : x \rightarrow 1 - x$. Essa função é uma **involução** do reticulado $R = ([0, 1], \leq)$. Em lógica fuzzy, tal aplicação é dita **negação forte**.

Definição 2.7. Uma aplicação $\eta : [0, 1] \rightarrow [0, 1]$ é uma **negação forte** se satisfaz as seguintes condições:

n_1) fronteiras: $\eta(0) = 1$ e $\eta(1) = 0$;

n_2) involução: $\eta(\eta(x)) = x$;

n_3) monotonicidade: η é não crescente.

Se ela satisfaz somente as duas primeiras condições é dita **negação**.

t-conorma

Definição 2.8. O operador $\nabla : [0, 1] \times [0, 1] \rightarrow [0, 1]$, $\nabla(x, y) = x \nabla y$, é uma **t-conorma**, se satisfaz as seguintes condições:

c_1) elemento neutro: $\nabla(1, x) = 1 \nabla x = x$;

c_2) comutativa: $\nabla(x, y) = x \nabla y = y \nabla x = \nabla(y, x)$;

c_3) associativa: $x \nabla (y \nabla z) = (x \nabla y) \nabla z$;

c_4) monotonicidade: se $x \leq u$ e $y \leq v$, então $x \nabla y \leq u \nabla v$.

Observação 2.7. O operador **t-conorma** estende o operador $\vee$. Note que $\nabla \mid \{0, 1\} \times \{0, 1\} = \vee$ é a disjunção clássica “ou”.

⁴Idem.

2.4.3 Implicações Fuzzy

Definição 2.9. Qualquer operação $\longrightarrow: [0, 1] \times [0, 1] \rightarrow [0, 1]$ que estenda a tabela verdade da implicação clássica é chamada **implicação fuzzy**.

Destaca-se, abaixo, três tipos de implicações fuzzy:

1. *S-implicação*: $(x \longrightarrow y) = \eta(x) \nabla y$;
2. *Q-implicação*: $(x \longrightarrow y) = \eta(x) \nabla (x \triangle y)$;
3. *R-implicação*: $(x \longrightarrow y) = \sup\{z \in [0, 1] : x \triangle z \leq y\}$.

Exemplo 2.7. São exemplos de implicações fuzzy: ⁵

1. Implicação de Gaines-Rescher:

$$(x \rightarrow y) = g_r(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ 0 & \text{se } x > y \end{cases}.$$

2. Implicação de Gödel:

$$(x \rightarrow y) = g(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ y & \text{se } x > y \end{cases}.$$

3. Implicação de Goguen:

$$(x \rightarrow y) = g_n(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ \frac{y}{x} & \text{se } x > y \end{cases}.$$

4. Implicação de Lukasiewicz:

$$(x \rightarrow y) = \ell(x, y) = \min(1 - x + y, 1).$$

5. Implicação de Kleene-Dienes:

$$(x \rightarrow y) = k_d(x, y) = \max(1 - x, y).$$

6. Implicação de Reichenbach:

$$(x \rightarrow y) = r(x, y) = (1 - x + xy) \vee 0.$$

⁵Ibidem.

7. Implicação de Wu:

$$(x \rightarrow y) = w(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ \min(1 - x, y) & \text{se } x > y \end{cases}.$$

8. Implicação de Zadeh:

$$(x \rightarrow y) = k_d(x, y) = \max(1 - x, \min(x, y)).$$

Conforme mencionado no Capítulo 1, uma família de conjuntos fuzzy de grande importância são os chamados números fuzzy. A seguir serão apresentadas sua definição formal, sua representação usual e como operar com eles.

3 Números Fuzzy

Em muitos problemas concretos os números são idealizações de informações imprecisas que envolvem valores concretos, por exemplo, quando se mede a altura de várias pessoas com o objetivo de montar uma tabela. [11] Em tais situações valores imprecisos surgem a todo instante, seja pela limitação de instrumentos de medida, quando se decide se um indivíduo é alto ou baixo, por limitações na tabela onde os dados devem ser inseridos, entre outras. Após a medição é necessário classificar a pessoa por altura e colocar as informações em algum lugar da tabela. Escolhe-se, então, o valor mais próximo da altura da pessoa e classifica-se sua altura de acordo com ele. Esse tipo de situação fornece um exemplo de aplicação dos chamados números fuzzy.

3.1 Introdução

Definição 3.1. Um conjunto fuzzy A é chamado de **número fuzzy** quando o conjunto universo no qual φ_A está definida, é o conjunto dos números reais $\mathbb{R}$ e satisfaz às condições:

1. todos os α -níveis de A são não vazios, com $0 \leq \alpha \leq 1$;
2. todos os α -níveis de A são intervalos fechados de $\mathbb{R}$;
3. $\text{supp } A = \{x \in \mathbb{R} : \varphi_A(x) > 0\}$ é limitado.

Definição 3.2. Um número fuzzy A é dito **triangular** se sua função de pertinência é da forma

$$\varphi_A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{u-a} & \text{se } a < x \leq u \\ \frac{u-x}{u-b} & \text{se } u \leq x < b \\ 0 & \text{se } x \geq b \end{cases}.$$

Observação 3.1. O gráfico da função de pertinência de um número fuzzy triangular tem a forma de um triângulo, tendo como base o intervalo $[a, b]$ e, como único vértice

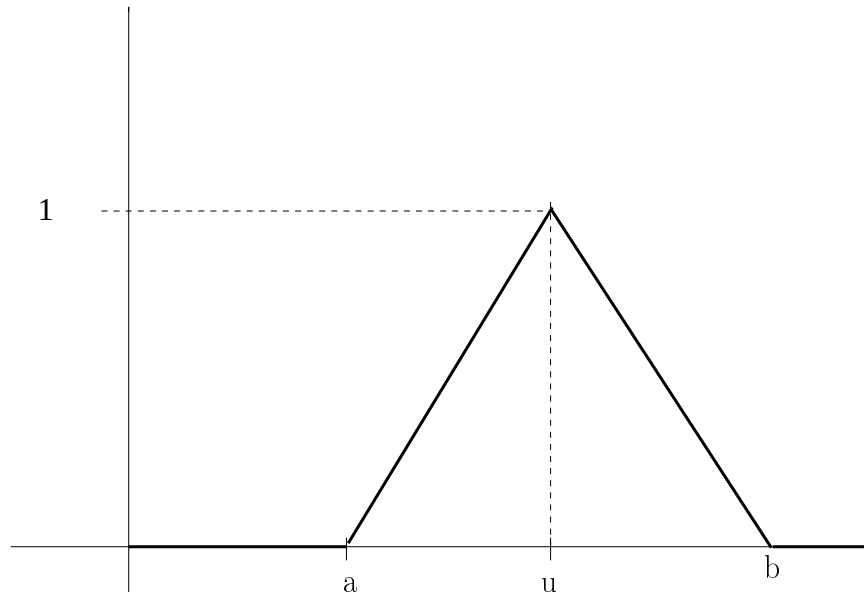


Figura 3.1: Número Fuzzy triangular

fora desta base o ponto $(u, 1)$. Representa-se um número fuzzy triangular por $(a; u; b)$ (Figura 3.1).

Definição 3.3. Um número fuzzy A é dito **trapezoidal** se sua função de pertinência é da forma

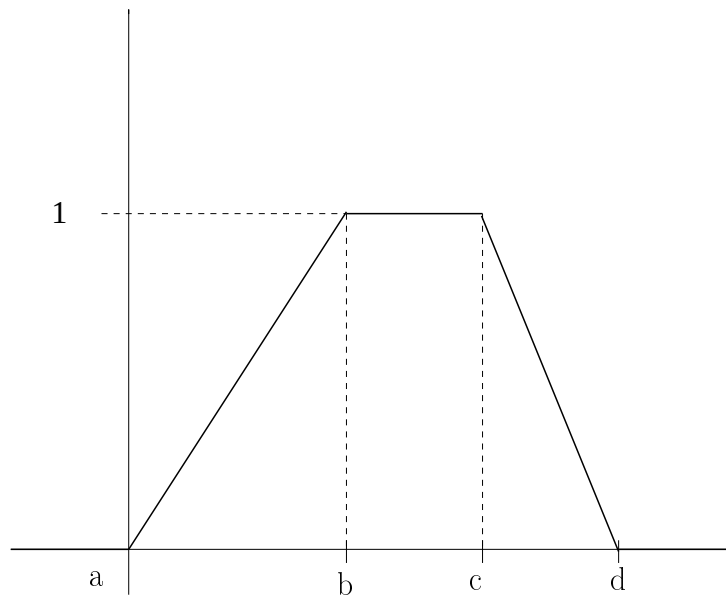


Figura 3.2: Número Fuzzy trapezoidal

$$\varphi_A(x) = \begin{cases} \frac{x-a}{b-a} & \text{se } a \leq x < b \\ 1 & \text{se } b \leq x \leq c \\ \frac{d-x}{d-c} & \text{se } c < x \leq d \\ 0 & \text{caso contrário} \end{cases}.$$

Observação 3.2. Representa-se um número fuzzy trapezoidal pela notação $(a; b; c; d)$ (Figura 3.2).

Definição 3.4. Um número fuzzy A é dito ter **forma de sino** se sua função de pertinência segue aproximadamente o perfil de uma distribuição de probabilidades. (Figura 3.3).

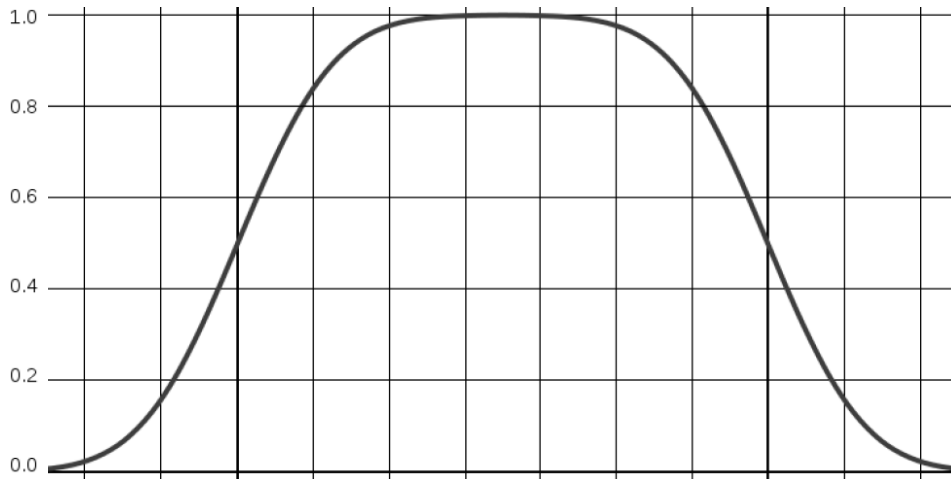


Figura 3.3: Exemplo de um Número Fuzzy em forma de sino

3.2 Operações com Números Fuzzy

Nesta seção as operações sobre números reais serão estendidas para números fuzzy. Como ocorre com as operações lógicas é de se esperar que uma extensão da aritmética dos números reais para números fuzzy não preserve todas as propriedades usuais.

Uma vez que, no caso dos números fuzzy, os α -níveis são sempre intervalos da reta, inicialmente estas operações serão estendidas para tais intervalos e, depois, para números fuzzy. Isso permite associar números fuzzy a intervalos da reta o que, por sua vez, produz métodos práticos para se obter mais facilmente os resultados de tais operações entre os números fuzzy.

3.2.1 Operações com Intervalos

Definição 3.5. Sejam $\lambda \in \mathbb{R}$ e, $A=[a_1, a_2]$ e $B=[b_1, b_2]$ dois intervalos fechados da reta. Defina-se:

1. A **soma** entre A e B é o intervalo

$$A + B = [a_1 + b_1, a_2 + b_2].$$

2. A **diferença** entre A e B é o intervalo

$$A - B = [a_1 - b_2, a_2 - b_1].$$

3. A **multiplicação** de A por um **escalar** λ é o intervalo

$$\lambda A = \begin{cases} [\lambda a_1, \lambda a_2] & \text{se } \lambda \geq 0 \\ [\lambda a_2, \lambda a_1] & \text{se } \lambda < 0 \end{cases}.$$

4. A **multiplicação** de A por B é o intervalo

$$AB = [\min P, \max P]$$

$$\text{onde } P = \{a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2\}.$$

5. A **divisão** de A por B , se $0 \notin B$, é o intervalo

$$\frac{A}{B} = [a_1, a_2] \left[\frac{1}{b_2}, \frac{1}{b_1} \right].$$

Tais operações são chamadas **intervalares**.

Exemplo 3.1. Sejam $A = [2, 5]$, $B = [6, 8]$ e $\lambda = 3$. Então:

$$\begin{aligned} A + B &= [8, 13], \\ A - B &= [-6, -1], \\ \lambda A &= [6, 15]. \end{aligned}$$

Adicionalmente, $P_1 = \{12, 16, 30, 40\}$ e, daí,

$$A.B = [12, 40].$$

Para $\frac{A}{B}$ calcula-se $P_2 = \left\{ \frac{1}{4}, \frac{1}{3}, \frac{5}{8}, \frac{5}{6} \right\}$ e, então,

$$\frac{A}{B} = \left[\frac{1}{4}, \frac{5}{6} \right].$$

Teorema 3.1. (*Princípio de Extensão para Intervalos da Reta*) Sejam A e B dois intervalos de $\mathbb{R}$, e $\otimes$ uma das operações aritméticas entre números reais. Então

$$\chi_{A \otimes B}(x) = \sup_{\{(y,z): y \otimes z = x\}} \min(\chi_A(y), \chi_B(z)).$$

Demonstração. Fazendo a demonstração para a soma tem-se:

$$\min_{\{(y,z):x=y+z\}} (\chi_A(y), \chi_B(z)) = \begin{cases} 1 & \text{se } y \in A \text{ e } z \in B \\ 0 & \text{se } y \notin A \text{ ou } z \notin B \end{cases}.$$

Daí, se $y \in A = [a_1, a_2]$, então $y = a_1 + t$, e $0 \leq t \leq a_2 - a_1$. Se $z \in B = [b_1, b_2]$, segue que $z = b_1 + u$, e $0 \leq u \leq b_2 - b_1$.

Logo, $x = y + z = a_1 + b_1 + (t + u)$ e $0 \leq a_2 - b_2 - (a_1 - b_1)$. Logo, $x \in A + B$.

Por outro lado, se $x \in A + B$, quer-se concluir que existem $y \in A$ e $z \in B$ tal que $x = y + z$. Mas $a_1 + b_1 \leq x \leq a_2 + b_2$. Tomando $y \in A = [a_1, a_2]$ e subtraindo esta expressão da anterior tem-se:

$$a_1 + b_1 - a_1 \leq x - y \leq a_2 + b_2 - a_2$$

e $b_1 \leq x - y \leq b_2$. Basta, então tomar $z = x - y$ e obtem-se o resultado desejado.

As demais operações são demonstradas de maneira análoga. □

Corolário 3.1. *Os α -níveis do conjunto clássico $A + B$, com função característica χ_{A+B} , são dados por*

$$[A + B]^\alpha = A + B$$

para todo $\alpha \in [0, 1]$, onde A e B são conjuntos clássicos quaisquer.

Demonstração. Seja $x \in A + B$. Então $\chi_{A+B}(x) = 1$. Mas, então, $1 = \chi_{A+B}(x) \geq \alpha, \forall \alpha \in [0, 1]$. Logo $x \in [A + B]^\alpha$. □

3.2.2 Operações com Números Fuzzy

Definição 3.6. *Sejam A e B dois números fuzzy e λ um número real.*

1. *A soma dos números fuzzy A e B é um número fuzzy, denotado por $A + B$, cuja função de pertinência é dada por*

$$\varphi_{(A+B)}(z) = \sup_{\{(x,y) \in A \times B: x+y=z\}} \min(\varphi_A(x), \varphi_B(y)).$$

2. *A multiplicação de λ por A é um número fuzzy, denotado por λA , cuja função de pertinência é dada por:*

$$\varphi_{(\lambda A)}(z) = \sup |\varphi_A(x)| = \begin{cases} \varphi_A(\lambda^{-1}z) & \text{se } \lambda \neq 0 \\ 0 & \text{se } \lambda = 0 \end{cases}.$$

3. A diferença entre A e B é um número fuzzy, denotado por $A - B$ cuja função de pertinência é dada por

$$\varphi_{(A-B)}(z) = \sup_{\{(x,y) \in A \times B : x-y=z\}} \min(\varphi_A(x), \varphi_B(y)).$$

4. A multiplicação de A por B é um número fuzzy, denotado por $A.B$ cuja função de pertinência é dada por:

$$\varphi_{(A.B)}(z) = \sup_{\{(x,y) \in A \times B : x.y=z\}} \min(\varphi_A(x), \varphi_B(y)).$$

5. A divisão de A por B , se $0 \notin \text{supp } B$ é um número fuzzy, denotado por A/B , cuja função de pertinência é dada por:

$$\varphi_{(A/B)}(z) = \sup_{\{(x,y) \in A \times B : x/y=z\}} \min(\varphi_A(x), \varphi_B(y)).$$

Como se pode perceber facilmente desta definição a manipulação de tais operações não é simples nem intuitiva. Por causa disso é vantajoso procurar uma conexão entre tais operações e os α -níveis dos números fuzzy, que são intervalos da reta. As chamadas operações intervalares fornecem essa conexão como mostra a proposição seguinte.

Proposição 3.1. Sejam A e B números fuzzy, com α -níveis $[A]^\alpha = [a_1^\alpha, a_2^\alpha]$ e $[B]^\alpha = [b_1^\alpha, b_2^\alpha]$. São válidas as seguintes propriedades:¹

1. A soma entre A e B é o número fuzzy $A + B$ cujos α -níveis são:

$$[A + B]^\alpha = [A]^\alpha + [B]^\alpha = [a_1^\alpha + b_1^\alpha, a_2^\alpha + b_2^\alpha].$$

2. A diferença entre A e B é o número fuzzy $A - B$ cujos α -níveis são:

$$[A - B]^\alpha = [A]^\alpha - [B]^\alpha = [a_1^\alpha - b_2^\alpha, a_2^\alpha - b_1^\alpha].$$

3. A multiplicação de λ por A é o número fuzzy λA cujos α -níveis são:

$$[\lambda A]^\alpha = \lambda[A]^\alpha = \begin{cases} [\lambda a_1^\alpha, \lambda a_2^\alpha] & \text{se } \lambda \geq 0. \\ [\lambda a_2^\alpha, \lambda a_1^\alpha] & \text{se } \lambda < 0. \end{cases}$$

4. Seja $X = \{a_1^\alpha b_1^\alpha, a_1^\alpha b_2^\alpha, a_2^\alpha b_1^\alpha, a_2^\alpha b_2^\alpha\}$. A multiplicação de A por B é o número fuzzy $A.B$ cujos α -níveis são

$$[A.B]^\alpha = [A]^\alpha [B]^\alpha = [\min X, \max X].$$

¹Para detalhes veja *Tópicos de Lógica Fuzzy e Biomatemática* [11].

5. A divisão de A por B , se $0 \notin \text{supp } B$, é o número fuzzy cujos α -níveis são

$$\left[\frac{A}{B} \right]^\alpha = \frac{[A]^\alpha}{[B]^\alpha} = [a_1^\alpha, a_2^\alpha] \left[\frac{1}{b_2^\alpha}, \frac{1}{b_1^\alpha} \right].$$

Exemplo 3.2. Sejam $X = (3; 4; 5)$ e $Y = (7; 8; 9)$ dois números fuzzy triangulares, expressando “quase 4” e “quase 8”. Calculando seus α -níveis chega-se a:

$$[A]^\alpha = [3 + \alpha, 5 - \alpha]$$

e

$$[B]^\alpha = [7 + \alpha, 9 - \alpha].$$

Daí, pelo item anterior, vale:

1. $[A + B]^\alpha = [A]^\alpha + [B]^\alpha = [10 + 2\alpha, 14 - 2\alpha]$. Logo, $A + B = (10; 12; 14)$.
2. $[A - B]^\alpha = [A]^\alpha - [B]^\alpha = [-6 + 2\alpha, -2 - 2\alpha]$. Portanto, $A - B = (-6; -4; -2)$.
3. Seja $\lambda = 6$. $[A]^\alpha = 6[A]^\alpha = [18 + 6\alpha, 30 - 6\alpha]$. Daí, $6A = (18; 24; 30)$.
4. Para a multiplicação $A.B$ obtém-se, primeiro, o conjunto X , que no exemplo é dado por: $X = \{(3 + \alpha)(7 + \alpha), (3 + \alpha)(9 - \alpha), (5 - \alpha)(7 + \alpha), (5 - \alpha)(9 - \alpha)\}$. Então $[A.B]^\alpha = [A]^\alpha[B]^\alpha = [(3 + \alpha)(7 + \alpha), (5 - \alpha)(9 - \alpha)]$.
5. A divisão de A por B , se $0 \notin \text{supp } B$, é o número fuzzy cujos α -níveis são

$$\left[\frac{A}{B} \right]^\alpha = \frac{[A]^\alpha}{[B]^\alpha} = [3 + \alpha, 5 - \alpha] \left[\frac{1}{7 + \alpha}, \frac{1}{9 - \alpha} \right].$$

Calculando o conjunto X para este produto chega-se a

$$X = \left\{ \frac{3 + \alpha}{9 - \alpha}, \frac{3 + \alpha}{5 - \alpha}, \frac{5 - \alpha}{7 + \alpha}, \frac{5 - \alpha}{9 - \alpha} \right\}.$$

Portanto,

$$\left[\frac{A}{B} \right]^\alpha = \left\{ \frac{3 + \alpha}{9 - \alpha}, \frac{5 - \alpha}{7 + \alpha} \right\}.$$

Parte II

Aplicação: Representação Geométrica dos Conjuntos Fuzzy

4 “Fuzziness” e Probabilidade

4.1 Conjuntos Fuzzy e Aleatoriedade

Esse capítulo analisa “fuzziness” como alternativa a aleatoriedade na descrição da incerteza. Para isso utiliza-se a visão geométrica dos conjuntos fuzzy finitos, que identifica um conjunto fuzzy com um ponto em um hipercubo unitário. Como mencionado no Capítulo 1, neste hipercubo os conjuntos clássicos são identificados com os vértices. Os demais pontos do hipercubo (e existem tantos destes quanto há números reais) representam os conjuntos fuzzy não-clássicos.

Na teoria fuzzy um elemento sempre pertence a um conjunto fuzzy com algum grau entre 0 e 1. Esta visão pode ser ampliada para conjuntos por considerar um conjunto sempre incluído em outro com algum grau. Deste ponto de vista tanto a pertinência de elementos, como o conceito de probabilidade podem ser vistos como inclusão de conjuntos.

4.1.1 “Fuzziness” e Probabilidade

Muitos acreditam que a incerteza e a aleatoriedade são a mesma coisa e que indício e probabilidade permitem considerar todo tipo de incerteza que existe. Por exemplo, o físico E.T. Jaynes [12] afirma que “qualquer método de inferência que apresente graus de razoabilidade por números reais, é necessariamente equivalente a [probabilidade] de Laplace, ou inconsistente”. Ele afirma que o físico R. T. Cox[13] provou isto como um teorema.

Porém, através de exemplos e teoremas tem sido demonstrado que “fuzziness” e aleatoriedade são diferentes tanto teórica como conceitualmente, embora apresentem algumas similaridades. Como um exemplo destas similaridades, tem-se que ambos os processos descrevem as incertezas de forma numérica. Apesar disto, a forma como estas teorias tratam o conjunto A e seu complementar A^c , mostra que, com certeza, existem muitas diferenças entre elas.

A teoria das probabilidades concorda com a teoria clássica dos conjuntos quando afirma que $A \cap A^c = \emptyset$. Deste ponto de vista, $A \cap A^c$ representa um evento probabilisticamente impossível. Por sua vez, “fuzziness” começa justamente quando $A \cap A^c \neq \emptyset$.

Levanta-se as seguintes questões:

1. Será que, de fato, existe algum caso em que $A \cap A^c \neq \emptyset$? Ou seja, “fuzziness” existe?
2. Pode-se concluir ou supor que o operador probabilístico condicional

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad (4.1)$$

aplica-se a todos os sistemas envolvendo incerteza?

Durante o decorrer deste capítulo ficará claro que a primeira questão leva à segunda. A princípio, deve-se entender o que é medir a “fuzziness” de um conjunto fuzzy A . Evidentemente, na teoria das probabilidades a relação $A \cup A^c \subset A \cap A^c$ tem probabilidade zero, pois

$$\begin{aligned} P(A \cap A^c | A \cup A^c) &= \frac{P((A \cup A^c) \cap (A \cap A^c))}{P(A \cup A^c)} = \frac{P((A \cup A^c) \cap \emptyset)}{P(A \cup A^c)} = \\ &= \frac{P(X \cap \emptyset)}{P(X)} = \frac{P(\emptyset)}{P(X)} = \frac{P(\emptyset \cap X)}{P(X)} = P(\emptyset | X) = 0, \end{aligned}$$

onde X denota o espaço todo, ou seja, o evento certo.

Então, medir a “fuzziness” de um conjunto fuzzy A significa medir o quanto ele se afasta desse padrão, isto é, o quanto $A \cup A^c$ está contido em $A \cap A^c$.

A primeira pergunta coloca em questão as leis da lógica clássica, a saber, a não-contradição e o terceiro excluído, as leis $A \cap A^c = \emptyset$ e $A \cup A^c = X$. “Fuzziness” de conjuntos ocorre se, e somente se, estas leis são violadas. A lógica e a teoria dos conjuntos clássicas assumem que isto nunca ocorre. Desta forma, “fuzziness” começa onde a lógica clássica termina e onde, do ponto de vista desta última, a contradição começa.

4.1.2 A Diferença entre SE e QUANTO

O objetivo de “fuzziness” é medir o QUANTO um evento ocorre e não SE ele ocorre. Por outro lado, a teoria da probabilidade descreve a aleatoriedade de um evento. SE ele ocorre ou não. De que maneira ele é incerto? É possível, sem ambiguidade, distinguir um evento de outros?

SE um evento ocorre ou não é aleatoriedade, mas O QUANTO ele ocorre é “fuzzy”. Se a ocorrência de um evento ambíguo (como quando se diz que há 20 por cento de chance do sol nascer amanhã) envolver componentes INCERTOS, envolverá a probabilidade de um evento “fuzzy”.

Geralmente confunde-se probabilidade com eventos “fuzzy”. Por exemplo, estudantes que tiram “A” podem incluir estudantes que tiraram “B+”. Ou seja, na prática

um evento pode incluir em maior ou em menor grau eventos que são diferentes dele. Mas na teoria, nos livros textos, um evento e, em especial, seu evento complementar, são considerados como pretos ou brancos. Quine [14] menciona: “Uma colina é considerada uma montanha se ela tem x metros, mas não é considerada assim se tiver um microm a menos do que x metros.”

Como um outro exemplo, seja o evento “estacionar o carro nas faixas pintadas de um estacionamento”. Estaciona-se o carro em uma das faixas com alguma probabilidade. Após isto, ocupa-se uma das faixas com toda a probabilidade e todas as outras com nenhuma probabilidade. O número probabilístico reflete com que frequência uma vaga é totalmente ocupada pelo carro. No entanto, é possível imaginar que se ocupa todas as faixas com um certo grau. O carro parcial, e deterministicamente, ocupará todo espaço. Na prática, ele ocupará muitos espaços com grau zero. Finalmente, o intervalo $[0, 1]$ é usado para descrever, para cada faixa de estacionamento, a probabilidade de ocorrência de cada grau de ocupação parcial - ou seja, probabilidades de eventos “fuzzy”.

A todo instante imagina-se que eventos não são ambíguos, assim como a retirada de bolas de uma urna não o é. Porém, uma análise do universo físico mostra que a maior parte das afirmações de ambiguidade ou não de um evento não passam de hipóteses empíricas. Quando a teoria das probabilidades é usada no estudo de tais eventos este fato é desconsiderado. Simplesmente assume-se que eventos não são ambíguos. No entanto, quando se pensa um pouco mais nesse assunto se percebe que $A \cap A^c = \emptyset$ é o “postulado das paralelas” da teoria dos conjuntos e da lógica clássica. De que forma?

Se ele for colocado de lado e postular-se que a negação dele vale, ou seja, $A \cap A^c \neq \emptyset$, “fuzziness” será um tipo válido de incerteza. E, se “fuzziness” existe, as consequências físicas são universais. A observação mostra que “fuzziness” é um tipo de incerteza determinística e a ambiguidade é uma propriedade de fenômenos físicos. A probabilidade pode ser posta de lado conforme a informação aumenta, pois uma vez que um fato ocorreu é desnecessário analisar a sua probabilidade. Isso não acontece com “fuzziness”. “Fuzziness” permanece sempre.

Então, a questão é: Será que $\varphi_A(x)$, o grau de pertinência de um elemento x a um conjunto fuzzy A é igual a probabilidade com que x pertence a A ? Ou seja, a afirmação

$$\varphi_A(x) = Prob\{x \in A\} \quad (4.2)$$

é sempre verdadeira?

Não é sem razão que de tempos em tempos os probabilistas se perguntem se a probabilidade descreve alguma coisa real. De Demócrito a Einstein tem se suspeitado que, como David Hume [15] mencionou: “Pense que não há coisa como *chance* no mundo, nossa ignorância é a causa real de qualquer evento ter alguma influência sobre o entendimento e requerer alguma espécie de crença.” A fim de reduzir essa “ignorância” será analisado o conceito de entropia do ponto de vista dos conjuntos fuzzy numa tentativa de medir “fuzziness”.

4.2 Retomando a Visão de Conjuntos como Pontos

A representação de conjuntos fuzzy como pontos, mencionada no primeiro capítulo, é útil na descrição de “fuzziness”, na definição de conceitos fuzzy e na demonstração de teoremas da teoria dos conjuntos fuzzy. Visualizar esta geometria fornece ainda um poderoso argumento para afirmar que “fuzziness” existe, ou seja, que nem sempre a afirmação em (4.1) é verdadeira.

Todos os subconjuntos fuzzy A em um determinado conjunto X podem ser representados como pontos num hipercubo. Então, (X, I^n) define o espaço fundamental da teoria fuzzy (finita). Para um conjunto com n elementos existem 2^n vetores de *bits* (zeros ou uns) de comprimento n que definem os vértices do hipercubo I^n . Assim, os vértices também representam o conjunto potência clássico dos n elementos $x_1, \dots, x_n$, o conjunto de todos os subconjuntos clássicos de n elementos de X . Nos elementos desse conjunto um *bit* de valor 0 na i -ésima posição do vetor de *bits* indica a ausência do elemento x_i no subconjunto. Um *bit* de valor 1 indica a presença de x_i no subconjunto. O vetor de bits (1 0 1 0 0) indica o subconjunto $\{x_1, x_3\}$ do conjunto $\{x_1, x_2, x_3, x_4, x_5\}$.

Note que, uma vez que os valores *fit* varrem o intervalo $[0, 1]$ um *bit* é um caso particular de um *fit*. Considerando um ponto no hipercubo associado a um vetor *fit*, o valor *fit* na i -ésima coluna será igual ao valor da função de pertinência, ou seja, ao grau com que x_i pertence a A . Por exemplo, o valor *fit* $\frac{1}{5}$ indica que x_i pertence apenas um pouco a A , o valor $\frac{1}{2}$ indica que ele pertence de forma igual tanto a A como a A^c .

Os exemplos na próxima seção mostrarão como operar com os conjuntos fuzzy nesta visão geométrica.

4.2.1 Operando com a Representação Geométrica

Escolhendo-se o mínimo, o máximo e a negação para operar com pontos do hipercubo, os operadores da lógica de Lukasiewicz, serão dados por

$$\begin{aligned}\varphi_{A \cap B} &= \min(\varphi_A, \varphi_B), \\ \varphi_{A \cup B} &= \max(\varphi_A, \varphi_B), \\ \varphi_{A^c} &= 1 - \varphi_A.\end{aligned}$$

Observação 4.1. Às vezes, para simplificar a notação, usa-se $A(x)$ para $\varphi_A(x)$. Neste caso, as definições anteriores ficam:

$$\begin{aligned}(A \cap B)(x) &= A(x) \wedge B(x), \\ (A \cup B)(x) &= A(x) \vee B(x), \\ A^c(x) &= 1 - A(x).\end{aligned}$$

Exemplo 4.1. Sejam $X = \{x_1, x_2\}$ e A , um subconjunto fuzzy em X . Se $x_1 \in A$ com grau $\frac{1}{3}$ e $x_2 \in A$ com grau $\frac{3}{4}$, A pode ser representado como $A = \left(\frac{1}{3} \frac{3}{4}\right)$ ou, ainda,

$A(x_1) = \frac{1}{3}$ e $A(x_2) = \frac{3}{4}$. Então, A é um ponto no quadrado formado por todos os subconjuntos fuzzy em $\tilde{X}$.

Exemplo 4.2. Sejam $A = (1, 0 \ 0, 9 \ 0, 5 \ 0, 3)$ e $B = (0, 7 \ 1, 0 \ 0, 6 \ 0, 4)$, conjuntos fuzzy. Então

$$\begin{aligned} A \cap B &= (0, 7 \ 0, 9 \ 0, 5 \ 0, 3), \\ A \cup B &= (1, 0 \ 1, 0 \ 0, 6 \ 0, 4), \\ A^c &= (0, 0 \ 0, 1 \ 0, 5 \ 0, 7), \\ A \cap A^c &= (0, 0 \ 0, 1 \ 0, 5 \ 0, 3), \\ A \cup A^c &= (1, 0 \ 0, 9 \ 0, 5 \ 0, 7), \\ B^c &= (0, 3 \ 0, 0 \ 0, 4 \ 0, 6), \\ B \cap B^c &= (0, 3 \ 0, 0 \ 0, 4 \ 0, 4), \\ B \cup B^c &= (0, 7 \ 1, 0 \ 0, 6 \ 0, 6). \end{aligned}$$

Esse exemplo ilustra que $A \cap A^c$ pode ser diferente de $\emptyset$ (representado por $(0 \ 0 \ 0 \ 0)$) e que $A \cup A^c$ pode ser diferente do universo (representado por $(1 \ 1 \ 1 \ 1)$).

Observação 4.2. As leis da lógica clássica valem apenas nos vértices do conjunto I^n , ou seja, para os conjuntos clássicos (veja a Definição 1.4 de conjuntos não-clássicos). Em todos os outros pontos as violações das leis da contradição e do terceiro excluído ocorrem em maior ou menor grau. Isso mostra que, na teoria dos conjuntos fuzzy, as violações são a regra não a exceção. Ou seja, os conjuntos clássicos, que obedecem as leis da lógica clássica, são a “minoria” localizando-se nos vértices dos hipercubos fuzzy.

4.3 O Ponto Médio do Quadrado Fuzzy

Seja o conjunto fuzzy A definido pelo vetor-linha $\left(\frac{1}{3} \ \frac{4}{5}\right)$. Encontrando seu complementar A^c , sua sobreposição $A \cup A^c$ e sua subposição $A \cap A^c$ chega-se a:

$$\begin{aligned} A &= \left(\frac{1}{3} \ \frac{4}{5}\right), \\ A^c &= \left(\frac{2}{3} \ \frac{1}{5}\right), \\ A \cap A^c &= \left(\frac{1}{3} \ \frac{1}{5}\right), \\ A \cup A^c &= \left(\frac{2}{3} \ \frac{4}{5}\right). \end{aligned}$$

A visão geométrica dos conjuntos fuzzy mostra que estes quatro pontos juntos formam um retângulo e que, quando um se move, os outros se moverão de forma natural. Pode-se intuir este movimento por se observar com atenção a Figura 4.1.

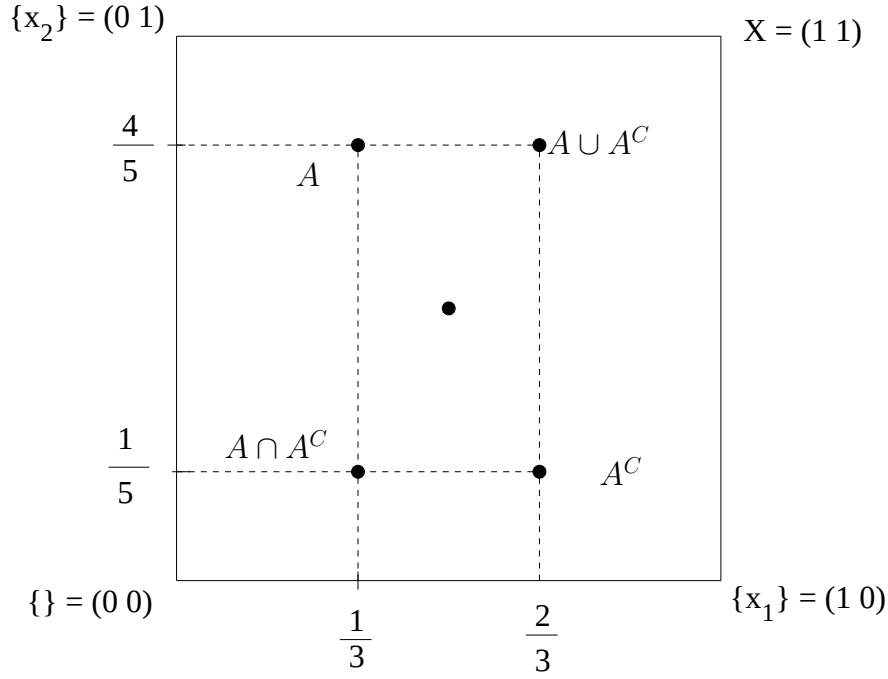


Figura 4.1: Simetria em relação ao Ponto Médio

Observação 4.3. Note que nem sempre existem quatro pontos distintos. Nos casos onde $A = \{x_1, x_2\}$ é tal que $\varphi_A(x_1) < \frac{1}{2}$ e $\varphi_A(x_2) < \frac{1}{2}$, vale que $A = A \cap A^c$ e $A^c = A \cup A^c$. De fato, considerando o caso geral para um n qualquer tem-se que, se

$$\varphi_A(x_i) < \frac{1}{2}, \quad \text{para todo } x \in X,$$

então

$$\varphi_{A^c}(x_i) = 1 - \varphi_A(x_i) \geq \frac{1}{2}, \quad \text{para todo } x_i \in X.$$

Desta maneira,

$$\varphi_{A \cap A^c}(x_i) = \min(\varphi_A(x_i), \varphi_{A^c}(x_i)) = \min(\varphi_A(x_i), 1 - \varphi_A(x_i)) = \varphi_A(x_i), \text{ e}$$

$$\varphi_{A \cup A^c}(x_i) = \max(\varphi_A(x_i), \varphi_{A^c}(x_i)) = \max(\varphi_A(x_i), 1 - \varphi_A(x_i)) = \varphi_{A^c}(x_i).$$

Portanto, para $n = 2$, serão apenas dois pontos, sendo que as coordenadas do ponto médio do quadrado determinam o limite entre estas duas situações.

Exemplo 4.3. Como um exemplo, seja $A = \left(\frac{1}{3} \frac{1}{5}\right)$. Então

$$\begin{aligned} A^c &= \left(\frac{2}{3} \frac{4}{5}\right), \\ A \cap A^c &= \left(\frac{1}{3} \frac{1}{5}\right) = A, \\ A \cup A^c &= \left(\frac{2}{3} \frac{4}{5}\right) = A^c, \end{aligned}$$

conforme ilustrado na Figura 4.2.

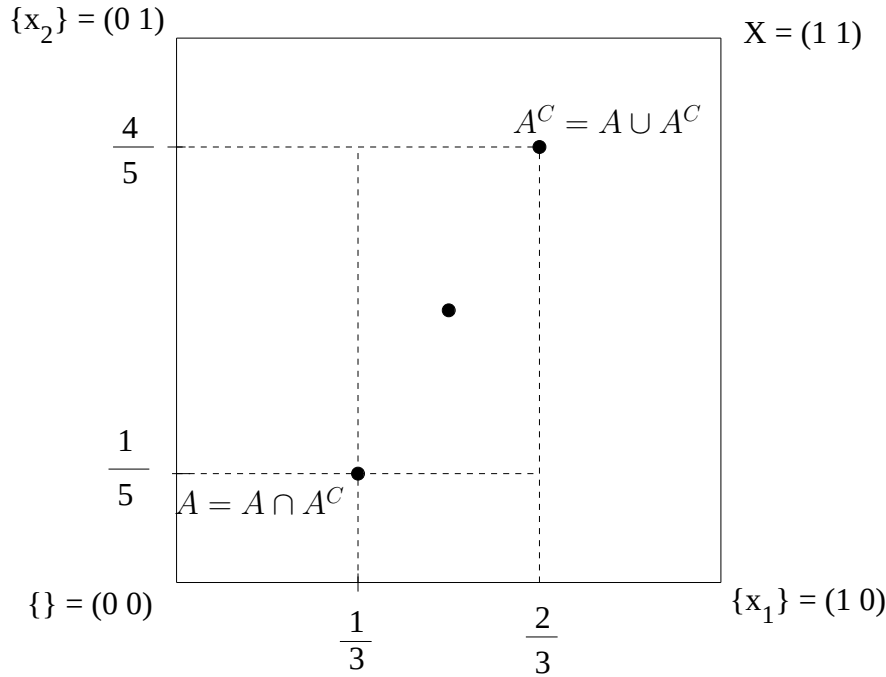


Figura 4.2: Simetria em relação ao Ponto Médio com $\varphi(x_i) < \frac{1}{2}, \forall i$

Observando os exemplos, quanto mais os pontos A , A^c , $A \cap A^c$ e $A \cup A^c$ se aproximarem do ponto médio, mais a “fuzziness” de A se aproxima do valor máximo. Por outro lado, quanto mais eles se aproximarem dos vértices booleanos do quadrado, mais ela se aproxima de zero. Essa contração ou expansão ocorre independentemente do número de dimensões consideradas. Ou seja, fixada uma dimensão n os 2^n conjuntos fuzzy definidos por todas as combinações de $\varphi_A(x_1)$ e $\varphi_{A^c}(x_1)$, $\dots$, $\varphi_A(x_n)$ e $\varphi_{A^c}(x_n)$ sofrem esta contração ou expansão conforme se aproximam ou se afastam do ponto médio. No ponto médio não há distinção, a “fuzziness” é máxima. Nos vértices todas as coisas são distintas, ou seja, a “fuzziness” é igual a zero.

Resumidamente, no ponto médio nada se distingue e nos vértices tudo se distingue. São os extremos da lógica e da teoria dos conjuntos. Sendo como é, o ponto médio pode ser considerado, metaforicamente falando, o buraco negro da teoria dos conjuntos, uma vez que nele é impossível distinguir qualquer coisa.

O ponto médio, portanto, é o local onde os paradoxos da lógica bivalente ocorrem. Por meio dessa visão geométrica é possível responder questões tais como “Quão fuzzy é um conjunto fuzzy?” e “Quanto um conjunto fuzzy pode ser subconjunto de outro?” Para alcançar este objetivo utiliza-se os teoremas da Entropia Fuzzy e da Inclusão.

O ponto médio do hipercubo I^n possui “fuzziness” máxima, uma vez que nele todos os valores são iguais a $\frac{1}{2}$. Por causa disso, ele é único em dois aspectos:

1. É o único conjunto igual ao seu complementar, igual à sua sobreposição e igual à sua subposição, isto é, denotando o ponto médio por M , vale que

$$M = M \cap M^c = M \cup M^c = M^c,$$

ou seja, é no ponto médio que ocorre a violação máxima das leis da contradição e do terceiro excluído; e

2. é o único ponto no hipercubo I^n que é equidistante de cada um dos 2^n vértices.

É devido a estas características que todos os paradoxos da lógica clássica ocorrem no ponto médio. A lógica clássica e a teoria tradicional dos conjuntos, evitam o ponto médio através dos axiomas da contradição e do terceiro excluído que, por sua vez, geram os chamados “paradoxos” dos sistemas bivalentes. Por meio de tais axiomas, consagrados no pensamento ocidental, quando os fenômenos que ocorrem no ponto médio aparecem, os teóricos tradicionalmente os chamam de “paradoxos” ou, então, os desconsideram totalmente.

Para ilustrar considere o paradoxo do barbeiro, que surge na afirmação de um barbeiro que anunciou que faria a barba apenas daqueles que não se barbeassem. Então, quem faz a barba do barbeiro? Se ele barbeia a si mesmo, então por definição ele não poderia fazer isto. Por outro lado, se ele não faz isto, então ele pode se barbear. Então ele o faz e ele não o faz, ou seja, ocorre uma contradição. Seja S a proposição que afirma que o barbeiro barbeia a si mesmo e $\sim S$ a que afirma que ele não o faz. Então desde que S implica $\sim S$ e $\sim S$ implica S são verdadeiros, as duas proposições são logicamente equivalentes e tem o mesmo valor verdade, ou seja,

$$t(S) = t(\sim S) = 1 - t(s).$$

Resolvendo $t(S)$ chega-se ao ponto médio do intervalo $[0, 1]$ (que é o “cubo” unidimensional): $t(S) = \frac{1}{2}$. O ponto médio é equidistante dos vértices 0 e 1. No caso bivalente (dois-valores), um paradoxo ocorre. Na realidade, estas equações descrevem a forma lógica de muitos paradoxos, embora cada um deles tenha uma descrição diferente.

Na lógica bivalente as sentenças S e $\sim S$ devem ter valor 0 ou 1. A resolução fuzzy do paradoxo usa somente o fato de que os valores verdade são iguais. Isso não restringe seu contra-domínio. O valor do ponto médio $\frac{1}{2}$ emerge da estrutura do problema e inverte o efeito da negação.

Kosko [16] afirma que, os paradoxos da lógica clássica e da teoria dos conjuntos ilustram o preço pago quando se insiste, de forma arbitrária, no uso de uma lógica bivalente. Os cientistas frequentemente insistem na bivalência em nome da ciência. Mas, no fim, esta insistência se reduz a uma mera preferência cultural, reflexo de uma predileção educacional que remonta, pelo menos, até Aristóteles. “Fuzziness” mostra que há limites para a certeza lógica. Não se pode exigir que as leis da não contradição e do terceiro excluído nunca sejam violadas usando a afirmação de que, desta forma, sempre haverá uma certeza. Permitindo essa violação uma liberdade maior é alcançada.

Por outro lado, para utilizar o conceito de “fuzziness” é preciso explicar como ele funciona em sistemas bivalentes e vice-versa. É preciso explicar porque durante tanto tempo “fuzziness” não foi considerada pela grande maioria. Isso ocorreu porque o costume é aproximar, quantificar, simplificar a vida e baratear custos. Se todo mundo concorda, não há nenhum problema em chamar algo que esteja “quase vazio” de vazio.

Na realidade, de forma inconsciente os pontos do hipercubo fuzzy são sempre associados ao vértice mais próximo. Mas isto é exatamente o que não se pode fazer com o ponto médio, uma vez que ele é equidistante dos vértices. Quando se usa este artifício da lógica clássica no ponto médio, insistindo na bivalência declara-se que paradoxos existem.

Esse dilema leva à próxima questão, a saber, quão bem a aleatoriedade descreve a incerteza? Para responder a esta pergunta são necessários instrumentos que quantifiquem a “fuzziness” de um conjunto. Iniciando a apresentação de tais instrumentos a próxima seção traz a definição de contador sigma.

4.4 O Contador Sigma

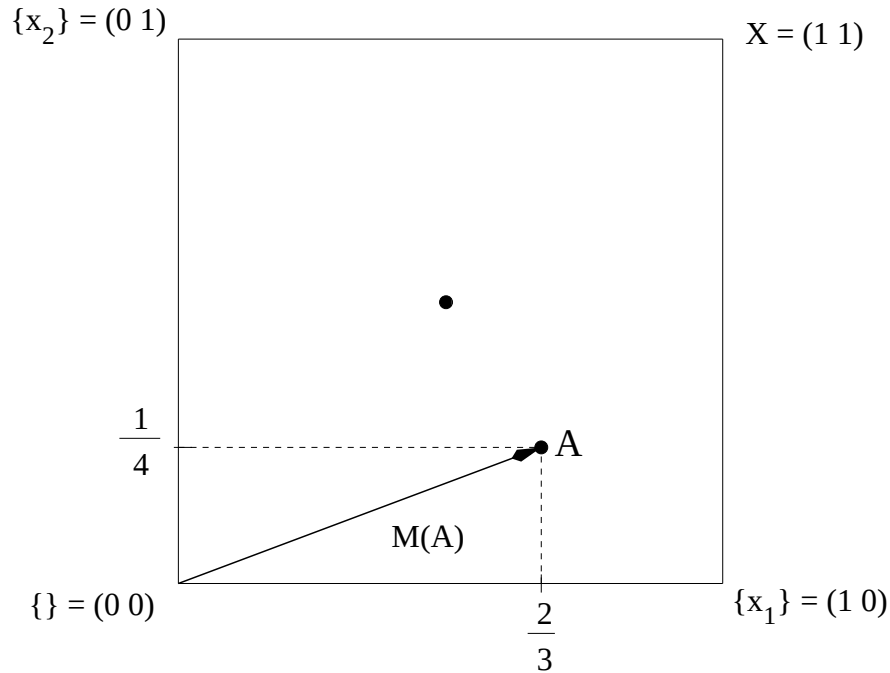


Figura 4.3: Medida M do Conjunto Fuzzy A

Como generalizar o conceito de cardinalidade de um conjunto fuzzy? Zadeh [17] introduziu uma medida para o tamanho de A , denotada por $M(A)$ e chamada de **contador sigma**. Essa medida é dada por

$$\ell_1(A, \emptyset) = M(A) = \sum_{i=1}^n \varphi_A(x_i), \quad x_i \in X,$$

ou seja, a soma dos graus de pertinência dos elementos x_i a A .

Exemplo 4.4. O contador de $A = \left(\frac{2}{3} \frac{1}{4}\right)$ é igual a $M(A) = \frac{2}{3} + \frac{1}{4} = \frac{11}{12}$.

A medida M tem uma interpretação geométrica natural no conjunto I^n . $M(A)$ é igual à distância da origem até o conjunto fuzzy, como ilustrado na Figura 4.3. Qual a forma mais conveniente de medir esta distância?

A definição geral desta distância entre os conjuntos fuzzy A e B em I^n , que é denotada por ℓ^p , é dada por:

$$\ell^p(A, B) = \sqrt[p]{\sum_{i=1}^n |\varphi_A(x_i) - \varphi_B(x_i)|^p}, \quad (4.3)$$

onde $1 \leq p < \infty$. A distância ℓ^2 é a distância euclidiana usual. A distância mais adequada dentre estas é ℓ^1 , a soma das distâncias absolutas. Esta é uma generalização da distância de Hamming, aplicada à teoria fuzzy, e que é chamada, por analogia, de **distância Hamming fuzzy**.

Terminadas as considerações sobre a representação geométrica, o próximo capítulo apresentará alguns teoremas envolvendo os conjuntos fuzzy sob este ponto de vista.

5 Teoremas sobre Conjuntos Fuzzy

5.1 Teorema da Entropia Fuzzy

Quão fuzzy é um conjunto? Para responder a esta pergunta esta seção descreverá uma medida da “fuzziness” de um conjunto chamada **entropia fuzzy**. Como analisado anteriormente, a entropia é uma noção genérica que pode ser utilizada para medir a incerteza de uma mensagem. No caso da entropia fuzzy, busca-se uma medida de incerteza que não seja, necessariamente, probabilística. Ao usar a representação de um conjunto fuzzy como um ponto, pode-se considerar essa representação como uma mensagem onde cada grau de pertinência é um *fit*, ou seja, uma unidade de informação fuzzy. Neste contexto, ao medir a incerteza desta mensagem mede-se a “fuzziness” do conjunto fuzzy associado.

5.1.1 Entropia e Conjuntos Fuzzy

No hipercubo unitário I^n , a entropia de A , $E(A)$, varia de 0 a 1. Nos vértices do hipercubo a entropia será zero, já que conjuntos clássicos não são ambíguos. Em outras palavras, ao medir sua entropia, não se obtém informação adicional sobre eles. Tudo o que é possível saber, já se sabe. Por outro lado, no ponto médio a entropia é máxima, ou seja, 1. É como se fosse necessário saber tudo a respeito dele. Conforme um ponto do vértice movimenta-se em direção ao ponto médio, a entropia fuzzy vai aumentando lentamente.

Conhecendo este fato é possível encontrar, de forma intuitiva, uma definição natural para a entropia fuzzy.

Definição 5.1. *Sejam*

- A_p : um dos vértices mais próximos do conjunto A ;
- A_l : complementar de A_p ;
- a : distância $\ell^1(A, A_p)$ de A a A_p ; e
- b : distância $\ell^1(A, A_l)$ de A a A_l .

A entropia fuzzy é definida, então, como a razão de a para b , ou seja,

$$E(A) = \frac{a}{b} = \frac{\ell^1(A, A_p)}{\ell^1(A, A_l)}.$$

Observação 5.1. Deve-se lembrar que o ponto médio é equidistante de todos os vértices. Assim, no ponto médio M , tem-se que $M_p = M_l$. Ou seja, neste caso, qualquer um dos vértices do hipercubo pode ser escolhido como o mais próximo.

Adicionalmente, quando existir um ou mais x_i tais que $\varphi_A(x_i) = \frac{1}{2}$, existirá mais de um vértice que pode ser escolhido como A_p .

Note que, com esta definição, a ideia intuitiva de entropia fuzzy é formalizada, uma vez que, com ela, no ponto médio a entropia vale 1 e nos vértices ela vale 0.

Exemplo 5.1. A Figura 5.1 mostra a interpretação geométrica da medida de entropia fuzzy. Dado $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$, tem-se $A_p = (1 \ 0)$ e $A_l = (0 \ 1)$. Então

$$a = |\varphi_A(x_1) - \varphi_{A_p}(x_1)| + |\varphi_A(x_2) - \varphi_{A_p}(x_2)| = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

e

$$b = |\varphi_A(x_1) - \varphi_{A_l}(x_1)| + |\varphi_A(x_2) - \varphi_{A_l}(x_2)| = \frac{2}{3} + \frac{3}{4} = \frac{17}{12}.$$

E assim, $E(A) = \frac{7}{17}$.

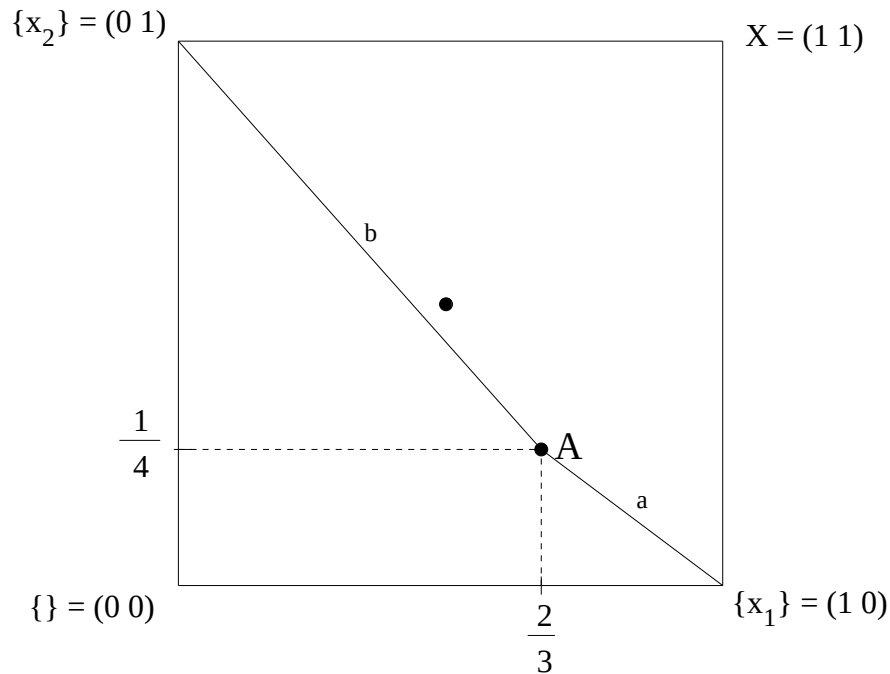


Figura 5.1: Interpretação Geométrica da Medida de Entropia Fuzzy

Note agora, na Figura 5.2, a representação do quadrado fuzzy unitário. Nela estão destacados o conjunto A , seu complementar A^c , sua sobreposição $A \cup A^c$ e sua

subposição $A \cap A^c$. Além disso, o contador sigma de cada um destes conjuntos são representados como flechas. Movendo o conjunto A estes outros conjuntos se moverão junto com ele. Tendo esse movimento em mente, o passo seguinte é observar os contadores sigma da sobreposição e da subposição.

Estudando estes contadores percebe-se que quanto mais o conjunto A se aproxima de um vértice, mais o contador $M(A \cap A^c)$ se aproxima de 0. Por sua vez, o contador $M(A \cup A^c)$ se aproxima de 2. Ou seja, neste caso, a razão entre eles vai para 0.

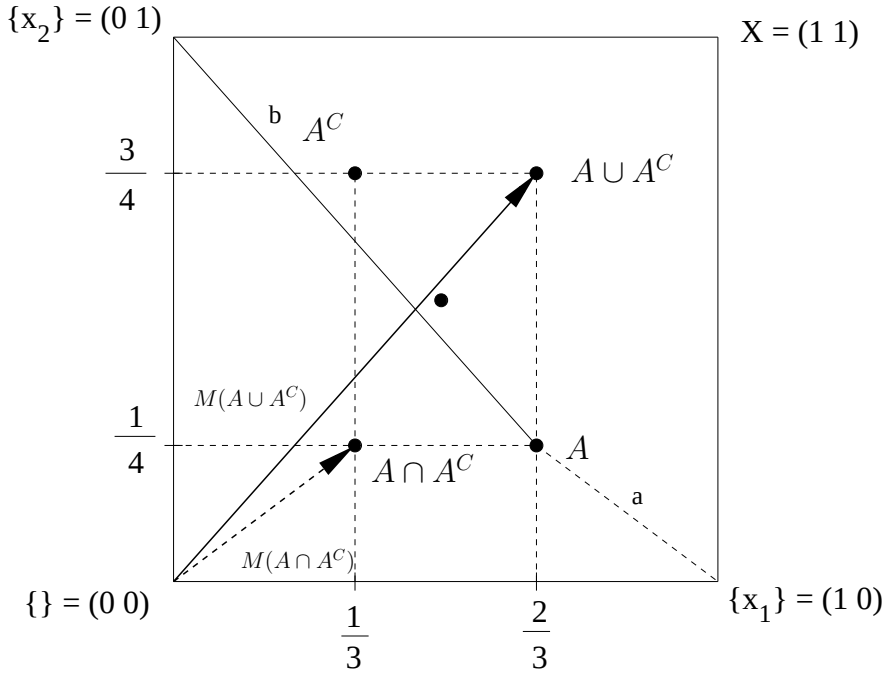


Figura 5.2: Geometria do Teorema da Entropia Fuzzy

Além disso, quanto mais o conjunto A se aproxima do ponto médio, mais o contador $M(A \cap A^c)$ se aproxima do contador $M(A \cup A^c)$. Ou seja, a razão entre eles vai aumentando lentamente conforme o conjunto A se movimenta em direção ao ponto médio, até atingir 1 neste ponto.

Isso mostra que é possível caracterizar a “fuzziness” dos conjuntos no interior do hipercubo pela sua sobreposição e pela sua subposição. Por este argumento, nota-se ainda que a razão entre a subposição e a sobreposição de um conjunto aumenta da mesma forma que sua entropia fuzzy, o que indica que existe alguma relação entre elas. Desta forma, chega-se ao primeiro resultado importante deste capítulo, o Teorema da Entropia Fuzzy.

Teorema 5.1 (TEOREMA DA ENTROPIA FUZZY). *Seja A um ponto no hipercubo I^n . Então sua entropia é dada por:*

$$E(A) = \frac{M(A \cap A^c)}{M(A \cup A^c)}.$$

Demonstração. Basta analisar os valores $a = \ell^1(A, A_p)$ e $b = \ell^1(A, A_l)$.

Observe que

$$|\varphi_A(x_i) - \varphi_{A_p}(x_i)| = \begin{cases} \varphi_A(x_i) & \text{se } \varphi_{A_p}(x_i) = 0 \\ 1 - \varphi_A(x_i) & \text{se } \varphi_{A_p}(x_i) = 1 \end{cases}.$$

Além disso, se $\varphi_A(x_i) < \frac{1}{2}$ vale que $\varphi_{A_p}(x_i) = 0$. Daí, $|\varphi_A(x_i) - \varphi_{A_p}(x_i)| = \varphi_A(x_i) < \frac{1}{2}$.

Também, se $\varphi_A(x_i) > \frac{1}{2}$ vale que $\varphi_{A_p}(x_i) = 1$. Disto segue que, $|\varphi_A(x_i) - \varphi_{A_p}(x_i)| = |\varphi_A(x_i) - 1| = 1 - \varphi_A(x_i) < \frac{1}{2}$.

Se $\varphi_A(x_i) = \frac{1}{2}$ então $|\varphi_A(x_i) - \varphi_{A_p}(x_i)| = \left| \pm \frac{1}{2} \right| = \frac{1}{2}$.

Portanto, nestes três casos tem-se

$$|\varphi_A(x_i) - \varphi_{A_p}(x_i)| = \min(\varphi_A(x_i), 1 - \varphi_A(x_i)).$$

Logo,

$$\begin{aligned} \ell^1(A, A_p) &= \sum_{i=1}^n |\min(\varphi_A(x_i), 1 - \varphi_A(x_i))| = \sum_{i=1}^n |\min(\varphi_A(x_i), \varphi_{A^c}(x_i))| \\ &= \sum_{i=1}^n \varphi_{A \cap A^c}(x_i) = M(A \cap A^c). \end{aligned}$$

Por outro lado,

$$|\varphi_A(x_i) - \varphi_{A_l}(x_i)| = \begin{cases} \varphi_A(x_i) & \text{se } \varphi_{A_l}(x_i) = 0 \\ 1 - \varphi_A(x_i) & \text{se } \varphi_{A_l}(x_i) = 1 \end{cases}.$$

Se $\varphi_A(x_i) < \frac{1}{2}$ vale que $\varphi_{A_l}(x_i) = 1$. Disto segue que, $|\varphi_A(x_i) - \varphi_{A_l}(x_i)| = |\varphi_A(x_i) - 1| = 1 - \varphi_A(x_i) > \frac{1}{2}$.

Também, se $\varphi_A(x_i) > \frac{1}{2}$ valerá que $\varphi_{A_l}(x_i) = 0$. Daí, $|\varphi_A(x_i) - \varphi_{A_l}(x_i)| = \varphi_A(x_i) > \frac{1}{2}$.

Se $\varphi_A(x_i) = \frac{1}{2}$ então $|\varphi_A(x_i) - \varphi_{A_l}(x_i)| = \left| \pm \frac{1}{2} \right| = \frac{1}{2}$.

Da mesma maneira, nestes três casos tem-se

$$|\varphi_A(x_i) - \varphi_{A_l}(x_i)| = \min(\varphi_A(x_i), 1 - \varphi_A(x_i)).$$

Logo,

$$\begin{aligned}
\ell^1(A, A_l) &= \sum_{i=1}^n |\max(\varphi_A(x_i), 1 - \varphi_A(x_i))| = \sum_{i=1}^n |\max(\varphi_A(x_i), \varphi_{A^c}(x_i))| \\
&= \sum_{i=1}^n \varphi_{A \cup A^c}(x_i) = M(A \cup A^c).
\end{aligned}$$

Portanto,

$$E(A) = \frac{a}{b} = \frac{\ell^1(A, A_p)}{\ell^1(A, A_l)} = \frac{M(A \cap A^c)}{M(A \cup A^c)}.$$

□

O quadrado da Figura 5.2 ilustra o teorema para o caso $n = 2$. Para o caso geral basta considerar a simetria, pois, cada um dos pontos $A, A^c, A \cap A^c$ e $A \cup A^c$ é equidistante do respectivo vértice mais próximo. A distância comum é igual a a . Da mesma maneira, cada ponto é igualmente distante do vértice mais distante. A distância comum é igual a b . Como já observado, a menor distância será igual ao contador $M(A \cap A^c)$ e a maior distância igual ao contador $M(A \cup A^c)$. Desta forma, o teorema da entropia fuzzy estabelece que “fuzziness” baseia-se em contar violações da lei da contradição, balanceado com contar violações da lei do terceiro excluído.

5.1.2 Consequências

O teorema da entropia fuzzy explica porque a lógica fuzzy começa onde a lógica clássica termina. Quando conjuntos ou proposições obedecem às leis da não-contradição e do terceiro excluído, a sobreposição é vazia e a subposição é exaustiva. Então, sempre se tem $M(A \cap A^c) = 0$ e $M(A \cup A^c) = n$, e daí $E(A) = 0$.

O teorema da entropia fuzzy também permite entender porque as principais operações sobre conjuntos fuzzy propostas por Lukasiewicz e Zadeh, foram o mínimo (intersecção), o máximo (união) e a negação (complementar). As normas triangulares ou *T-normas*, que generalizam a conjunção ou intersecção, e as normas triangulares duais, as *co-normas*, que generalizam a disjunção ou união, não são tão eficazes como o mínimo, o máximo e a negação, pois se tais normas forem utilizadas como operações sobre os conjuntos fuzzy o Teorema da Entropia nem sempre será útil, já que para as desigualdades das normas triangulares valerá

$$T(x, y) \leq \min(x, y) \leq \max(x, y) \leq C(x, y),$$

o que implica que, substituir o mínimo com qualquer T no termo $M(A \cap A^c)$, tornará menor o numerador de $E(A)$. Além disso, substituir o máximo por qualquer C no termo $M(A \cup A^c)$ tornará maior o denominador de $E(A)$. Então, quaisquer T ou C diferentes do mínimo ou do máximo, respectivamente, tornarão a razão menor,

realmente, estritamente menor se A é não-clássico. Neste caso, o Teorema da Entropia não é interessante, pois a medida pseudo-entrópica resultante não assume o valor 1 no ponto médio, embora continue sendo máxima nele, como é possível ver dos exemplos abaixo.

Exemplo 5.2. Por exemplo, tomando a T-norma produto $T(x, y) = xy$ e sua co-norma dual de De Morgan $C(x, y) = 1 - T(1 - x, 1 - y) = x + y - xy$, e tomando $A = \left(\frac{1}{2} \frac{1}{2}\right)$ calcula-se a medida pseudo-entrópica dada por

$$E^*(A) = \frac{T(x_1, x_2)}{C(x_1, x_2)} = \frac{x_1 x_2}{x_1 + x_2 - x_1 x_2} = \frac{\frac{1}{4}}{\frac{1}{2} + \frac{1}{2} - \frac{1}{4}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

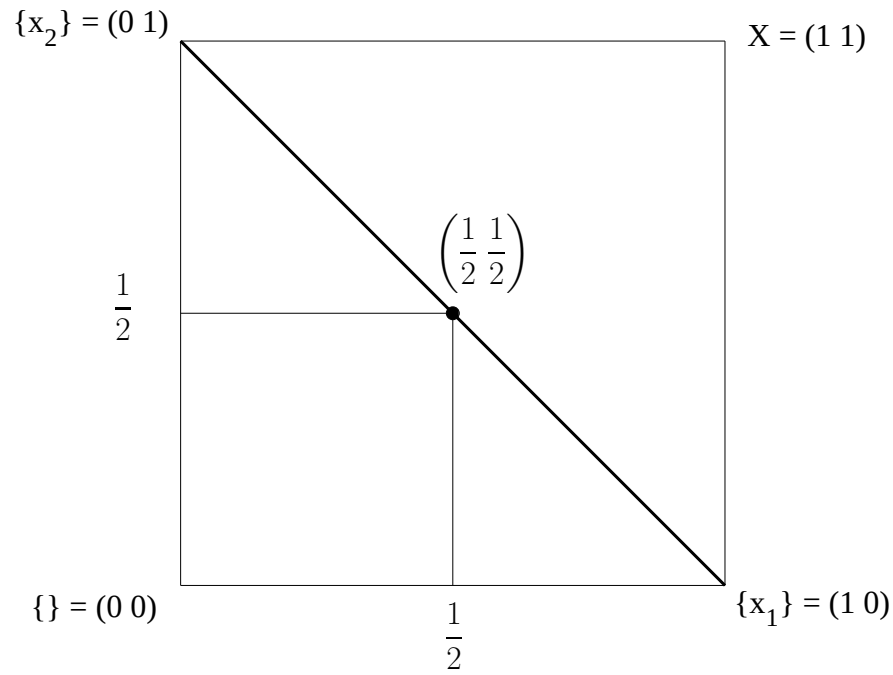
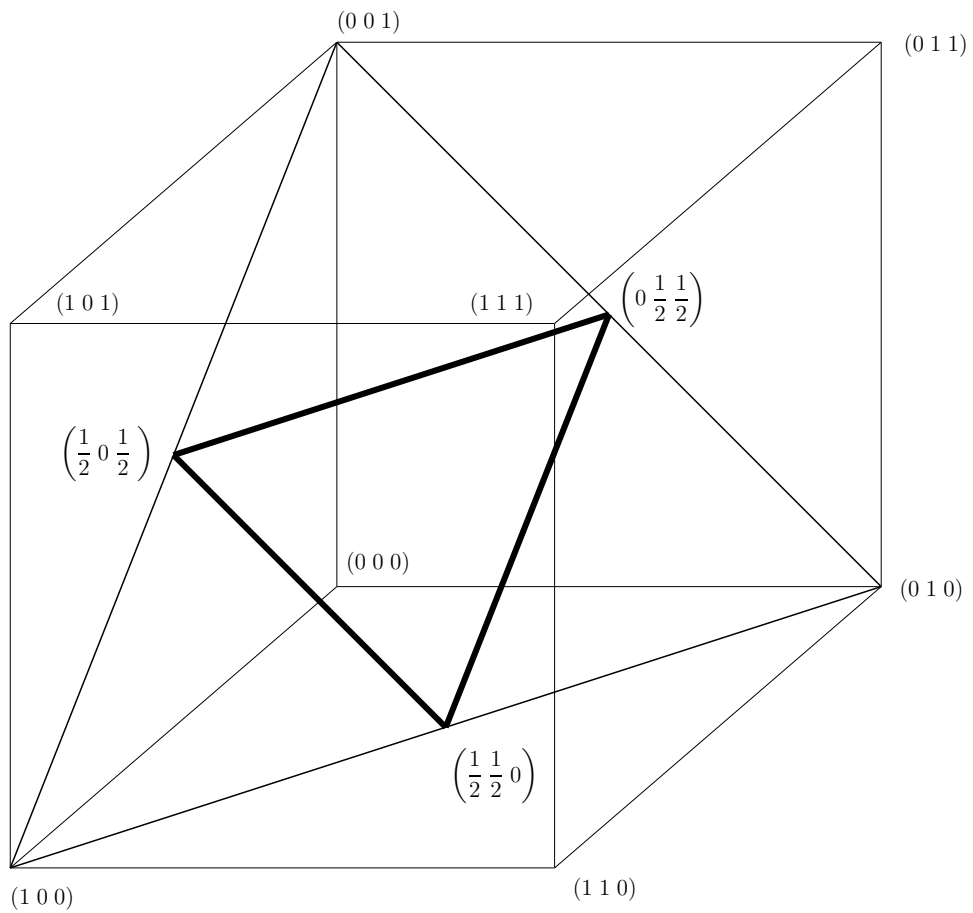
Exemplo 5.3. Tomando a T-norma da soma limitada $T(x, y) = \max(0, x + y - 1)$ e a dual de De Morgan $C(x, y) = \min(1, x + y)$, toma-se o ponto médio $M = \left(\frac{1}{2} \frac{1}{2}\right)$ e obtém-se a medida pseudo-entrópica dada por

$$\overline{E}(M) = \frac{T(x_1, x_2)}{C(x_1, x_2)} = \frac{\max(0, x_1 + x_2 - 1)}{\min(1, x_1 + x_2)} = \frac{\max(0, \frac{1}{2} + \frac{1}{2} - 1)}{\min(1, \frac{1}{2} + \frac{1}{2})} = \frac{0}{1} = 0.$$

Por outro lado, o Teorema da Entropia falha, em geral, quando o operador negação ou complementação $N(x) = 1 - x$ é trocado por outro. Por exemplo, falha se a negação for trocada pelo operador parametrizado $N_a(x) = \frac{(1 - x)}{(1 + ax)}$, onde $a > -1$ e $a \neq 0$. Isso ocorre porque, com esta troca, a definição de entropia fuzzy é alterada.

5.1.3 Aplicação

É interessante notar o que ocorre quando se toma, no hipercubo I^n , as distribuições de probabilidade, ou seja, todos os conjuntos A em I^n com $M(A) = 1$. Nesse caso obtém-se, para cada n , que a união destes conjuntos formam o simplexo S^{n-1} . No quadrado unitário, o simplexo é igual à linha diagonal principal de inclinação igual a -1 (S^1). No cubo unitário, de três dimensões, é igual a um triângulo sólido (S^2), cujos vértices são os conjuntos fuzzy dados por $\left(\frac{1}{2} \frac{1}{2} 0\right)$, $\left(\frac{1}{2} 0 \frac{1}{2}\right)$ e $\left(0 \frac{1}{2} \frac{1}{2}\right)$. No hipercubo de quatro dimensões é igual a um tetraedro (S^3), cujos vértices são os conjuntos fuzzy com três coordenadas iguais a $\frac{1}{2}$ e a outra igual a 0, e assim por diante. As Figuras 5.3 e 5.4 ilustram os casos para as dimensões 2 e 3.

Figura 5.3: Simplexo S^2 no Quadrado I^2 Figura 5.4: Simplexo S^3 no Cubo I^3

5.1.4 Aplicação em uma Distribuição de Probabilidades

Incluindo a condição de que o valor x_i na distribuição de probabilidade A satisfaça $p_i \leq \frac{1}{2}$ tem-se que $E(A) = \frac{1}{n-1}$.

De fato, pois, conforme visto na Observação 4.3 quando $x_i \leq \frac{1}{2}, \forall x_i \in X$, vale que $A = A \cap A^c$ e $A^c = A \cup A^c$. Então, o Teorema da Entropia Fuzzy leva a

$$E(A) = \frac{M(A \cap A^c)}{M(A \cup A^c)} = \frac{M(A)}{M(A^c)}.$$

Como A é uma distribuição de probabilidade $M(A) = 1$. Por outro lado,

$$M(A^c) = \sum_{i=1}^n (1 - \varphi_A(x_i)) = n - \sum_{i=1}^n \varphi_A(x_i) = n - 1.$$

Então,

$$E(A) = \frac{M(A \cap A^c)}{M(A \cup A^c)} = \frac{M(A)}{M(A^c)} = \frac{1}{n-1}.$$

Por outro lado, se não for uma distribuição de probabilidade, Kosko [18] mostra que neste caso $E(A) < \frac{1}{n-1}$.

Exemplo 5.4. Sejam o hipercubo I^4 e o conjunto fuzzy $A = \left(\frac{1}{4} \frac{1}{4} \frac{1}{4} \frac{1}{4}\right)$. Então,

$$A^c = \left(\frac{3}{4} \frac{3}{4} \frac{3}{4} \frac{3}{4}\right), \quad A \cap A^c = \left(\frac{1}{4} \frac{1}{4} \frac{1}{4} \frac{1}{4}\right),$$

$$A \cup A^c = \left(\frac{3}{4} \frac{3}{4} \frac{3}{4} \frac{3}{4}\right), \quad M(A \cap A^c) = 1,$$

$$M(A \cup A^c) = 3 \text{ e}$$

$$E(A) = \frac{M(A \cap A^c)}{M(A \cup A^c)} = \frac{1}{3} = \frac{1}{4-1}.$$

Por outro lado, para $B = \left(\frac{1}{8} \frac{5}{8} \frac{1}{8} \frac{1}{8}\right)$ obtém-se

$$B^c = \left(\frac{7}{8} \frac{3}{8} \frac{7}{8} \frac{7}{8}\right), \quad B \cap B^c = \left(\frac{1}{8} \frac{3}{8} \frac{1}{8} \frac{1}{8}\right),$$

$$B \cup B^c = \left(\frac{7}{8} \frac{5}{8} \frac{7}{8} \frac{7}{8}\right), \quad M(B \cap B^c) = \frac{3}{4},$$

$$M(B \cup B^c) = \frac{13}{4} \text{ e}$$

$$E(B) = \frac{M(B \cap B^c)}{M(B \cup B^c)} = \frac{3}{13} < \frac{1}{3} = \frac{1}{4-1}.$$

Exemplo 5.5. Como último exemplo, sejam o hipercubo I^4 e o conjunto fuzzy $A = \left(\frac{1}{5} \frac{2}{5} \frac{1}{5} \frac{1}{5}\right)$. Então

$$A^c = \left(\frac{4}{5} \frac{3}{5} \frac{4}{5} \frac{4}{5}\right), \quad A \cap A^c = \left(\frac{1}{5} \frac{2}{5} \frac{1}{5} \frac{1}{5}\right),$$

$$A \cup A^c = \left(\frac{4}{5} \frac{3}{5} \frac{4}{5} \frac{4}{5}\right), \quad M(A \cap A^c) = 1,$$

$$M(A \cup A^c) = 3 \text{ e}$$

$$E(A) = \frac{M(A \cap A^c)}{M(A \cup A^c)} = \frac{1}{3} = \frac{1}{4-1}.$$

Esse resultado, como os exemplos ilustram, mostra que a distribuição uniforme $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ maximiza a entropia fuzzy sobre S^n , mas não de uma única forma. Portanto, a entropia fuzzy diverge da medida pela entropia probabilística, já que nesta última a distribuição uniforme é maximizada de forma única.

5.2 O Teorema da Inclusão

Existem duas formas de analisar a medida do grau de inclusão, uma algébrica e outra geométrica. Nesta seção serão analisadas estas duas formas de dedução e será demonstrada a equivalência entre elas.

5.2.1 Dedução Algébrica da Medida do Grau de Inclusão

Será estudado, agora, o significado algébrico do grau de inclusão $S(A, B)$ levando em conta a visão geométrica de conjuntos como pontos dada na Figura 1.4.

O objetivo é medir $S(A, B)$, ou seja, o grau com que A é um subconjunto de B . Para fazer isso suponha que, para um determinado i , o elemento $x_i \in X$ não satisfaça a inclusão. Uma vez que define-se $A \subseteq B$ se, e somente se, $\varphi_A(x) \leq \varphi_B(x), \forall x \in X$, dizer que um elemento x_i de X não satisfaz a inclusão significa dizer que $\varphi_A(x_i) > \varphi_B(x_i)$. Quando isso acontece diz-se que o elemento x_i viola a inclusão. Nos casos em que ocorrem violações, A não é totalmente um subconjunto de B .

Exemplo 5.6. Sejam A e B dois subconjuntos fuzzy em X . Suponha que a desigualdade de pertinência-inclusão, $\varphi_A(x) \leq \varphi_B(x)$ seja verdadeira para todos os elementos $x \in X$, exceto um. Ou seja, se X contiver 100 elementos, 99 satisfazem a inclusão. Neste caso é natural tomar-se $S(A, B) = 0.99$. É claro que, nesta situação, A será, em grande medida, subconjunto de B . Por outro lado, também é intuitivo que, se X tivesse um trilhão de elementos e apenas um não satisfizesse a desigualdade, A seria

praticamente um subconjunto de B . O grau de inclusão, nesse caso, pode ser tomado naturalmente como $S(A, B) = 0.999999999999$.

Esse exemplo sugere que é interessante contar o quanto a relação de inclusão é violada em grandeza para cada x_i , ou seja, o quanto $\varphi_A(x_i)$ é maior do que $\varphi_B(x_i)$. Cada vez que isso ocorre tem-se uma violação da inclusão $A \subseteq B$. Consequentemente, o grau de inclusão diminui. Além disso, é necessário levar em conta a frequência com que isto ocorre, ou seja, para quantos x_i a violação acontece pois, naturalmente, o grau de inclusão diminuirá conforme a frequência destas violações aumentar.

Formalmente, para cada $x_i \in X$,

- se $\varphi_A(x_i) > \varphi_B(x_i)$, então $0 < \varphi_A(x_i) - \varphi_B(x_i) < 1$;
- caso contrário, $\varphi_A(x_i) - \varphi_B(x_i) = 0$.

Ou seja, para cada x_i , obtém-se o quanto ele viola a relação de inclusão por considerar-se o valor de $\max(0, \varphi_A(x_i) - \varphi_B(x_i))$.

A contagem da soma destas violações relaciona-se, por outro lado, com o grau com que A contém B , ou seja, com o que se chama de grau de incluir. Uma vez que o grau de incluir aumenta conforme o grau de inclusão diminui tem-se, denotando-se a medida do grau de incluir por $S^c(A, B)$, que, obviamente,

$$S^c(A, B) = 1 - S(A, B),$$

ou seja, o grau com que A contém B é o complementar do grau com que A está contido em B . A soma destes máximos, dada por,

$$\sum_{x_i \in X} \max(0, \varphi_A(x_i) - \varphi_B(x_i)).$$

Para chegar a uma medida de grau de inclusão que possa ser associada a uma função de pertinência o valor dessa soma deve estar entre 0 e 1 devendo ser, portanto, normalizado. O contador $M(A)$ provê um fator de normalização simples, mas eficiente para esse valor. Chega-se a esta conclusão pelo exame de casos limites na abordagem geométrica dos conjuntos fuzzy. Pode-se supor que $M(A) > 0$, pois $M(A) = 0$ se, e somente se, A é vazio, e o conjunto vazio satisfaz trivialmente a inclusão e, portanto, é subconjunto de qualquer conjunto. A medida do grau de incluir normalizada é dada por:

$$S^c(A, B) = \frac{\sum_{x \in X} \max(0, \varphi_A(x_i) - \varphi_B(x_i))}{M(A)}.$$

A medida do grau de inclusão é o complementar desta razão, logo,

$$S(A, B) = 1 - \frac{\sum_{x \in X} \max(0, \varphi_A(x_i) - \varphi_B(x_i))}{M(A)} \geq 0.$$

A medida de grau de inclusão pode parecer complicada no início, mas ela se comporta como deveria, pois $S(A, B) = 1$ se, e somente se, todo elemento de A satisfaz a relação de inclusão de Zadeh com B e $S(A, B) = 0$ se, e somente se, B é o conjunto vazio. Assim, $0 \leq S(A, B) \leq 1$.

Exemplo 5.7. Sejam os vetores $A = (0, 3 \ 0, 0 \ 0, 6 \ 0, 7)$ e $B = (0, 7 \ 0, 3 \ 0, 4 \ 0, 9)$. Nenhum conjunto é totalmente subconjunto do outro. A é quase subconjunto de B , mas não completamente, desde que $\varphi_A(x_3) - \varphi_B(x_3) = 0, 6 - 0, 4 = 0, 2 > 0$. Daí $S(A, B) = 1 - \frac{0, 2}{1, 6} = 1 - \frac{1}{8} = \frac{7}{8}$. Da mesma maneira: $\varphi_B(x_1) - \varphi_A(x_1) = 0, 7 - 0, 3 = 0, 4 > 0$, $\varphi_B(x_2) - \varphi_A(x_2) = 0, 3 - 0, 0 = 0, 3 > 0$ e $\varphi_B(x_4) - \varphi_A(x_4) = 0, 9 - 0, 7 = 0, 2 > 0$. Portanto, $S(B, A) = 1 - \frac{0, 4}{2, 3} = \frac{19}{23}$.

5.2.2 Dedução Geométrica da Medida do Grau de Inclusão

Agora, será analisada uma dedução geométrica do grau de inclusão. Ou o conjunto A está no hiperretângulo $F(2^B)$ ou não está. De forma intuitiva, $S(A, B)$ se aproxima da unidade conforme A se aproxima do conjunto potência fuzzy $F(2^B)$. Por outro lado, $S(A, B)$ deve diminuir, e a medida do grau de incluir $1 - S(A, B)$ (complementar de $S(A, B)$) aumentar, conforme A se move para fora de $F(2^B)$. Então o ponto chave é a medida: Quão perto A está de $F(2^B)$?

Observação 5.2. A fim de simplificar a notação, a distância ℓ^p da fórmula (4.3) para $p = 1$ será denotada por $d(A, F(2^B))$, ou seja, $d(A, F(2^B)) = \ell^1(A, F(2^B))$.

Seja $B' \subset B$, um ponto qualquer no hiperretângulo $F(2^B)$. Considere $d(A, B')$ a distância de A a B' . Então, a distância $d(A, F(2^B))$ será igual ao ínfimo destas distâncias pois, o hiperretângulo é fechado e limitado, logo é compacto e, desta forma, existe um B^* de $F(2^B)$ tal que $d(A, B^*)$ assume o valor mínimo, ou seja,

$$d(A, F(2^B)) = \inf \{d(A, B') : B' \in F(2^B)\} = d(A, B^*).$$

Uma interpretação natural define a relação de inclusão como essa distância, uma vez que o grau de incluir aumenta com ela e o grau de inclusão diminui. Para manter tanto o grau de incluir como o grau de inclusão no intervalo unitário essa distância precisa ser normalizada.

Escolha do Fator de Normalização

A maneira simples de normalizar um valor é através de uma constante. A distância máxima no hipercubo unitário, $n^{\frac{1}{p}}$ no caso geral ℓ^p e n no caso ℓ^1 , parece uma boa

candidata para tal constante. Desta forma, o grau de inclusão poderia ser tomado como

$$S(A, B) = 1 - \frac{d(A, B^*)}{n}.$$

No entanto, analisando os casos extremos percebe-se que este candidato à grau de inclusão falha quando B é o conjunto vazio, já que este não tem subconjuntos. Nesta situação, $d(A, B^*) = d(A, B) = M(A)$. Então, a expressão anterior poderia ser escrita como:

$$S(A, \emptyset) = 1 - \frac{M(A)}{n},$$

ou seja, sem depender de B^* . Daí, $S(A, \emptyset) = 0$ e $S(\emptyset, A) = 1$.

Por outro lado, a expressão vale para $A = X$ e, neste caso, $M(A) = n$. Assim, $M(A)$ é uma escolha melhor para fator de normalização, pois atende aos dois casos extremos.

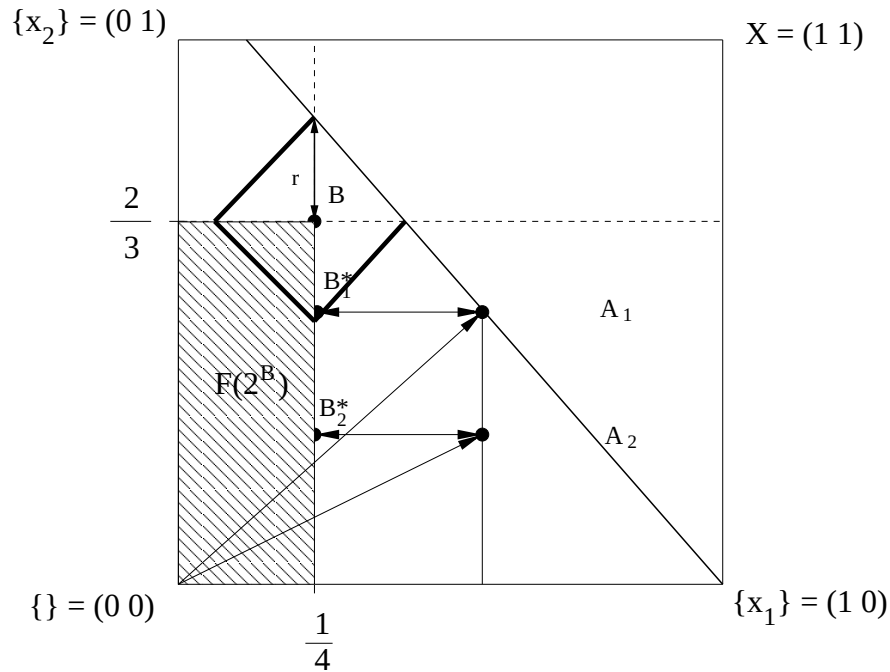


Figura 5.5: Dependência da Inclusão em relação ao Contador $M(A)$

Ao normalizar por n equaliza-se todos os pontos equidistantes de $F(2^B)$. Considere os pontos A_1 e A_2 na Figura 5.5. Ambos são equidistantes do ponto mais próximo de $F(2^B)$: $d(A_1, B_1^*) = d(A_2, B_2^*)$. Mas A_1 é mais próximo de B que A_2 . Em particular, A_1 é mais próximo da linha horizontal definida pelo valor $\varphi_B(x_2) = \frac{2}{3}$. O contador $M(A)$ reflete isto, já que $M(A_1) > M(A_2)$. A diferença $M(A_1) - M(A_2)$ mede a diferença dos x_1 e reflete em $d(A_1, B) < d(A_2, B)$. Percebe-se, então, que o contador $M(A)$ está relacionado a esta distância. Isto pode ser confirmado por se estudar os casos extremos de proximidade de A a B na Figura 5.5. Observando-se o “círculo” (na

distância ℓ^1) centrado em B , percebe-se que se $\varphi_A(x_i) > \varphi_B(x_i)$ para algum i tem-se que $d(A, B) = M(A) - M(B)$.

Como $F(2^B)$ se ajusta à origem, o contador $M(A)$, em qualquer um dos outros $2^n - 1$ hiperretângulos poderia ser maior que o contador $M(B^*)$ do ponto B^* de $F(2^B)$ mais próximo de A . A escolha de n para a normalização do candidato a grau de inclusão funciona tanto para os $2^n - 1$ casos de hiperretângulos que contém A , quanto para aqueles nos quais A está no hiperretângulo $F(2^B)$. Cada ponto de cada hiperretângulo envolve uma combinação diferente de condições violadas e satisfeitas. A escolha de $M(A)$ para a normalização reflete esta estrutura de violações e satisfaz de forma única as condições extremas.

A escolha de $M(A)$ para a normalização conduz à seguinte definição de grau de inclusão:

$$S(A, B) = 1 - \frac{d(A, B^*)}{M(A)}. \quad (5.1)$$

Essa medida equivale a medida $S(A, B)$ da seção 5.2.1 calculada algebricamente.

Observação 5.3. Pode-se pensar numa versão do Teorema de Pitágoras para a distância ℓ^p no hipercubo fuzzy. Para o caso onde $p = 1$ tem-se:

$$d(A, B) = d(A, C) + d(B, C).$$

A versão para ℓ^2 utilizará o quadrado destas distâncias e dá a versão usual do Teorema de Pitágoras. Para o caso geral tem-se

$$||A - B||^p = ||A - B^*||^p + ||B^* - B||^p,$$

ou seja,

$$\sum_{i=1}^n |a_i - b_i|^p = \sum_{i=1}^n |a_i - b_i^*|^p + \sum_{i=1}^n |b_i^* - b_i|^p.$$

Essa igualdade vale para todo $p \geq 1$ porque sempre se tem $b_i^* = a_i$ ou $b_i^* = b_i$. Isso sugere que o conjunto mais próximo B^* tem um status especial. O teorema da inclusão de conjuntos confirma esta sugestão. O termo “ortogonalidade” é usado para referir-se a essa relação pitagórica.

Demonstrando a Equivalência

Seja B' um subconjunto qualquer de B . Então, por definição, o subconjunto mais próximo de A , B^* satisfaz a desigualdade:

$$\sqrt[p]{\sum_{i=1}^n |a_i - b_i^*|^p} \leq \sqrt[p]{\sum_{i=1}^n |a_i - b_i'|^p}$$

onde, por conveniência, utiliza-se $a_i = \varphi_A(x_i)$, $b_i = \varphi_B(x_i)$, $b'_i = \varphi_{B'}(x_i)$ e $b_i^* = \varphi_{B^*}(x_i)$.

Lema 5.1.

$$\max(0, a_i - b_i) = |a_i - b_i^*|.$$

Demonstração. Tem-se $a_i \geq b_i^*$. Considere dois casos:

- se i é tal que: $a_i = b_i^*$ e $|a_i - b_i^*| = 0$ e não ocorre violação $a_i > b_i$, logo $a_i \leq b_i$ e $\max(0, a_i - b_i) = 0 = |a_i - b_i^*|$.

Então se $a_i = b_i^*$ vale para todo i que $A = B^*$.

- se i é tal que $a_i > b_i^*$ então ocorre uma violação $a_i > b_i$, assim pode-se tomar $b_i = b_i^*$ e $a_i - b_i = a_i - b_i^*$ e $\max(0, a_i - b_i) = |a_i - b_i^*|$ e vale para todo i que $B = B^*$.

□

Então as medidas de grau de inclusão $S(A, B)$ apresentadas na Equação 5.1 e 5.2 são equivalentes.

Essa demonstração também fornece uma caracterização precisa do subconjunto ótimo B^* , a saber,

Lema 5.2.

$$B^* = A \cap B.$$

Demonstração. Se uma violação ocorre, então $a_i > b_i$ e $b_i = b_i^*$. Então $\min(a_i, b_i) = b_i^*$. Por outro lado, se $a_i = b_i^*$ então $\min(a_i, b_i) = b_i^*$ já que B^* é o subconjunto de B mais próximo de A . Logo $B^* = A \cap B$. □

Adicionalmente, uma vez que $B^* = A^*$, tanto B^* é o subconjunto de B mais próximo de A , quanto o subconjunto de A mais próximo de B .

$$d(B, F(2^A)) = d(B, A^*) = d(B, B^*) \text{ e } d(A, F(2^B)) = d(A, B^*) = d(A, A^*).$$

A Figura 5.6 ilustra que $B^* = A \cap B = A^*$ e identifica o conjunto com o hiperretângulo $F(2^A) \cap F(2^B)$ que tem o contador sigma igual a $M(A \cap B)$.

Note ainda que vale

$$d(A, B^*) = M(A) - M(A \cap B).$$

Dividindo ambos os lados dessa igualdade por $M(A)$ e rearranjando obtém-se

$$\frac{d(A, B^*)}{M(A)} = \frac{M(A) - M(A \cap B)}{M(A)} = 1 - \frac{M(A \cap B)}{M(A)}. \quad (5.2)$$

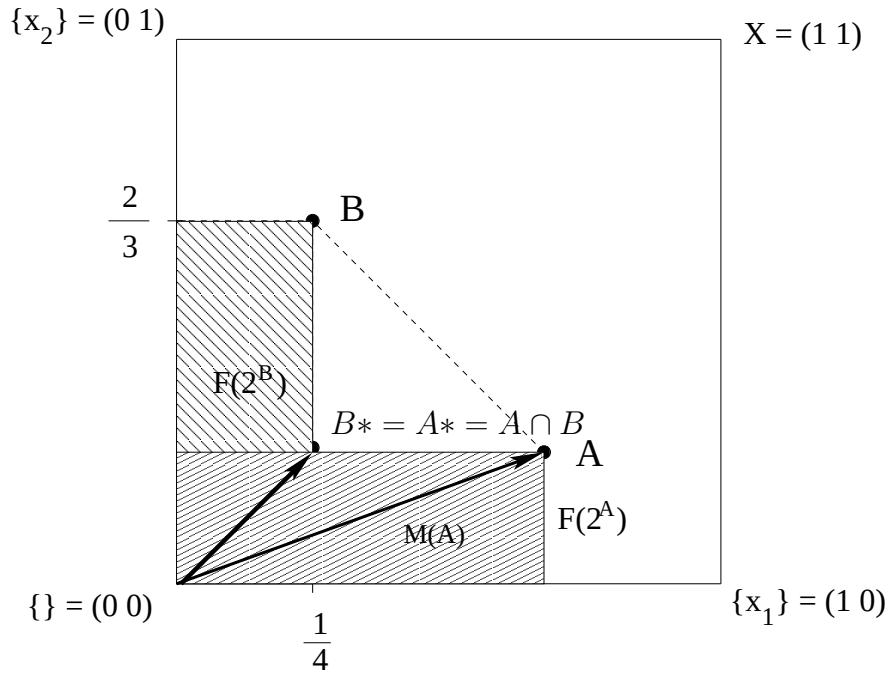


Figura 5.6: Ilustração do Teorema da Inclusão

ou seja, uma caracterização estrutural detalhada de inclusão, o teorema da inclusão.

Observe ainda que, $M(A) = 0$ se, e somente se, $A = \emptyset$. Além disso, neste caso, $S(\emptyset, B) = 1$, para qualquer conjunto B . Assim, por hipótese, no teorema $M(A)$ será tomado como maior do que zero.

Teorema 5.2 (TEOREMA DA INCLUSÃO). *Sejam A e B dois conjuntos fuzzy. Então, se $M(A) > 0$,*

$$S(A, B) = \frac{M(A \cap B)}{M(A)}.$$

Demonstração. Pela definição

$$S(A, B) = 1 - \frac{\sum_{x \in X} \max(0, \varphi_A(x_i) - \varphi_B(x_i))}{M(A)}.$$

Mas, como já observado,

$$d(A, B^*) = \sum_{i=1}^n \max(0, \varphi_A(x_i) - \varphi_B(x_i)).$$

Daí,

$$S(A, B) = 1 - \frac{d(A, B^*)}{M(A)}.$$

Então, por (5.2), chega-se a

$$S(A, B) = 1 - 1 + \frac{M(A \cap B)}{M(A)} = \frac{M(A \cap B)}{M(A)}.$$

□

Corolário 5.1. *Sejam A e B dois conjuntos fuzzy. Então, se $M(A) > 0$ vale que*

$$S(A, B) = \frac{M(B)S(B, A)}{M(A)}.$$

Demonstração. Consequência direta do Teorema. Uma vez que $S(B, A) = \frac{M(A \cap B)}{M(B)}$ obtém-se

$$M(A \cap B) = M(B)S(B, A).$$

Dividindo os dois lados por $M(A)$ chega-se a

$$\frac{M(A \cap B)}{M(A)} = \frac{M(B)S(B, A)}{M(A)}$$

e, então,

$$S(A, B) = \frac{M(B)S(B, A)}{M(A)}.$$

□

A forma quociente da medida do grau de inclusão $S(A, B)$ é a mesma da condição de probabilidade $P(B|A)$ em (4.1.1), ou seja, é similar à do Teorema de Bayes para distribuições de probabilidade. A forma quociente para a medida do grau de inclusão $S(A, B)$ foi calculada, mas a condição de probabilidade foi assumida por hipótese. Uma vez que toda probabilidade é uma probabilidade condicional, já que,

$$P(A) = \frac{P(A)}{1} = \frac{P(A \cap X)}{P(X)} = P(A|X),$$

fica implícita uma sugestão de que probabilidade pode ser reduzida à inclusão. Para isso considere a interpretação física da aleatoriedade como frequência relativa $\frac{n_A}{n}$, onde $n_A = P(A)$, denota o número de sucessos que ocorrem em n tentativas. Assim, a frequência relativa de probabilidade pode ser vista como $S(X, A)$, o grau com que o superconjunto X , o espaço todo, é um subconjunto de seu subconjunto A .

Corolário 5.2. *Seja A é um subconjunto clássico de X . Então*

$$S(X, A) = P(A|X).$$

Demonstração. De fato,

$$S(X, A) = \frac{M(A \cap X)}{M(X)} = \frac{M(A)}{M(X)} = \frac{n_A}{n} = \frac{P(A)}{P(X)} = \frac{P(A \cap X)}{P(X)} = P(A|X).$$

□

5.3 O Teorema da Entropia-Inclusão

O Teorema da Entropia Fuzzy e o Teorema da Inclusão foram deduzidos de maneira independente, como resultados principais da geometria dos conjuntos-como-pontos num hipercubo unitário. Os dois teoremas envolvem razões de cardinalidade, portanto, é natural pensar que deve existir uma conexão entre eles.

O Teorema da Entropia-Inclusão mostrará que a conexão envolve $A \cap A^c$ e $A \cup A^c$. O teorema associa entropia fuzzy ao grau de inclusão. Sendo assim, a inclusão se torna fundamental para a caracterização da quantidade de “fuzziness” e de probabilidade também, uma vez que, como visto no Corolário 5.2, o grau de inclusão pode ser associado à probabilidade no caso de conjuntos clássicos.

Teorema 5.3 (TEOREMA DA ENTROPIA-INCLUSÃO). *Seja A um conjunto fuzzy no hipercubo I^n . Então a entropia de A é dada por:*

$$E(A) = S(A \cup A^c, A \cap A^c).$$

Demonstração. Basta substituir B e A no teorema da inclusão por $A \cap A^c$ e $A \cup A^c$, respectivamente. Tem-se então que

$$S(A \cup A^c, A \cap A^c) = \frac{M((A \cup A^c) \cap (A \cap A^c))}{M(A \cup A^c)} = \frac{M(A \cap A^c)}{M(A \cup A^c)} = E(A).$$

□

O Teorema da Entropia-Inclusão descreve uma relação peculiar. Ele mede a “fuzziness” ou a ambiguidade com a qual o superconjunto $A \cup A^c$ é um subconjunto de seu subconjunto $A \cap A^c$, o grau com o qual o todo é uma parte de uma de suas partes, ou seja, uma relação impossível na lógica clássica.

Essa relação viola a intuição que um diagrama de Venn nos dá sobre a não-ambiguidade da inclusão de conjuntos. Somente no ponto médio de I^n a união $A \cup A^c$ pode estar totalmente contida na intersecção $A \cap A^c$. Nos vértices do hipercubo essa possibilidade vale zero. Isso pode ser comparado com a relação de frequência relativa $S(X, A) = \frac{n_A}{n}$, onde o conjunto clássico A contém em algum grau seu superconjunto clássico X .

A Figura 5.7 ilustra o Teorema da Entropia-Inclusão. Mostra que d^* , a menor distância da união $A \cup A^c$ ao hiperretângulo que define o conjunto potência fuzzy da intersecção $A \cap A^c$ é igual à distância $d(A \cup A^c, A \cap A^c) = d(A, A^c)$ e igual a diferença dos vetores grandeza: $d^* = M(A \cup A^c) - M(A \cap A^c)$.

5.3.1 Principal Consequência

Qual a principal consequência do Teorema da Entropia-Inclusão? Suponha que algo tal como “fuzziness” não exista, mas que “fuzziness” seja simplesmente um tipo

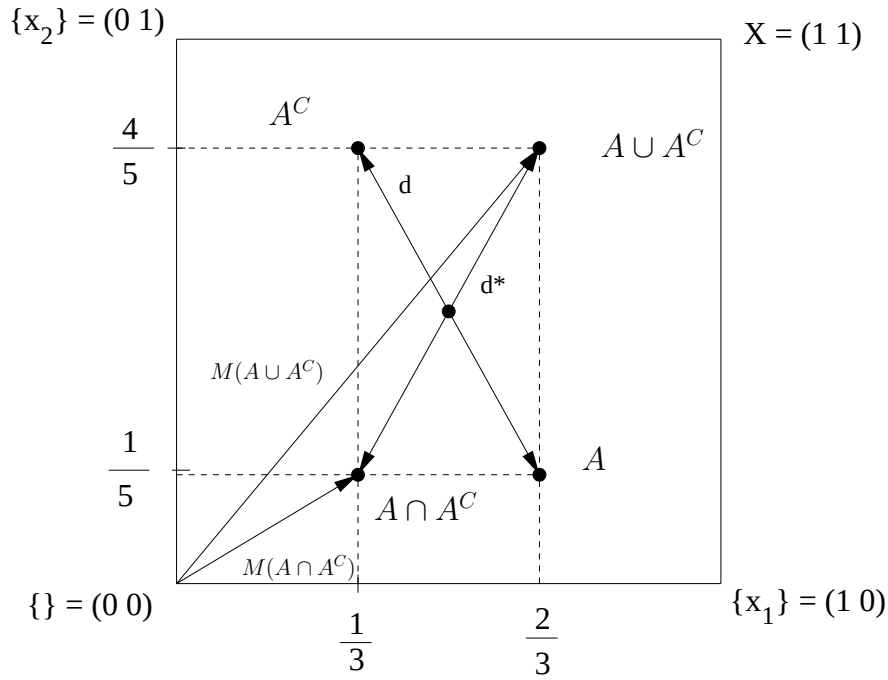


Figura 5.7: Ilustração do Teorema da Entropia-Inclusão em duas dimensões

de probabilidade disfarçada e que a probabilidade seja adequada para medir todos os problemas envolvendo incertezas. A que conclusão se chega? Se este fosse o caso, existiria uma medida de probabilidade P , tal que $P = E$, que mediria a “fuzziness” de um subconjunto fuzzy A qualquer, de X . Uma vez que $P(X) = 1$ existe pelo menos um conjunto A onde $P(A) > 0$ e como, por hipótese, $P(A) = E(A)$ tem-se que

$$0 < P(A) = E(A).$$

Mas, se $A = X$ obtém-se

$$A \cap A^c = \emptyset \quad e \quad A \cup A^c = X.$$

Então, aplicando o teorema da entropia-inclusão chega-se a

$$0 < P(A) = E(A) = S(A \cup A^c, A \cap A^c) = S(X, \emptyset).$$

No entanto, X somente pode ser um subconjunto com um grau de inclusão diferente de zero do conjunto vazio se ele mesmo for o conjunto vazio e, portanto, somente se $A = X = \emptyset$. Mas se isto ocorre, o evento certo será impossível, pois $P(X) = P(\emptyset) = 0$. Ou, então o evento impossível é certo com $P(X) = P(\emptyset) = 1$. Isso resulta em uma contradição dupla, que é impossível de ser normalizada. Então, não existe uma medida de probabilidade P que meça “fuzziness”.

Assim sendo, existem problemas envolvendo incertezas, que não podem ser medidos pela probabilidade. Portanto, “fuzziness” existe.

6 Conclusão

Nesse trabalho foi feita uma descrição das propriedades básicas dos Conjuntos Fuzzy e discutiu-se a representação geométrica dos Conjuntos Fuzzy finitos que permite comparar os conceitos de “fuzziness” e probabilidade.

“Fuzziness” existe. Uma frase simples, mas de amplo significado para a lógica e para todo o pensamento matemático. Uma frase simples, mas que esclarece paradoxos levantados na lógica bivalente, amplia nosso entendimento do universo e talvez permita, no futuro, programar computadores para “pensar” de maneira semelhante aos humanos. O correto entendimento de “fuzziness” mostra que a incerteza pode ser medida, que é possível usar a quantidade de informação para compreender o quanto um evento ocorre e esclarece a diferença entre SE ele ocorre e o QUANTO ocorre. Da mesma forma que a negação do “postulado das paralelas” conduziu a novas geometrias, negar as leis clássicas da lógica bivalente, a saber, $A \cap A^c = \emptyset$ e $A \cup A^c = X$, tem conduzido a novas lógicas ampliando assim as possibilidades de pesquisa na área da informação.

Referências

- [1] PEDRYCZ, W. *Fuzzy Control and Fuzzy Systems*. 2nd, extend. ed. West Sussex: John Wiley & Sons, 1992.
- [2] LUKASIEWICZ, J. O logice trójwartosciowej. *Ruch filozoficzny*, v. 5, p. 170–171, 1920.
- [3] LUKASIEWICZ, J. *Selected works (Studies in logic and the foundations of mathematics)*. 1st. ed. Amsterdam: North-Holland Pub. Co, 1970.
- [4] ZADEH, L. A. Fuzzy Sets. *Information and Control*, v. 8, p. 338–353, 1965.
- [5] KRUSE, R.; GEBHARDT, J.; KLAWONN, F. *Foundations of Fuzzy Systems*. 1st. ed. New York: John Wiley & Sons, 1994.
- [6] SHANNON, C. A mathematical theory of communication. *Bell System Technical Journal*, v. 27, p. 379–423, 1948.
- [7] NGUYEN, H. T.; WALKER, E. A. *A First Course in Fuzzy Logic*. 2nd. ed. New York: Chapman & Hall/CRC, 2000.
- [8] BARONETT, S. *Lógica Uma introdução voltada para as ciências*. 1a. ed. Porto Alegre: IMECC, 2009.
- [9] MACHADO, N. J.; CUNHA, M. O. *Lógica e linguagem cotidiana, verdade, coerência, comunicação, argumentação*. 1a. ed. Belo Horizonte: Autêntica, 2005.
- [10] ENDERTON, H. E. *A Mathematical Introduction to Logic*. 1a. ed. New York: Academic Press, 1972.
- [11] BARROS, L. C.; BASSANEZI, R. C. *Tópicos de Lógica Fuzzy e Biomatemática*. 1a. ed. Campinas: IMECC, 2006.
- [12] JAYNES, E. T. Where do we stand on maximum entropy? *The Maximum Entropy Formalism*, v. 1, p. 15–118, 1979.
- [13] JAYNES, E. T. Probability, frequency, and reasonable expectations. *American Journal of Physics*, v. 14, p. 1–13, 1946.

-
- [14] QUINE, W. V. What price bivalence? *The Journal of Philosophy*, v. 2, p. 90–95, 1981.
- [15] HUME, D. *An Inquiry Concerning Human Understanding*. 1st. ed. New Jersey: Prentice-Hall, Inc., 1748.
- [16] KOSKO, B. *Neural Networks and Fuzzy Systems - A Dynamical Systems Approach to Machine Intelligence*. 1st. ed. New Jersey: Prentice-Hall, Inc., 1992.
- [17] ZADEH, L. A. A computacional approach to fuzzy quantifiers in natural languages. *Computers and Mathematics*, v. 9, p. 149–184, 1983.
- [18] KOSKO, B. *Foundations of Fuzzy Estimation Theory*. Tese (Doutorado) — University of California at Irvine, 1987.
- [19] CIGNOLI, R. L. O.; D’OTTAVIANO, I. M. L.; MUNDICI, D. *Álgebra das Lógicas de Lukasiewics*. 1a. ed. Campinas: IMECC, 1995.
- [20] NEGOITA, C. V.; RALESCU, D. A. *Applications of Fuzzy Sets to Systems Analysis*. 1st. ed. New York: John Wiley & Sons, 1975.