

Vinícius Oliari Couto Dias

Wavelets e uma Aplicação na Espectrometria

Brasil

2017

Vinícius Oliari Couto Dias

Wavelets e uma Aplicação na Espectrometria

Dissertação apresentada ao Departamento de Matemática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Mestrado em Matemática.

Universidade Federal do Espírito Santo – UFES

Departamento de Matemática

Mestrado em Matemática

Orientador: Etereldes Gonçalves

Brasil

2017

Vinícius Oliari Couto Dias

Wavelets e uma Aplicação na Espectrometria/ Vinícius Oliari Couto Dias. –
Brasil, 2017-

120 p. : il. (algumas color.) ; 30 cm.

Orientador: Etereldes Gonçalves

Dissertação – Universidade Federal do Espírito Santo – UFES

Departamento de Matemática

Mestrado em Matemática, 2017.

1. Wavelets 2. Fuzzy I. Etereldes Gonçalves II. Universidade Federal do
Espírito Santo. III. Departamento de Matemática. IV. Wavelets e uma Aplicação
na Espectrometria

Vinícius Oliari Couto Dias

Wavelets e uma Aplicação na Espectrometria

Dissertação apresentada ao Departamento de Matemática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Mestrado em Matemática.

Brasil, 03 de fevereiro de 2017:

Etereldes Gonçalves
Orientador

Convidado 1

Convidado 2

Brasil
2017

Resumo

O uso das Wavelets tem crescido bastante nos últimos anos. Suas aplicações contemplam diversas áreas, desde a matemática pura até processamento de imagens, por exemplo. A principal propriedade das Wavelets é a representação de funções num espaço de escalas, o que permite a observação da função de maneira eficaz em um intervalo reduzido.

Nesta dissertação apresentamos um desenvolvimento matemático da construção de uma família de Wavelets, além de sua aplicação em um problema da espectrometria para a obtenção de propriedades do petróleo. Para tal, fazemos uso de outras ferramentas, como teoria Fuzzy, Algoritmo Genético e clusterização, obtendo um método de predição das propriedades. Por fim, apresentamos os resultados do método desenvolvido por meio dos espectros de amostras de petróleo, fornecidos pelo LabPetro-UFES.

Abstract

Wavelets have played an important role on recent researches. Its applications go from pure math to image processing and many others. Their main property is the representation of functions in a scale space, which allows an efficient observation of the function in an small interval.

In this thesis, a mathematical development of the construction of a Wavelet family is presented, as well as an application on spectrometry to obtain petroleum properties. To do so, many others tools are used, such as Fuzzy theory, Genetic Algorithm and clusterization, resulting in a method of properties forecasting. Finally, the method results are presented, using the spectrum of samples provided by LabPetro-UFES.

Lista de ilustrações

Figura 1	– Condições sobre as funções Scaling φ para obtermos uma análise multiresolucional	40
Figura 2	– Condições sobre os coeficientes a_0 e a_1 para que o conjunto $\{2^{j/2}\psi_1(2^j x - \gamma)\}_{\gamma, j \in \mathbb{Z}}$ seja base ortonormal de $L^2(\mathbb{R})$	40
Figura 3	– Condições sobre os coeficientes a_0 e a_1 para garantir a ortonormalidade das funções ψ_k	41
Figura 4	– Condições sobre os coeficientes A_0 e A_1 para obtenção de uma análise multiresolucional	61
Figura 5	– Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a quinta potência da identidade trigonométrica	65
Figura 6	– Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a segunda potência da identidade trigonométrica	65
Figura 7	– Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a terceira potência da identidade trigonométrica	66
Figura 8	– Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) Daubechies 4	66
Figura 9	– Gráfico do Espectro da Amostra 1	67
Figura 10	– a) - Gráfico do Espectro da Amostra 1. b) - Gráfico do Espectro da Amostra 47. c) - Gráfico do Espectro da Amostra 57.	68
Figura 11	– Regiões do Espectro utilizadas para obtenção das Propriedades	68
Figura 12	– Fluxograma do Método	69
Figura 13	– Exemplo bidimensional do método Duplex (até à quarta etapa apenas)	70
Figura 14	– Wavelet Daubechies 3	71
Figura 15	– Wavelet Daubechies 4	71
Figura 16	– a) Região selecionada no espectro. b) Divisão da região selecionada em 5 partes iguais (subregiões), devido à escala 5 escolhida.	72
Figura 17	– Possíveis pontos iniciais e finais da primeira região	73
Figura 18	– Possíveis pontos iniciais e finais da segunda região	73
Figura 19	– Wavelets ajustadas ao tamanho de cada subregião para uma escala de 5 na primeira região com ponto inicial de 900 e final de 1400	74
Figura 20	– Representação Genérica de um Sistema Fuzzy	75
Figura 21	– Exemplo de uma função de pertinência definida no intervalo $[0, 10]$	76
Figura 22	– Pertinência da proximidade com o número 5 dos números de 1 a 10	77
Figura 23	– Função de pertinência de um conjunto não-convexo	78
Figura 24	– Função de pertinência de um conjunto convexo	78
Figura 25	– Funções de pertinência de A e B	80

Figura 26	– Função de pertinência de $A \cap B$	81
Figura 27	– Função de pertinência de $A \cup B$	81
Figura 28	– Função de pertinência de A^c	82
Figura 29	– Densidade de probabilidade D dos valores que uma resistência produzida pode assumir	83
Figura 30	– Gráfico da fuzzyficação do valor 4 em um <i>singleton</i>	84
Figura 31	– Conjuntos Fuzzy respectivos à entrada n	87
Figura 32	– Conjuntos Fuzzy respectivos à entrada T_a	88
Figura 33	– Conjuntos Fuzzy respectivos à saída P	88
Figura 34	– Função de Pertinência $\mu_{\tilde{B}_k}$ via implicação Mandani em comparação com μ_{B_k}	91
Figura 35	– Função de Pertinência $\mu_{\tilde{B}_k}$ via implicação Gödel em comparação com μ_{B_k}	92
Figura 36	– Função de agregação pelo máximo	93
Figura 37	– Função de agregação pelo mínimo	93
Figura 38	– Obtenção da coordenada v do centro de área da função de pertinência de \tilde{B}_T	95
Figura 39	– Representação do Primeiro e Último Ponto de Máximo da função de pertinência de \tilde{B}_T	96
Figura 40	– Ilustração do potencial P gerada pelo histograma de um conjunto de pontos bidimensionais. Quanto maior o valor na barra (referente a uma cor), maior o potencial correspondente do ponto.	97
Figura 41	– Funções de pertinência relativas a cada entrada (para duas regras apenas)	109
Figura 42	– Regra Fuzzy para um sistema com duas entradas	109
Figura 43	– Defuzzyficação das funções de pertinência de saída resultantes	109
Figura 44	– Densidade Medida Vs. Densidade Estimada pelo Sistema Fuzzy para a Densidade	112
Figura 45	– Erro entre o Valor Estimado e o Medido para a Densidade	112
Figura 46	– TIAC Medida Vs. Densidade Estimada pelo Sistema Fuzzy para a TIAC	114
Figura 47	– Erro entre o Valor Estimado e o Medido para a TIAC	114

Lista de tabelas

Tabela 1 – Densidade e TIAC da Amostra 1	67
Tabela 2 – Densidade e TIAC das Amostras 1, 47 e 57	68
Tabela 3 – Valores da Função de Pertinência μ_{p5}	77
Tabela 4 – Pertinência em uma Relação entre Dois Conjuntos de Três Pessoas	82
Tabela 5 – Pertinência do conjunto PZ	86
Tabela 6 – Consequência de Saída (Potência PX) para as Combinações de Número de Pessoas (NX) e Temperatura (TX)	89
Tabela 7 – Notação dos Parâmetros Tamanho e Número de Pessoas referentes a cada Sala	103
Tabela 8 – Valores de δ_k , relativos à Probabilidade de Mutação de cada Parâmetro	105
Tabela 9 – Características utilizadas para Geração do Sistema Fuzzy para a Propriedade Densidade	111
Tabela 10 – Parâmetros do Melhor Indivíduo obtido pelo Algoritmo Genético (A.G.) para a Propriedade Densidade	111
Tabela 11 – Valores de Erro entre Valor Estimado e Medido para a Densidade	112
Tabela 12 – Características utilizadas para geração do sistema Fuzzy para a propriedade TIAC	113
Tabela 13 – Parâmetros do melhor indivíduo obtido pelo Algoritmo Genético (A.G.) para a propriedade TIAC	113
Tabela 14 – Valores de Erro entre Valor Estimado e Medido para a TIAC	114
Tabela 15 – Comparação de resultados API	115
Tabela 16 – Comparação de resultados TIAC	115

Sumário

1	Introdução	15
1.1	História e aplicações	15
1.2	Propriedades principais	15
1.3	Comparação entre Wavelet e Fourier	16
1.4	Ortogonalidade das funções bases das transformadas de Fourier e Haar	17
1.5	Descrição da Aplicação	21
1.6	Organização do Texto	21
2	Wavelets	23
2.1	Wavelet Haar	23
2.1.1	Função Scaling Haar	23
2.1.2	Espaços gerados a partir da Wavelet Haar e sua Scaling	24
2.1.3	Wavelet Haar e a base para $L^2(\mathbb{R})$	28
2.2	Análise Multiresolucional	30
2.2.1	Condições necessárias para uma análise multiresolucional	33
2.3	Wavelets	37
2.3.1	Propriedades das Wavelets	38
2.4	Transformada de Fourier e Wavelets	41
2.4.1	Propriedades no domínio da Transformada de Fourier	42
2.4.2	Ortonormalidade do conjunto $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$	58
2.5	A construção de uma Wavelet	61
3	Método de Obtenção de Propriedades do Petróleo	67
3.1	Ordenação das amostras	69
3.2	Escolha da Wavelet	70
3.3	Escolha da sscala da Wavelet	71
3.4	Escolha dos pontos de início e fim das regiões	72
3.5	Produto interno	73
3.6	Modelagem Fuzzy	74
3.6.1	Teoria Fuzzy	75
3.6.2	Funcionamento de um sistema Fuzzy	82
3.6.3	Fuzzyficação	83
3.6.4	Afirmação Fuzzy	84
3.6.5	Regras	86
3.6.6	Operação Fuzzy implicação	89
3.6.7	Agregação	91

3.6.8	Defuzzyficação	94
3.7	Clusterização via Método da Montanha	96
3.7.1	Obtenção dos núcleos	96
3.7.2	Critério de parada	98
3.7.3	Determinação do modelo Fuzzy	99
3.7.4	Otimização do Modelo Fuzzy	100
3.8	Algoritmo Genético	102
3.8.1	Indivíduo	102
3.8.2	População	103
3.8.3	Geração	104
3.8.4	Operações em indivíduos	104
3.8.5	Mutação	104
3.8.6	<i>Cross-Over</i>	105
3.8.7	Eliminação	106
3.8.8	Função <i>fitness</i>	107
3.8.9	Elitismo	107
3.8.10	Critério de parada	108
3.9	Sistema Fuzzy utilizado e desempenho	108
3.10	Resultados	111
4	Conclusão	117
	Referências	119

1 Introdução

1.1 História e aplicações

As funções Wavelets são consideradas um objeto de estudo recente e já possuem diversas aplicações em vários campos de pesquisa. A primeira Wavelet, a Wavelet Haar, foi mencionada apenas em 1909, por A. Haar ([HAAR, 1910](#)), e o próprio nome "Wavelet" começou a ser utilizado apenas na década de 80 ([MORLET et al., 1982](#)). Desde então, o uso das Wavelets cresceu em grande proporção ([DAUBECHIES, 1992](#)).

Suas aplicações contemplam várias áreas além da matemática pura ([DAUBECHIES, 1992](#)). Por exemplo, na física, há trabalhos sobre estados coerentes e grupo de renormalização. Na química, foi introduzida na década de 90 e hoje possui um papel importante na quimiometria e espectrofotometria ([CHAU et al., 2004](#)). No campo da engenharia, encontram-se trabalhos baseados em Wavelets sobre codificação de sub-banda, [Orthogonal frequency-division multiplexing \(OFDM\)](#), análise de sinais, compressão de dados, remoção de ruído, dentre outros. Computacionalmente, trabalha-se com a transformada Wavelet discreta, que pode ser vista como uma representação da Wavelet por apenas um número finito de pontos, sendo que sua teoria também é bastante desenvolvida ultimamente.

1.2 Propriedades principais

A principal vantagem da transformada Wavelet é que ela representa um sinal no domínio da frequência e do tempo, enquanto a transformada de Fourier trabalha apenas na frequência. Tal fato permite avaliar variações ocorridas em sinais em um intervalo de tempo específico, o que seria similar ao processo de janelamento utilizado com a transformada de Fourier. Isso ocorre porque a transformada Wavelet trabalha com um espaço de escalas, que divide o sinal em grupos de intervalos transladados, nos quais cada grupo possui um comprimento de intervalo específico. Dessa maneira, grupos com um intervalo de comprimento maior darão uma visão global do sinal e os de comprimento menor uma visão mais detalhada.

No que diz respeito à transformada Wavelet discreta, uma das principais propriedades é que ela guarda a maior parte da energia do sinal em poucos coeficientes, o que permite que o sinal seja recuperado apenas por um reduzido número deles. Esse fato é de grande importância na compressão de dados, na qual se escolhe uma determinada Wavelet que armazene a maior quantidade possível de energia no menor número de coeficientes e

que permita com eles a recuperação do sinal sem que sofra grandes deformações.

1.3 Comparação entre Wavelet e Fourier

Uma transformada Wavelet, analiticamente falando, muito se assemelha com a transformada de Fourier. Ambas permitem que um sinal seja representado por uma série de coeficientes que multiplicam suas funções base. Como exemplo, uma função f definida no intervalo $[0, 1]$ que seja quadrado integrável (REED; SIMON, 1980) é representada da seguinte maneira em uma série de Fourier:

$$f(x) = b_0 + \sum_{k=1}^{\infty} (b_k \cos(2\pi kx) + a_k \sin(2\pi kx))$$

Por outro lado, considerando a função Wavelet Haar, definida por:

$$\psi(x) = \begin{cases} 1 & \text{se } 0 \leq x < \frac{1}{2} \\ -1 & \text{se } \frac{1}{2} \leq x < 1 \\ 0 & \text{caso contrário} \end{cases} \quad (1.1)$$

a função f em uma série de Haar é dada por:

$$f(x) = c_0 + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{kj} \psi(2^j x - k) \quad (1.2)$$

Em ambas as séries, temos um termo independente e uma soma de funções com determinados pesos: senos e cossenos na série de Fourier e a própria Wavelet Haar na série de Haar. Além disso, ambas as séries são expansões em termos de funções ortogonais em $L^2[0, 1]$ (a prova consta na próxima seção). Tal fato facilita na obtenção dos coeficientes da expansão em séries:

$$b_k = \int_0^1 \overline{f(x)} \cos(2\pi kx) dx$$

$$a_k = \int_0^1 \overline{f(x)} \sin(2\pi kx) dx$$

$$c_{kj} = \int_0^1 \overline{f(x)} \psi(2^j x - k) dx$$

$$b_0 = c_0 = \int_0^1 \overline{f(x)} dx$$

onde a barra representa o conjugado complexo.

A Wavelet Haar possui uma importante propriedade que os senos e cossenos não possuem, um suporte compacto.

Definição 1.3.1. Suporte: O suporte de uma função $f : X \mapsto V$ é o conjunto dado por:

$$\text{supp}(f) = \overline{\{x \in X; f(x) \neq 0\}}$$

onde a barra significa o fecho do conjunto.

Um conjunto compacto na reta real significa que o conjunto é limitado e fechado. Dizemos que o suporte de f é compacto quando o conjunto $\text{supp}(f)$ é compacto.

A vantagem advinda de tal propriedade pode ser resumida no seguinte fato: para descrever a função em um intervalo $[a, b]$, precisamos de todos os coeficientes da série de Fourier, enquanto que para a Wavelet Haar apenas aqueles $I_{k,j} = [2^{-j}k, 2^{-j}(k+1)]$ que intersectam $[a, b]$.

Na equação (1.2) notamos também a propriedade de escala (comentada na seção anterior) pois, considerando apenas a soma parcial da expansão em Haar (ou seja, para $0 \leq j \leq N$), temos a representação de f considerando detalhes (Wavelets com uma determinada escala e coeficiente) de magnitude 2^{-N} ou maiores.

1.4 Ortogonalidade das funções bases das transformadas de Fourier e Haar

O produto interno entre duas funções em $L^2[a, b]$ será definido por:

$$\begin{aligned} \langle \cdot, \cdot \rangle : L^2[a, b] \times L^2[a, b] &\mapsto \mathbb{C} \\ \langle f, g \rangle &= \int_a^b \overline{f(x)}g(x) dx \end{aligned}$$

onde a barra significa o conjugado complexo. A ortogonalidade entre duas funções f, g ocorre quando $\langle f, g \rangle = 0$. Com base no conceito de produto interno definido, será então provada a ortogonalidade das funções bases das transformadas de Fourier e Haar.

- Ortogonalidade da Wavelet Haar

Sejam ψ_1, ψ_2 dadas por

$$\begin{aligned} \psi_1 &= \psi(2^{j_1}x - k_1) \\ \psi_2 &= \psi(2^{j_2}x - k_2) \end{aligned}$$

onde ψ é a função Haar definida na equação (1.1) e $k_1, k_2, j_1, j_2 \in \mathbb{Z}$. Para provar a ortogonalidade, basta verificar que $\langle \psi_1, \psi_2 \rangle = 0$ no caso em que $j_1 \neq j_2$ ou $k_1 \neq k_2$. Para isso, podemos partir da definição de produto interno:

$$\langle \psi_1, \psi_2 \rangle = \int_{-\infty}^{\infty} \overline{\psi(2^{j_1}x - k_1)} \psi(2^{j_2}x - k_2) dx$$

onde $k_1, k_2, j_1, j_2 \in \mathbb{Z}$. Temos que:

$$\begin{aligned} \langle \psi_1, \psi_2 \rangle &= \int_{\frac{k_1}{2^{j_1}}}^{\frac{k_1+1}{2^{j_1}}} \psi(2^{j_2}x - k_2) dx - \int_{\frac{k_1+\frac{1}{2}}{2^{j_1}}}^{\frac{k_1+1}{2^{j_1}}} \psi(2^{j_2}x - k_2) dx = \\ &= \int_{U_1} 1 dx - \int_{U_2} 1 dx - \int_{U_3} 1 dx + \int_{U_4} 1 dx \end{aligned} \quad (1.3)$$

onde temos:

$$\begin{aligned} U_1 &= \left(\frac{k_1}{2^{j_1}}, \frac{k_1 + \frac{1}{2}}{2^{j_1}} \right) \cap \left(\frac{k_2}{2^{j_2}}, \frac{k_2 + \frac{1}{2}}{2^{j_2}} \right) \\ U_2 &= \left(\frac{k_1}{2^{j_1}}, \frac{k_1 + \frac{1}{2}}{2^{j_1}} \right) \cap \left(\frac{k_2 + \frac{1}{2}}{2^{j_2}}, \frac{k_2 + 1}{2^{j_2}} \right) \\ U_3 &= \left(\frac{k_1 + \frac{1}{2}}{2^{j_1}}, \frac{k_1 + 1}{2^{j_1}} \right) \cap \left(\frac{k_2}{2^{j_2}}, \frac{k_2 + \frac{1}{2}}{2^{j_2}} \right) \\ U_4 &= \left(\frac{k_1 + \frac{1}{2}}{2^{j_1}}, \frac{k_1 + 1}{2^{j_1}} \right) \cap \left(\frac{k_2 + \frac{1}{2}}{2^{j_2}}, \frac{k_2 + 1}{2^{j_2}} \right) \end{aligned}$$

Se $j_1 = j_2$ e $k_1 \neq k_2$, claramente U_1, U_2, U_3 e $U_4 = \emptyset$ e logo $\langle \psi_1, \psi_2 \rangle = 0$. Se $j_1 \neq j_2$, considerando sem perda de generalidade $j_1 > j_2$, para que U_1, U_2, U_3 e U_4 não sejam todos vazios, precisamos que, para $d \in \left\{0, \frac{1}{2}, 1\right\}$:

$$\frac{k_2}{2^{j_2}} \leq \frac{k_1 + d}{2^{j_1}} \leq \frac{k_2 + \frac{1}{2}}{2^{j_2}} \quad (1.4)$$

ou

$$\frac{k_2 + \frac{1}{2}}{2^{j_2}} \leq \frac{k_1 + d}{2^{j_1}} \leq \frac{k_2 + 1}{2^{j_2}} \quad (1.5)$$

As inequações são verdadeiras pois:

$$\begin{aligned} \frac{k_1}{2^{j_1}} < \frac{k_2}{2^{j_2}} &\implies k_1 < 2^{j_1-j_2} k_2 \\ &\implies k_1 + 1 \leq 2^{j_1-j_2} k_2 \\ &\implies \frac{k_1 + 1}{2^{j_1}} \leq \frac{k_2}{2^{j_2}} \end{aligned} \quad (1.6)$$

e

$$\begin{aligned}
\frac{k_1 + 1}{2^{j_1}} > \frac{k_2 + 1}{2^{j_2}} &\implies k_1 + 1 > 2^{j_1 - j_2}(k_2 + 1) \\
&\implies k_1 \geq 2^{j_1 - j_2}(k_2 + 1) \\
&\implies \frac{k_1}{2^{j_1}} \geq \frac{k_2 + 1}{2^{j_2}}
\end{aligned} \tag{1.7}$$

As equações (1.6) e (1.7) mostram que para $\frac{k_2}{2^{j_2}} \leq \frac{k_1 + d}{2^{j_1}}$ e $\frac{k_1 + d}{2^{j_1}} \leq \frac{k_2 + 1}{2^{j_2}}$ não satisfeitas temos U_1, U_2, U_3 e $U_4 = \emptyset$. Basta agora mostrar que, com elas satisfeitas temos $\frac{k_1 + d}{2^{j_1}} \leq \frac{k_2 + \frac{1}{2}}{2^{j_2}}$ ou $\frac{k_2 + \frac{1}{2}}{2^{j_2}} \leq \frac{k_1 + d}{2^{j_1}}$. Para tal, basta mostrar que $\frac{k_2 + \frac{1}{2}}{2^{j_2}} \notin \left(\frac{k_1}{2^{j_1}}, \frac{k_1 + 1}{2^{j_1}} \right)$. Suponha por absurdo que pertença, então:

$$\frac{k_1}{2^{j_1}} < \frac{k_2 + \frac{1}{2}}{2^{j_2}} < \frac{k_1 + 1}{2^{j_1}} \tag{1.8}$$

Mas

$$\begin{aligned}
\frac{k_1}{2^{j_1}} < \frac{k_2 + \frac{1}{2}}{2^{j_2}} &\implies k_1 < 2^{j_1 - j_2} \left(k_2 + \frac{1}{2} \right) \\
&\implies k_1 + 1 \leq 2^{j_1 - j_2} \left(k_2 + \frac{1}{2} \right) \\
&\implies \frac{k_1 + 1}{2^{j_1}} \leq \frac{k_2 + \frac{1}{2}}{2^{j_2}}
\end{aligned}$$

o que é absurdo por (1.8).

Agora, podemos assumir que ou (1.4) ou (1.5) vale. Sendo assim, considere que (1.4) vale (o caso para (1.5) é análogo). Desse modo temos que

$$U_2, U_4 = \emptyset \tag{1.9}$$

$$U_1 = \left(\frac{k_1}{2^{j_1}}, \frac{k_1 + \frac{1}{2}}{2^{j_1}} \right)$$

$$U_3 = \left(\frac{k_1 + \frac{1}{2}}{2^{j_1}}, \frac{k_1 + 1}{2^{j_1}} \right)$$

Como U_1 e U_3 possuem o mesmo comprimento, usando a equação (1.9), temos por (1.3) que:

$$\langle \psi_1, \psi_2 \rangle = \int_{U_1} 1 \, dx - \int_{U_2} 1 \, dx - \int_{U_3} 1 \, dx + \int_{U_4} 1 \, dx = 0$$

No caso em que $j_1 = j_2 = j$ e $k_1 = k_2 = k$, ou seja, $\psi_1 = \psi_2 = \psi$, temos que:

$$\begin{aligned}
\langle \psi, \psi \rangle &= \int_0^1 \overline{\psi(2^j x - k)} \psi(2^j x - k) \, dx \\
&= \int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} \psi(2^j x - k) \, dx - \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} \psi(2^j x - k) \, dx \\
&= \left(\frac{k + \frac{1}{2}}{2^j} - \frac{k}{2^j} \right) + \left(\frac{k + 1}{2^j} - \frac{k + \frac{1}{2}}{2^j} \right) \\
&= \frac{1}{2^j}
\end{aligned}$$

Isso mostra que a Wavelet Haar não é normalizada. Para tal, devemos multiplica-la por um fator $2^{\frac{j}{2}}$.

- Ortogonalidade da Base de Fourier

A ortogonalidade dos senos e cossenos, que são as bases da série de Fourier, segue diretamente da definição de produto interno. Sejam $k_1, k_2 > 0$ números inteiros, então temos que:

$$\begin{aligned}
\langle \cos(2\pi k_1 \cdot), \cos(2\pi k_2 \cdot) \rangle &= \int_0^1 \overline{\cos(2\pi k_1 x)} \cos(2\pi k_2 x) \, dx \\
&= \frac{1}{2} \int_0^1 (\cos(2\pi(k_1 + k_2)x) + \cos(2\pi(k_1 - k_2)x)) \, dx \\
&= \begin{cases} 0, & \text{se } k_1 \neq k_2 \\ \frac{1}{2}, & \text{se } k_1 = k_2 \end{cases}
\end{aligned}$$

$$\begin{aligned}
\langle \sen(2\pi k_1 \cdot), \sen(2\pi k_2 \cdot) \rangle &= \int_0^1 \overline{\sen(2\pi k_1 x)} \sen(2\pi k_2 x) \, dx \\
&= \frac{1}{2} \int_0^1 (\cos(2\pi(k_1 - k_2)x) - \cos(2\pi(k_1 + k_2)x)) \, dx \\
&= \begin{cases} 0, & \text{se } k_1 \neq k_2 \\ \frac{1}{2}, & \text{se } k_1 = k_2 \end{cases}
\end{aligned}$$

$$\begin{aligned}
\langle \sen(2\pi k_1 \cdot), \cos(2\pi k_2 \cdot) \rangle &= \int_0^1 \overline{\sen(2\pi k_1 x)} \cos(2\pi k_2 x) \, dx \\
&= \frac{1}{2} \int_0^1 (\sen(2\pi(k_1 + k_2)x) + \sen(2\pi(k_1 - k_2)x)) \, dx \\
&= 0 \quad \forall k_1, k_2 \in \mathbb{N}
\end{aligned}$$

Assim concluímos que as bases da série de Fourier são ortogonais.

1.5 Descrição da Aplicação

Neste trabalho foi realizada uma aplicação da Teoria de Wavelets no campo da Espectrometria. Ela se baseia na predição de propriedades de amostras de petróleo a partir do espectro refletido pela emissão de infravermelho nas amostras, feita pelo espectômetro, um aparelho disponível nos laboratórios do LabPetro - UFES.

Atualmente são utilizadas técnicas que fazem uso de ferramentas mais estatísticas, como regressões multivariável e outros. Porém, nesse trabalho quisemos gerar um método mais flexível (com várias partes que podem ser mudadas e adaptadas) e diferente dos já feitos, de modo a obter uma nova linha de aproximação do problema.

O uso da Wavelet no método se baseia na extração de variáveis do espectro, as quais serão utilizadas no processo de predição. Tal procedimento foi proposto devido à propriedade de escala das funções Wavelet, que será discutida nesta dissertação.

Devido à qualidade dos resultados gerados pelo método, um artigo foi desenvolvido para ser submetido à revista *Chemometrics and Intelligent Laboratory Systems*, explicando as etapas do método e sua aplicação nas amostras de petróleo fornecidas.

1.6 Organização do Texto

O segundo capítulo desta dissertação descreve um método de criar Wavelets, bem como as principais propriedades do espaço e comportamento de tais funções. A teoria contida nesse capítulo foi baseada no desenvolvimento feito em (STRICHARTZ, 1993). A diferença entre os dois está no detalhamento adicional feito nos teoremas e provas contidos em tal referência.

O terceiro capítulo descreve os principais métodos utilizados para a elaboração do preditor, culminando com a apresentação e resultados de sua utilização com espectros de amostras de petróleo fornecidas pelo LabPetro.

2 Wavelets

Neste capítulo serão demonstrados os principais teoremas e propriedades das Wavelets. Além disso, no final do capítulo, será mostrado como criar Wavelets com as propriedades especificadas. Para tal, começaremos com um tipo específico de Wavelet, a Haar, para depois generalizar para uma Wavelet qualquer que atenda às propriedades impostas.

2.1 Wavelet Haar

Esta seção trata da Wavelet Haar, um dos tipos mais simples de Wavelet. A teoria envolvida em tal função é de extrema importância para compreender como as Wavelets funcionam de maneira geral.

2.1.1 Função Scaling Haar

A função Scaling é uma função que está relacionada com uma dada Wavelet, e conhecendo funções Scaling podemos criar funções Wavelet. A função Scaling associada à Wavelet Haar é a função característica do intervalo $[0, 1]$, definida para $x \in \mathbb{R}$ por:

$$\varphi(x) = \begin{cases} 1 & \text{se } 0 \leq x < 1 \\ 0 & \text{caso contrário} \end{cases} \quad (2.1)$$

Tendo em vista as principais características de uma Wavelet, é necessário citar duas propriedades importantes de φ :

1. As translações de φ por inteiros ($\{\varphi(x-k), k \in \mathbb{Z}\}$) formam um conjunto de funções ortonormais em $L^2(\mathbb{R})$.

Demonstração. De fato temos, para $k_1, k_2 \in \mathbb{Z}$, $k_1 \neq k_2$:

$$\begin{aligned} \langle \varphi(\bullet - k_1), \varphi(\bullet - k_2) \rangle &= \int_{-\infty}^{+\infty} \overline{\varphi(x - k_1)} \varphi(x - k_2) dx \\ &= \int_{k_1}^{k_1+1} \varphi(x - k_2) dx \\ &= \int_{k_1}^{k_1+1} 0 dx \\ &= 0 \end{aligned}$$

E para $k_1 = k_2 = k$:

$$\begin{aligned} \langle \varphi(\bullet - k), \varphi(\bullet - k) \rangle &= \int_{-\infty}^{+\infty} \overline{\varphi(x - k)} \varphi(x - k) \, dx \\ &= \int_k^{k+1} \varphi(x - k) \, dx \\ &= \int_{k_1}^{k_1+1} 1 \, dx \\ &= 1 \end{aligned}$$

■

2. φ é auto-similar. Isso significa nesse contexto que ela pode ser expressa por uma combinação linear dela mesma em diferentes escalas:

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1) \quad (2.2)$$

2.1.2 Espaços gerados a partir da Wavelet Haar e sua Scaling

No que diz respeito a séries e transformadas, a questão fundamental é saber quais funções são possíveis de serem representadas/realizadas por elas. Com base nessa questão, o intuito desta subseção é determinar se é possível obter uma base para $L^2(\mathbb{R})$ (o espaço das funções quadrado integráveis, ou seja, as funções f tais que $\int |f(x)|^2 \, dx < \infty$ (REED; SIMON, 1980)) a partir da Wavelet Haar e a função φ .

Uma vez definida a função Scaling, o espaço natural gerado por ela é:

$$V_0 = \left\{ \sum_{k=-\infty}^{\infty} a_k \varphi(x - k), \sum |a_k|^2 < \infty \right\}$$

que é o fecho em $L^2(\mathbb{R})$ das combinações lineares das funções $\varphi(x - k)$, $k \in \mathbb{Z}$. V_0 representa o conjunto das funções constantes por partes, sendo cada parte um intervalo de comprimento 1 da forma $[k, k + 1)$, $k \in \mathbb{Z}$.

Como $\{\varphi(x - k), k \in \mathbb{Z}\}$ é um conjunto ortonormal, segue que ele é uma base ortonormal para V_0 . Porém, $V_0 \subsetneq L^2(\mathbb{R})$, já que até mesmo a função definida por

$$f(x) = \begin{cases} 1 & \text{se } 0 \leq x < \frac{1}{2} \\ 0 & \text{caso contrário} \end{cases}$$

não pertence a V_0 , mas pertence a $L^2(\mathbb{R})$. Sendo assim, para representar todas as funções de $L^2(\mathbb{R})$ em termos de Scalings e Wavelets, precisamos definir outros conjuntos.

Um novo conjunto pode ser gerado apenas reescalando o conjunto $\{\varphi(x-k), k \in \mathbb{Z}\}$. Para tal, considere o conjunto $(1/2)\mathbb{Z}$ definido por:

$$(1/2)\mathbb{Z} = \left\{ \frac{k}{2}, k \in \mathbb{Z} \right\} \quad (2.3)$$

e a função $\varphi_2(\bullet) = \varphi(2\bullet)$ (note que o suporte de φ_2 é o intervalo $[0, \frac{1}{2}]$). A ideia é gerar um conjunto similar ao V_0 , porém agora, ao invés de ser um conjunto de funções constantes em intervalos de comprimento 1, será um conjunto de funções constantes em intervalos de comprimento $\frac{1}{2}$, o que seria equivalente a mudar a escala de uma função de V_0 .

Podemos agora definir o conjunto dado por $\left\{ \left(2^{\frac{1}{2}}\right) \varphi_2(x - k_2), k_2 \in (1/2)\mathbb{Z} \right\}$. Tal conjunto é composto de funções ortonormais (prova análoga à ortogonalidade das funções de $\{\varphi(x - k), k \in \mathbb{Z}\}$) que são constantes em intervalos de comprimento $\frac{1}{2}$, e logo será usado como a base do novo espaço V_1 . Utilizando (2.3) e o fato de que $\varphi_2(x) = \varphi(2x)$, podemos definir V_1 por:

$$V_1 = \left\{ \sum_{k=-\infty}^{\infty} a_k \varphi(2x - k), \sum |a_k|^2 < \infty \right\}$$

Das definições acima, é possível dizer que:

$$f(\bullet) \in V_0 \iff f(2\bullet) \in V_1$$

Para justificar essa afirmação, seja $f \in V_0$ dada por $f(x) = \sum_{k=-\infty}^{\infty} a_k \varphi(x - k)$, $\sum |a_k|^2 < \infty$. Então $f(2x) = \sum_{k=-\infty}^{\infty} a_k \varphi(2x - k) \implies f(2\bullet) \in V_1$. Por fim, se $f(2\bullet) \in V_1$ então $f(2x) = \sum_{k=-\infty}^{\infty} a_k \varphi(2x - k)$, $\sum |a_k|^2 < \infty$. Logo $f(x) = \sum_{k=-\infty}^{\infty} a_k \varphi(x - k) \implies f(\bullet) \in V_0$.

Podemos agora redefinir a identidade escala (2.2) por sua versão transladada, para $k \in \mathbb{Z}$:

$$\varphi(x - k) = \varphi(2x - 2k) + \varphi(2x - 2k - 1) \quad (2.4)$$

A equação (2.4) afirma que $V_0 \subseteq V_1$, já que ela representa uma base de V_0 como uma combinação linear de elementos de V_1 .

Podemos agora repetir o processo de gerar V_1 a partir de V_0 e produzir, por indução, os conjuntos V_j , $j \in \mathbb{Z}$, que serão os conjuntos de funções constantes em intervalos de comprimento 2^{-j} . Por esse procedimento, cada um deles terá as seguintes propriedades:

- $\{2^{j/2} \varphi(2^j x - k), k \in \mathbb{Z}\}$ é uma base ortonormal de V_j
- $f(x) \in V_j \iff f(2^{k-j} x) \in V_k$

- $V_j \subseteq V_k$ se $j \leq k$ (obtida a partir de reescalonamento de (2.4))

A sequência $\{V_j\}$ é um tipo de análise multiresolucional (será detalhada na próxima seção) e possui duas propriedades importantes:

1. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$

Demonstração. Claramente $0 \in V_j \forall j \in \mathbb{Z}$. Além disso, se f for constante, segue que $f \equiv 0$ pois $f \in L^2(\mathbb{R})$. Suponha por absurdo que $f \in V_j \forall j \in \mathbb{Z}$ e que f não é constante. Então, como $f \in V_0$, temos que: $f(x) = \sum_{k=-\infty}^{\infty} a_k \varphi(x-k)$, $\sum |a_k|^2 < \infty$. Claramente não podemos ter a_k iguais para todo $k \in \mathbb{Z}$, pois se não teríamos $\sum |a_k|^2 = \infty$ ou $f \equiv 0$ caso $a_k = 0 \forall k \in \mathbb{Z}$. Logo, existe $k_0 \in \mathbb{Z} - \{-1, 0\}$ tal que $a_{k_0} \neq a_{k_0+1}$.

Seja j_0 um inteiro tal que $2^{j_0} > |k_0| + 1$. Então, caso $k_0 > 0$, f não é constante no intervalo $[0, 2^{j_0})$ ou, caso $k_0 < 0$, f não é constante no intervalo $[-2^{j_0}, 0)$, o que significa que $f \notin V_{-j_0}$, o que é uma contradição. ■

2. $\bigcup_{j \in \mathbb{Z}} V_j$ é denso em $L^2(\mathbb{R})$

Demonstração. Considere uma função qualquer f^* em $L^2(\mathbb{R})$. Dado $\epsilon > 0$, considere agora uma função $f \in \mathcal{S}(\mathbb{R})$ (o espaço de Schwartz (REED; SIMON, 1980), composto por funções de derivadas contínuas em qualquer ordem e de decaimento rápido, o qual é denso em $L^2(\mathbb{R})$) tal que tenhamos

$$\int |f(x) - f^*(x)|^2 dx < \frac{\epsilon}{4}$$

Como $\|f\|_2 < \infty$, existe $C \in \mathbb{R}$, $C > 0$, tal que

$$\left| \int_{-C}^C |f(x)|^2 dx - \|f\|_2^2 \right| < \frac{\epsilon}{8}$$

Considere agora o conjunto $I = \bigcup_{i \in T} [2^{-k}i, 2^{-k}(i+1)]$ onde $T = \{i : [2^{-k}i, 2^{-k}(i+1)] \cap [-C, C] \neq \emptyset, i \in \mathbb{Z}\}$ (note que $[-C, C] \subset I$). Seja F_k definida por

$$F_k(x) = \begin{cases} \sum_{i=-N}^{N-1} f(2^k i) \varphi(2^{-k}x - i) & \text{se } x \in I \\ 0 & \text{caso contrário} \end{cases}$$

onde N é o índice i dos conjuntos $[2^{-k}i, 2^{-k}(i+1)]$ em I tal que $2^{-k}i$ é mínimo (analogamente, que $2^{-k}(i+1)$ é máximo). Como se trata de um somatório finito, segue que

$$\sum_{i=-N}^{N-1} |f(2^k i)|^2 < \infty$$

e portanto $F_k \in \bigcup_{j \in \mathbb{Z}} V_j$.

Pela continuidade uniforme de f ((REED; SIMON, 1980)), temos que existe $k > 0$ tal que, para todo $i \in [-N, N - 1]$

$$|f(x) - F_k(x)|^2 = |f(x) - f(2^k i)|^2 < \frac{\epsilon}{8(2^{-k+1}N)}, \quad x \in [2^k i, 2^k(i + 1)].$$

Logo, temos que

$$\begin{aligned} \int |f(x) - F_k(x)|^2 dx &= \int_{-\infty}^{-2^{-k}N} |f(x)|^2 dx + \int_{-2^{-k}N}^{2^{-k}N} |f(x) - F_k(x)|^2 dx + \\ &+ \int_{2^{-k}N}^{\infty} |f(x)|^2 dx \\ &\leq \frac{\epsilon}{8} + \int_{-2^{-k}N}^{2^{-k}N} \frac{\epsilon}{8(2^{-k+1}N)} dx \\ &= \frac{\epsilon}{8} + \frac{\epsilon}{8} \\ &= \frac{\epsilon}{4} \end{aligned}$$

Sendo assim, vale que

$$\begin{aligned} \int |F_k(x) - f^*(x)|^2 dx &\leq 2 \int |F_k(x) - f(x)|^2 dx + 2 \int |f(x) - f^*(x)|^2 dx \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon \end{aligned}$$

Como $\bigcup_{j \in \mathbb{Z}} V_j \subset L^2(\mathbb{R})$, segue que $\bigcup_{j \in \mathbb{Z}} V_j$ é denso em $L^2(\mathbb{R})$. ■

O fato de $\bigcup_{j \in \mathbb{Z}} V_j$ ser denso em $L^2(\mathbb{R})$ não significa que, combinando todas as bases $\{2^{\frac{j}{2}}\varphi(2^j x - k)\}$ de V_j , obtemos uma única base ortonormal para $L^2(\mathbb{R})$. A explicação se deve ao fato de que, ainda que $V_j \subseteq V_{j+1}$, a base ortonormal $\{2^{\frac{j}{2}}\varphi(2^j x - k)\}$ de V_j não está contida na base ortonormal $\{2^{\frac{j+1}{2}}\varphi(2^{j+1} x - k)\}$ de V_{j+1} e existem elementos de $\{2^{\frac{j}{2}}\varphi(2^j x - k)\}$ que não são ortogonais com determinados elementos de $\{2^{\frac{j+1}{2}}\varphi(2^{j+1} x - k)\}$ (como $2^{\frac{j_1}{2}}\varphi(2^{j_1} x)$ e $2^{\frac{j_2}{2}}\varphi(2^{j_2} x)$ com $j_2 > j_1$, por exemplo). Portanto, apenas com a função φ não é possível obter uma base ortonormal para $L^2(\mathbb{R})$.

Como estamos interessados em obter bases para $L^2(\mathbb{R})$, considere que no restante do capítulo o limite de funções ($f_n \rightarrow f$) é tomado considerando a norma $L^2(\mathbb{R})$, a não ser que seja explicitada outra norma para um dado limite.

2.1.3 Wavelet Haar e a base para $L^2(\mathbb{R})$

Como descrito anteriormente, $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$ é uma base para V_0 e temos que $V_0 \subseteq V_1$. Porém, ao invés de utilizar a base de V_1 , será adotado um novo procedimento: partindo da base de V_0 , completamos a base de V_1 com um conjunto $\{\psi(x - k)\}_{k \in \mathbb{Z}}$ (ortogonal a $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$) para uma certa função ψ . Isso equivale a determinar uma base para o complemento ortogonal de V_0 em V_1 , que aqui será denotado de W_0 , de modo que $V_1 = V_0 \oplus W_0$.

Para tal, a função ψ será exatamente a função Haar, definida em (1.1), a qual pode ser expressada em função de φ :

$$\psi(x) = \varphi(2x) - \varphi(2x - 1)$$

Afirmção: $\{\psi(x - k)\}_{k \in \mathbb{Z}}$ é uma base ortogonal para W_0 .

Demonstração. Que $\{\psi(x - k)\}_{k \in \mathbb{Z}}$ é uma base ortogonal já foi provado na Introdução. Assim, a prova consiste em mostrar que $\{\psi(x - k)\}_{k \in \mathbb{Z}} \subseteq W_0$ e, por fim, mostrar que $\{\varphi(x - k)\}_{k \in \mathbb{Z}} \cup \{\psi(x - k)\}_{k \in \mathbb{Z}}$ é uma base de V_1 , o que significa que $\{\psi(x - k)\}_{k \in \mathbb{Z}}$ é base de W_0 , já que $V_1 = V_0 \oplus W_0$.

- $\{\psi(x - k)\}_{k \in \mathbb{Z}} \subseteq W_0$

Para $k_1, k_2 \in \mathbb{Z}$, temos que:

$$\begin{aligned} \langle \psi(\bullet - k_1), \varphi(\bullet - k_2) \rangle &= \int_{-\infty}^{+\infty} \overline{\psi(x - k_1)} \varphi(x - k_2) dx \\ &= \int_{-\infty}^{+\infty} \varphi(2x - k_1) \varphi(x - k_2) dx - \int_{-\infty}^{+\infty} \varphi(2x - k_1 - 1) \varphi(x - k_2) dx \\ &= \int_{k_1}^{k_1 + \frac{1}{2}} \varphi(x - k_2) dx - \int_{k_1 + \frac{1}{2}}^{k_1 + 1} \varphi(x - k_2) dx \\ &= 0 \quad \forall k_1, k_2 \in \mathbb{Z} \end{aligned}$$

Logo, como $\langle \psi(\bullet - k_1), \varphi(\bullet - k_2) \rangle = 0 \quad \forall k_1, k_2 \in \mathbb{Z}$, segue que $\{\psi(x - k)\}_{k \in \mathbb{Z}} \subseteq W_0$.

- $\{\varphi(x - k)\}_{k \in \mathbb{Z}} \cup \{\psi(x - k)\}_{k \in \mathbb{Z}}$ é uma base de V_1

Basta notar que

$$\sqrt{2}\varphi(2x - k) = \begin{cases} \frac{\sqrt{2}}{2}\varphi\left(x - \left\lfloor \frac{k}{2} \right\rfloor\right) + \frac{\sqrt{2}}{2}\psi\left(x - \left\lfloor \frac{k}{2} \right\rfloor\right) & \text{se } k \text{ é ímpar} \\ \frac{\sqrt{2}}{2}\varphi\left(x - \left\lfloor \frac{k}{2} \right\rfloor\right) - \frac{\sqrt{2}}{2}\psi\left(x - \left\lfloor \frac{k}{2} \right\rfloor\right) & \text{se } k \text{ é par} \end{cases}$$

onde $\lfloor p \rfloor$ representa o maior número inteiro menor ou igual ao número real p . Assim, podemos expressar todos os elementos da base de V_1 por meio de elementos de $\{\varphi(x - k)\}_{k \in \mathbb{Z}} \cup \{\psi(x - k)\}_{k \in \mathbb{Z}}$, o que prova a afirmação deste item.



Reescalando o espaço W_0 , obtemos que:

$$V_{j+1} = V_j \oplus W_j \quad (2.5)$$

e que $\{2^{\frac{j}{2}}\psi(2^j x - k)\}_{k \in \mathbb{Z}}$ é uma base ortonormal de W_j . Sendo $V_{j^*} \subseteq V_{j^*+1}$, temos que $V_{j^*} \cap V_{j^*+1} = V_{j^*}$ e

$$\bigcup_{j=-\infty}^{j^*+1} V_j = V_{j^*+1}. \quad (2.6)$$

Assim, pela propriedade $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$:

$$\bigcap_{j \in \mathbb{Z}} V_j = \lim_{j \rightarrow -\infty} V_j = \{0\} \quad (2.7)$$

Por (2.5), temos que:

$$V_{j^*+1} = V_{j^*} \oplus W_{j^*} = V_{j^*-1} \oplus W_{j^*-1} \oplus W_{j^*} = \dots = \left(\lim_{j \rightarrow -\infty} V_j \right) \oplus \bigoplus_{j'=-\infty}^{j^*} W_{j'} = \bigoplus_{j'=-\infty}^{j^*} W_{j'}$$

e que

$$\bigcup_{j=j^*+2}^{\infty} V_j = V_{j^*+1} \oplus \bigoplus_{j'=j^*+1}^{\infty} W_{j'} \quad (2.8)$$

Juntando (2.6), (2.7) e (2.8) com (2) temos que:

$$L^2(\mathbb{R}) = \bigcup_{j=-\infty}^{\infty} V_j = \bigoplus_{j=-\infty}^{\infty} W_j \quad (2.9)$$

Assim, a base $\{2^{\frac{j}{2}}\psi(2^j x - k)\}_{k, j \in \mathbb{Z}}$ é uma base ortonormal de $L^2(\mathbb{R})$.

Existe também uma variação de (2.9) fazendo $j^* = -1$ em (2.8) e (2.6):

$$L^2(\mathbb{R}) = V_0 \oplus \left(\bigoplus_{j=0}^{\infty} W_j \right) \quad (2.10)$$

A equação (2.10) mostra que $L^2(\mathbb{R})$ tem uma base dada por $\{\varphi(x - k)\}_{k \in \mathbb{Z}} \cup \{2^{\frac{j}{2}}\psi(2^j x - k)\}_{k \in \mathbb{Z}, j \in \mathbb{Z}_+}$, a qual é a mais parecida com a base considerada na expansão em Wavelet Haar vista em (1.2).

2.2 Análise Multiresolucional

Na seção anterior, toda a construção da base ortonormal para o $L^2(\mathbb{R})$ foi baseada na função característica (definida em (2.1) para o intervalo $[0, 1]$) e na Wavelet Haar. Nesta seção, será apresentada uma teoria generalizada para uma função Scaling φ qualquer, com o objetivo de construir uma análise multiresolucional.

Definição 2.2.1. Uma análise multiresolucional $\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots$ com uma função Scaling φ é uma sequência crescente de subespaços do $L^2(\mathbb{R})$ que satisfazem as seguintes condições:

1. (Densidade) $\bigcup_j V_j$ é denso em $L^2(\mathbb{R})$,
2. (Separação) $\bigcap_j V_j = \{0\}$,
3. (Escalonamento) $f(x) \in V_j \iff f(2^{-j}x) \in V_0$,
4. (Ortonormalidade) $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é uma base ortonormal de V_0 .

A partir da definição, podemos deduzir algumas propriedades da análise multiresolucional e da função φ :

Teorema 2.2.1. Considerando a função φ definida em 2.2.1, temos que o conjunto $\{2^{\frac{j}{2}}\varphi(2^j x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é uma base ortonormal de V_j

Demonstração. Considere a transformação linear T_j dada por:

$$\begin{aligned} T_j : L^2(\mathbb{R}) &\longrightarrow L^2(\mathbb{R}) \\ f(x) &\longrightarrow 2^{\frac{j}{2}}f(2^j x) \end{aligned}$$

Pelo item 3 da definição 2.2.1, obtemos que $T_j(V_0) = V_j$, e pela definição da transformação, $T_j(\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}) = \{2^{\frac{j}{2}}\varphi(2^j x - \gamma)\}_{\gamma \in \mathbb{Z}}$. Note que a transformação T_j é ortogonal, já que se, dadas duas funções $f, g \in L^2(\mathbb{R})$ tais que $\langle f, g \rangle = 0$, então

$$\begin{aligned} \langle T_j f, T_j g \rangle &= \langle 2^{\frac{j}{2}}f(2^j x), 2^{\frac{j}{2}}g(2^j x) \rangle \\ &= 2^j \langle f(2^j x), g(2^j x) \rangle \\ &= 2^j \int \overline{f(2^j x)} g(2^j x) dx \\ &= \int \overline{f(y)} g(y) dy \\ &= 0 \end{aligned}$$

Como transformações ortogonais preservam a propriedade de base ortonormal de um espaço, segue que $\{2^{\frac{j}{2}}\varphi(2^j x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é uma base ortonormal de V_j .



Pela definição da análise multiresolucional, temos $\varphi \in V_0 \subseteq V_1$. Portanto, é possível escrever φ em função da base de V_1 :

$$\varphi(x) = \sum_{\gamma \in \mathbb{Z}} a(\gamma) \varphi(2x - \gamma) \quad (2.11)$$

A equação (2.11) representa uma generalização de (2.2) e (2.4) e portanto também será chamada de identidade escala.

É interessante agora obter condições sobre os coeficientes $a(\gamma)$ da equação (2.11). Os teoremas 2.2.2 e 2.2.3 representam algumas delas.

Teorema 2.2.2. Seja $a(\gamma)$ como em (2.11), então

$$\sum_{\gamma \in \mathbb{Z}} |a(\gamma)|^2 = 2. \quad (2.12)$$

Demonstração. Como

$$\langle \varphi(x), \varphi(x) \rangle = \|\varphi(x)\|_2^2 = \left\| \sum_{\gamma \in \mathbb{Z}} a(\gamma) \varphi(2x - \gamma) \right\|_2^2 \quad (2.13)$$

pela ortogonalidade das funções $\varphi(2x - \gamma)$ em $L^2(\mathbb{R})$ temos

$$\left\| \sum_{\gamma \in \mathbb{Z}} a(\gamma) \varphi(2x - \gamma) \right\|_2^2 = \sum_{\gamma \in \mathbb{Z}} |a(\gamma)|^2 \|\varphi(2x - \gamma)\|_2^2 \quad (2.14)$$

Pela definição de norma em $L^2(\mathbb{R})$ e pelo item 4 da definição 2.2.1:

$$\begin{aligned} \|\varphi(2x - \gamma)\|_2^2 &= \int \overline{\varphi(2x - \gamma)} \varphi(2x - \gamma) \, dx \\ &= \frac{1}{2} \int \overline{\varphi(x)} \varphi(x) \, dx \\ &= \frac{1}{2} \|\varphi(x)\|_2^2 = \frac{1}{2} \end{aligned} \quad (2.15)$$

Substituindo (2.15) e (2.14) em (2.13):

$$1 = \|\varphi(x)\|_2^2 = \frac{1}{2} \sum_{\gamma \in \mathbb{Z}} |a(\gamma)|^2 \implies \sum_{\gamma \in \mathbb{Z}} |a(\gamma)|^2 = 2$$



Teorema 2.2.3. Seja $a(\gamma)$ como em (2.11), então

$$a(\gamma) = 2 \int \varphi(x) \overline{\varphi(2x - \gamma)} dx$$

Demonstração. A prova segue diretamente do fato de $\{\sqrt{2}\varphi(2x - \gamma)\}_{\gamma \in \mathbb{Z}}$ ser uma base ortonormal:

$$\begin{aligned} \varphi(x) &= \sum_{\gamma \in \mathbb{Z}} \langle \varphi(\bullet), \sqrt{2}\varphi(2\bullet - \gamma) \rangle \sqrt{2}\varphi(2x - \gamma) \\ &= \sum_{\gamma \in \mathbb{Z}} [2\langle \varphi(\bullet), \varphi(2\bullet - \gamma) \rangle] \varphi(2x - \gamma) \end{aligned}$$

Portanto, pela representação feita em (2.11), temos que:

$$a(\gamma) = 2\langle \varphi(\bullet), \varphi(2\bullet - \gamma) \rangle = \int \varphi(x) \overline{\varphi(2x - \gamma)} dx$$

■

Pela definição, uma função Scaling φ define uma análise multiresolucional. Porém a afirmação inversa não é verdadeira. Portanto, são necessárias condições que permitam determinar se uma dada função gera uma análise multiresolucional. Tais condições serão definidas por meio das propriedades dos coeficientes $a(\gamma)$, que seguem da definição 2.2.1.

Teorema 2.2.4. Seja $a(\gamma)$ como em (2.11), então

$$\sum_{\gamma' \in \mathbb{Z}} a(\gamma') \overline{a(2\gamma + \gamma')} = 2\delta(\gamma, 0) \quad (2.16)$$

onde δ representa a função delta de Kronecker.

Demonstração. Pelo item 4 da definição 2.2.1, concluímos que:

$$\int \varphi(x - \gamma) \overline{\varphi(x)} dx = \delta(\gamma, 0); \quad (2.17)$$

Substituindo (2.11) em (2.17):

$$\begin{aligned} \int \varphi(x - \gamma) \overline{\varphi(x)} dx &= \sum_{\gamma' \in \mathbb{Z}} a(\gamma') \int \varphi(2x - 2\gamma - \gamma') \overline{\varphi(x)} dx \\ &= \sum_{\gamma' \in \mathbb{Z}} \sum_{\gamma'' \in \mathbb{Z}} a(\gamma') \overline{a(\gamma'')} \int \varphi(2x - 2\gamma - \gamma') \overline{\varphi(2x - \gamma'')} dx \end{aligned}$$

Realizando uma mudança de variável $x \rightarrow 2^{-1}x$ e pela ortonormalidade de φ :

$$\sum_{\gamma' \in \mathbb{Z}} \sum_{\gamma'' \in \mathbb{Z}} a(\gamma') \overline{a(\gamma'')} \int \varphi(2x - 2\gamma - \gamma') \overline{\varphi(2x - \gamma'')} dx = 2^{-1} \sum_{\gamma'' = 2\gamma + \gamma'} a(\gamma') \overline{a(\gamma'')}$$

Substituindo em (2.17):

$$\sum_{\gamma' \in \mathbb{Z}} a(\gamma') \overline{a(2\gamma + \gamma')} = 2\delta(\gamma, 0)$$

■

Perceba que fazendo $\gamma = 0$ em (2.16), obtemos (2.12) como um caso especial.

Uma propriedade advinda da densidade, a qual é o item 1 da definição 2.2.1 (análise multiresolucional), é que $\int \varphi(x) dx \neq 0$. Caso contrário, isso seria verdade para todas as funções em todos os espaços V_j , e logo cada um deles apenas conteria funções cuja integral é nula. Assim, $\cup_j V_j$ não seria denso em $L^2(\mathbb{R})$. Tal propriedade será utilizada no seguinte teorema:

Teorema 2.2.5. Seja $a(\gamma)$ como em (2.11), então

$$\sum_{\gamma \in \mathbb{Z}} a(\gamma) = 2 \tag{2.18}$$

Demonstração. Integrando a equação (2.11), obtemos que:

$$\begin{aligned} \int \varphi(x) dx &= \sum_{\gamma \in \mathbb{Z}} a(\gamma) \int \varphi(2x - \gamma) dx \\ &= \sum_{\gamma \in \mathbb{Z}} a(\gamma) 2^{-1} \int \varphi(x) dx \\ &\implies \sum_{\gamma \in \mathbb{Z}} a(\gamma) = 2 \end{aligned}$$

onde a última implicação é devido a $\int \varphi(x) dx \neq 0$.

■

2.2.1 Condições necessárias para uma análise multiresolucional

Nas próximas seções, o objetivo será produzir uma Scaling que gere uma análise multiresolucional, para que a partir dela seja possível definir uma função Wavelet. Para tal, serão necessários alguns passos:

1. Produzir soluções $a(\gamma)$ para as identidades algébricas (2.16) e (2.18);

2. Definir uma função Scaling φ a partir da identidade escala (2.11);
3. Demonstrar que a função φ definida gera uma análise multiresolucional;
4. Construir uma função Wavelet.

No que se refere ao passo 3, a propriedade da análise multiresolucional mais complexa de se demonstrar é a Ortonormalidade (item 4 da definição de Análise Multiresolucional) e será tratada em seções posteriores.

Devido a tal fato, neste momento será mostrado que, uma vez que φ atenda à Ortonormalidade e a uma condição de normalização

$$\int \varphi(x) dx = 1 \quad (2.19)$$

então a condição de Densidade e de Separação da definição de análise multiresolucional também serão atendidas.

Teorema 2.2.6. Seja V_0 um subespaço qualquer de $L^2(\mathbb{R})$, que está contido em $L^\infty(\mathbb{R})$ e que tem a propriedade de que

$$\|f\|_\infty \leq c\|f\|_2 \quad \forall f \in V_0 \quad (2.20)$$

Defina V_j pela condição de Escala da definição de análise multiresolucional (sem a necessidade de que $V_j \subseteq V_{j+1}$). Então a propriedade de Separação da definição de análise multiresolucional vale.

Demonstração. Seja $f \in V_0$ que atenda à equação (2.20). Então, como pela propriedade de Escala da definição de análise multiresolucional $f(x) \in V_0 \iff f(2^j x) \in V_j$, temos que:

$$\|f\|_2 = \sqrt{\int |f(x)|^2 dx} = \sqrt{\int 2^j |f(2^j x)|^2 dx} = 2^{\frac{j}{2}} \|f_j\|_2$$

onde $f_j(\bullet) = f(2^j \bullet)$. Com relação à norma $L^\infty(\mathbb{R})$, obtemos

$$\|f\|_\infty = \sup\{|f(x)|, x \in \mathbb{R}\} = \sup\{|f(2^j x)|, x \in \mathbb{R}\} = \|f_j\|_\infty$$

Portanto, vale a seguinte desigualdade:

$$\|f_j\|_\infty = \|f\|_\infty \leq c\|f\|_2 = c2^{\frac{j}{2}} \|f_j\|_2 \quad \forall j \in \mathbb{Z} \quad (2.21)$$

Sendo assim, levando em conta a equação (2.20), podemos escrever que

$$\|f\|_\infty \leq c2^{\frac{j}{2}} \|f\|_2 \quad \forall f \in V_j \quad (2.22)$$

Considere agora uma função $f \in \cap_j V_j$ fixa, então a inequação (2.22) vale para todo $j \in \mathbb{Z}$. Logo, fazendo $j \rightarrow -\infty$ temos que $\|f\|_\infty = 0$. Logo, $f = 0$, o que implica que $\cap_j V_j = \{0\}$. ■

A estimativa proposta na equação (2.20) pode ser atendida considerando φ limitada e com suporte compacto, como é provado no seguinte lema:

Lema 2.2.7. Se φ é limitada, possui suporte compacto e a condição 4 de Ortonormalidade da definição 2.2.1 é atendida, então vale a inequação (2.20).

Demonstração. Seja $f \in V_0$. Pela item 4 da definição 2.2.1 (ortonormalidade):

$$\begin{aligned} f(x) &= \sum_{\gamma \in \mathbb{Z}} \varphi(x - \gamma) \int f(y) \overline{\varphi(y - \gamma)} dy \\ &= \int K(x, y) f(y) dy \end{aligned}$$

onde $K(x, y) = \sum_{\gamma \in \mathbb{Z}} \varphi(x - \gamma) \overline{\varphi(y - \gamma)}$. Então, pela desigualdade de Hölder:

$$\begin{aligned} |f(x)| &\leq \left(\int |K(x, y)|^2 dy \right)^{\frac{1}{2}} \|f\|_2 \\ &= \left(\int \left| \sum_{\gamma \in \mathbb{Z}} \varphi(x - \gamma) \overline{\varphi(y - \gamma)} \right|^2 dy \right)^{\frac{1}{2}} \|f\|_2 \\ &= \left(\int \sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2 |\overline{\varphi(y - \gamma)}|^2 dy \right)^{\frac{1}{2}} \|f\|_2 \\ &= \left(\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2 \int |\overline{\varphi(y - \gamma)}|^2 dy \right)^{\frac{1}{2}} \|f\|_2 \\ &= \left(\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2 \right)^{\frac{1}{2}} \|f\|_2 \end{aligned}$$

Como φ possui suporte compacto, segue que $\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2$ é uniformemente limitada, logo existe $c > 0$ tal que:

$$\left(\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2 \right)^{\frac{1}{2}} \leq c$$

o que prova o teorema. ■

Teorema 2.2.8. Suponha que φ tenha suporte compacto e satisfaça (2.11), (2.19) e a condição 4 de Ortonormalidade da definição 2.2.1. Então a condição 1 de Densidade da definição 2.2.1 é satisfeita.

Demonstração. Considere P_j a projeção ortogonal de $f \in L^2(\mathbb{R})$ em V_j dada por:

$$P_j f(x) = 2^j \sum_{\gamma \in \mathbb{Z}} \varphi(2^j x - \gamma) \int f(y) \overline{\varphi(2^j y - \gamma)} dy$$

A prova se baseia em mostrar que $\lim_{j \rightarrow \infty} P_j f = f$ para toda $f \in L^2(\mathbb{R})$ na norma $L^2(\mathbb{R})$, o que implica que $\cup_j V_j$ é denso em $L^2(\mathbb{R})$.

Como provado na seção anterior, a função Scaling φ dada pela função característica no intervalo $[0, 1]$ gera uma análise multiresolucional e portanto satisfaz a condição de densidade da definição 2.2.1. Logo, basta provar que, para todo intervalo compacto A , $\lim_{j \rightarrow \infty} P_j \chi_A = \chi_A$ na norma $L^2(\mathbb{R})$, onde χ_A é a função característica no intervalo A . Tal fato pode ser justificado pela seguinte argumento: considerando que $\lim_{j \rightarrow \infty} P_j \chi_A = \chi_A$, então

$$\begin{aligned} \|P_j f - f\|_2 &= \|P_j f - f + \sum_{i=1}^n c_i \chi_i - \sum_{i=1}^n c_i \chi_i + P_j \sum_{i=1}^n c_i \chi_i - P_j \sum_{i=1}^n c_i \chi_i\|_2 \\ &\leq \|P_j f - P_j \sum_{i=1}^n c_i \chi_i\|_2 + \|\sum_{i=1}^n c_i \chi_i - f\|_2 + \|P_j \sum_{i=1}^n c_i \chi_i - \sum_{i=1}^n c_i \chi_i\|_2 \\ &= \|P_j (\sum_{i=1}^n c_i \chi_i - f)\|_2 + \|\sum_{i=1}^n c_i \chi_i - f\|_2 + \|P_j \sum_{i=1}^n c_i \chi_i - \sum_{i=1}^n c_i \chi_i\|_2 \\ &\leq 2 \|\sum_{i=1}^n c_i \chi_i - f\|_2 + \|\sum_{i=1}^n c_i P_j \chi_i - \sum_{i=1}^n c_i \chi_i\|_2 \end{aligned}$$

Logo, como as funções características satisfazem a condição de densidade da definição 2.2.1 e $\lim_{j \rightarrow \infty} P_j \chi_A = \chi_A$, podemos obter $j > 0$ e um conjunto $\{\chi_i : i \in \mathbb{N}\}$ tal que $\|\sum_{i=1}^n c_i \chi_i - f\|_2 < \frac{1}{3}\epsilon$ e $\|\sum_{i=1}^n c_i P_j \chi_i - \sum_{i=1}^n c_i \chi_i\|_2 < \frac{1}{3}\epsilon$. Logo $\|P_j f - f\|_2 < \epsilon$.

Para provar que $\lim_{j \rightarrow \infty} P_j \chi_A = \chi_A$ na norma $L^2(\mathbb{R})$, note que:

$$\begin{aligned} \|P_j \chi_A\|_2^2 &= \int \left| 2^j \sum_{\gamma \in \mathbb{Z}} \varphi(2^j x - \gamma) \int_A \overline{\varphi(2^j y - \gamma)} dy \right|^2 dx \\ &= 2^{2j} \int \sum_{\gamma \in \mathbb{Z}} |\varphi(2^j x - \gamma)|^2 \left| \int_A \overline{\varphi(2^j y - \gamma)} dy \right|^2 dx \\ &= 2^{2j} \sum_{\gamma \in \mathbb{Z}} \left| \int_A \overline{\varphi(2^j y - \gamma)} dy \right|^2 \int |\varphi(2^j x - \gamma)|^2 dx \\ &= 2^j \sum_{\gamma \in \mathbb{Z}} \left| \int_A \overline{\varphi(2^j y - \gamma)} dy \right|^2 \\ &= 2^{-j} \sum_{\gamma \in \mathbb{Z}} \left| \int_{2^j A} \overline{\varphi(y - \gamma)} dy \right|^2 \end{aligned}$$

Considerando a normalização em (2.19), se todo o suporte de $\varphi(\bullet - \gamma)$ estiver contido em $2^j A$, então $\left| \int_{2^j A} \overline{\varphi(y - \gamma)} dy \right|^2 = 1$, mas se a interseção desse suporte com $2^j A$ for nula, então $\left| \int_{2^j A} \overline{\varphi(y - \gamma)} dy \right|^2 = 0$.

Seja D_1 o comprimento do intervalo A , D_2 o comprimento do suporte de φ e $M = \int |\overline{\varphi(x)}| dx$. Considerando j tal que $2^j D_1 > 10(1 + D_2 + M)$, podemos afirmar que:

$$2^{-j}(2^j D_1 - 2M - 2D_2) \leq 2^{-j} \sum_{\gamma \in \mathbb{Z}} \left| \int_{2^j A} \overline{\varphi(y - \gamma)} dy \right|^2 \leq 2^{-j}(2^j D_1 + 2M + 2D_2)$$

Fazendo $j \rightarrow \infty$, temos que:

$$\lim_{j \rightarrow \infty} \|P_j \chi_A\|_2^2 = D_1 = \|\chi_A\|^2$$

Portanto, sendo $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ uma base ortonormal, segue do teorema de Pitágoras que:

$$\lim_{j \rightarrow \infty} P_j \chi_A = \chi_A$$

■

2.3 Wavelets

A partir desta seção, será utilizada uma nova notação. A função φ será denotada por ψ_0 e fará parte de um par de elementos ψ_0, ψ_1 , sendo ψ_1 a função geradora da Wavelet. O objetivo é encontrar ψ_0, ψ_1 tais que $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}, k \in \{0,1\}}$ seja uma base ortonormal de V_1 . Como $\{\varphi_k(2x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é base ortonormal de V_1 , ψ_0 e ψ_1 devem ser combinações lineares das funções em $\{\varphi_k(2x - \gamma)\}_{\gamma \in \mathbb{Z}}$, e logo devem satisfazer a seguinte identidade:

$$\psi_k = \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) \varphi(2x - \gamma), \quad k \in \{0, 1\} \quad (2.23)$$

Porém, é necessário atribuir propriedades aos coeficientes a_k . Primeiramente, assim como (2.16), esperamos obter

$$\sum_{\gamma' \in \mathbb{Z}} a_j(\gamma') \overline{a_k(2\gamma + \gamma')} = 2\delta(j, k)\delta(\gamma, 0) \quad (2.24)$$

A condição obtida na Função Scaling da Haar, em que $\int \varphi(x) dx \neq 0$ não pode ser esperada da Wavelet ψ_1 , então, assim como (2.18), a outra condição que pode ser atribuída aos coeficientes a_k (em particular os a_0) é:

$$\sum_{\gamma \in \mathbb{Z}} a_0(\gamma) = 2 \quad (2.25)$$

2.3.1 Propriedades das Wavelets

Nesta subseção serão demonstrados alguns lemas e teoremas relacionados às Wavelets quando se estabelecem algumas condições.

Lema 2.3.1. Se $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal e se $a_j(\gamma)$ satisfaz (2.24) e (2.25), então $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$ é um conjunto ortonormal.

Demonstração. Basta mostrar que

$$\int \psi_j(x) \overline{\psi_k(x - \gamma)} dx = \delta(j, k) \delta(\gamma, 0)$$

De fato, utilizando (2.24):

$$\int \psi_j(x) \overline{\psi_k(x - \gamma)} dx = \sum_{\gamma' \in \mathbb{Z}} \sum_{\gamma'' \in \mathbb{Z}} a_j(\gamma') \overline{a_k(\gamma'')} \int \varphi(2x - \gamma') \overline{\varphi_k(2x - 2\gamma - 2\gamma'')} dx \quad (2.26)$$

Porém, devido à ortogonalidade de $\varphi(x - \gamma)$, temos que:

$$\begin{aligned} \int \varphi(2x - \gamma') \overline{\varphi_k(2x - 2\gamma - 2\gamma'')} dx &\stackrel{2x \rightarrow x}{=} \frac{1}{2} \int \varphi(x - \gamma') \overline{\varphi_k(x - 2\gamma - 2\gamma'')} dx \\ &= \frac{1}{2} \delta(\gamma', 2\gamma + \gamma'') \end{aligned}$$

Portanto, temos que (2.26) se resume a:

$$\sum \sum_{\gamma' = 2\gamma + \gamma''} a_j(\gamma') \overline{a_k(\gamma'')} \frac{1}{2}$$

Utilizando (2.24) obtemos:

$$\sum \sum_{\gamma' = 2\gamma + \gamma''} a_j(\gamma') \overline{a_k(\gamma'')} \frac{1}{2} = \delta(j, k) \delta(\gamma, 0)$$

■

Com esse lema, provamos que $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$ é um conjunto ortonormal em V_1 , porém ainda não mostra que ele é uma base de V_1 . Para tal, devemos mostrar que cada função $\varphi(2x - \gamma)$ é uma combinação linear de tais funções. Os coeficientes dessa combinação seriam, utilizando (2.23):

$$\int \varphi(2x - \tilde{\gamma}) \overline{\psi_k(x - \gamma)} dx = \sum \overline{a_k(\gamma')} \int \varphi(2x - \tilde{\gamma}) \varphi(2x - 2\gamma - \gamma') dx = \frac{1}{2} \overline{a_k(\tilde{\gamma} - 2\gamma)}$$

Portanto, é necessário mostrar que

$$\frac{1}{2} \sum_{k=0,1} \sum_{\gamma \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} - 2\gamma)} \psi_k(x - \gamma) \quad (2.27)$$

é igual a $\varphi(2x - \tilde{\gamma})$. Substituindo (2.23) na expressão (2.27):

$$\begin{aligned} & \frac{1}{2} \sum_{k=0,1} \sum_{\gamma \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} - 2\gamma)} \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma') \varphi(2x - 2\gamma - \gamma') \\ & \stackrel{2\gamma + \gamma' = \gamma''}{=} \frac{1}{2} \sum_{k=0,1} \sum_{\gamma \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} - 2\gamma)} \sum_{\gamma'' \in \mathbb{Z}} a_k(\gamma'' - 2\gamma) \varphi(2x - \gamma'') \\ & \stackrel{\gamma = -\gamma'''}{=} \frac{1}{2} \sum_{k=0,1} \sum_{\gamma''' \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} + 2\gamma''')} \sum_{\gamma'' \in \mathbb{Z}} a_k(\gamma'' + 2\gamma''') \varphi(2x - \gamma'') \\ & = \sum_{\gamma'' \in \mathbb{Z}} \left(\frac{1}{2} \sum_{k=0,1} \sum_{\gamma''' \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} + 2\gamma''')} a_k(2\gamma''' + \gamma'') \right) \varphi(2x - \gamma'') \end{aligned}$$

Então, basta mostrar que:

$$\sum_{k=0,1} \sum_{\gamma''' \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} + 2\gamma''')} a_k(2\gamma''' + \gamma'') = 2\delta(\tilde{\gamma}, \gamma'') \quad (2.28)$$

para $\tilde{\gamma} = 0$ ou 1 , por um argumento de translação. Tal desenvolvimento é utilizado para formular o seguinte lema:

Lema 2.3.2. A equação (2.28) sempre vale, i.e., $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$ é uma base ortonormal para V_1 .

A prova desse lema (2.3.2) será deixada para a próxima seção, pois envolverá a transformada de Fourier.

O próximo teorema mostrará que a Wavelet ψ_1 é uma base para $L^2(\mathbb{R})$ quando são estabelecidas algumas condições sobre os coeficientes a_k e a função Scaling φ .

Teorema 2.3.3. Suponha que φ gere uma análise multiresolucional e $a_k(\gamma)$ satisfaça (2.24) e (2.25) com ψ_k definida por (2.23) e $\psi_0 = \varphi$. Então as funções $\{2^{j/2}\psi_1(2^j x - \gamma)\}$ para $j \in \mathbb{Z}$, $\gamma \in \mathbb{Z}$, formam uma base ortonormal de $L^2(\mathbb{R})$

Demonstração. Seja W_0 o complemento ortogonal de V_0 em V_1 , ou seja, $V_1 = V_0 \oplus W_0$. Podemos dizer que $\{\psi_1(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é uma base ortonormal de W_0 . Tal afirmação é obtida pelo lema 2.3.2 apenas removendo $\{\psi_0(x - \gamma)\}_{\gamma \in \mathbb{Z}}$, que é uma base de V_0 .

Pela propriedade de escala:

$$V_{j+1} = V_j \oplus W_j$$

e $\{2^{j/2}\psi_1(2^j - \gamma)\}_{\gamma \in \mathbb{Z}}$ é uma base ortonormal de W_j . Mas

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$$

pela condição de densidade. Portanto as funções $\{2^{j/2}\psi_1(2^j x - \gamma)\}_{j \in \mathbb{Z}; \gamma \in \mathbb{Z}}$ formam uma base ortonormal de $L^2(\mathbb{R})$. ■

É importante ressaltar que, sendo $L^2(\mathbb{R}) = V_0 \oplus \left(\bigoplus_{j=0}^{\infty} W_j \right)$, o conjunto de funções $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ junto com $\{2^{j/2}\psi_1(2^j x - \gamma)\}_{j \geq 0; \gamma \in \mathbb{Z}}$ formam uma base ortonormal para $L^2(\mathbb{R})$. A vantagem dessa representação da base é que não é mais necessário utilizar as escalas de resolução $j < 0$ pela simples adição da base $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$.

Resumindo os resultados vistos até agora, temos as condições necessárias para que a função Scaling gere uma análise multiresolucional. Além disso, temos também condições para que as próprias Wavelets sejam uma base ortonormal para $L^2(\mathbb{R})$:

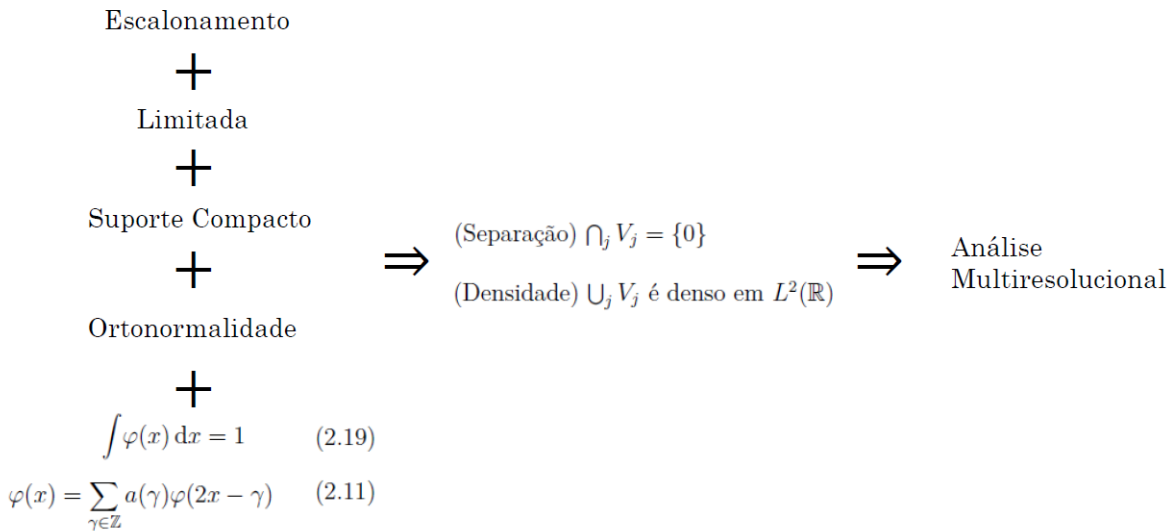


Figura 1 – Condições sobre as funções Scaling φ para obtermos uma análise multiresolucional

$\psi_k = \sum_{\gamma \in \mathbb{Z}} a_k(\gamma)\varphi(2x - \gamma), \quad k \in \{0, 1\}$	(2.23)		
$\sum_{\gamma' \in \mathbb{Z}} a_j(\gamma')\overline{a_k(2\gamma + \gamma')} = 2\delta(j, k)\delta(\gamma, 0)$	(2.24)	\Rightarrow	$\{2^{j/2}\psi_1(2^j x - \gamma)\}$ $j \in \mathbb{Z}, \gamma \in \mathbb{Z}$ é base ortonormal de $L^2(\mathbb{R})$
$\sum_{\gamma \in \mathbb{Z}} a_0(\gamma) = 2$	(2.25)		

Figura 2 – Condições sobre os coeficientes a_0 e a_1 para que o conjunto $\{2^{j/2}\psi_1(2^j x - \gamma)\}_{\gamma, j \in \mathbb{Z}}$ seja base ortonormal de $L^2(\mathbb{R})$

$$\begin{array}{ccc}
\text{Ortonormalidade } (\varphi) & & \\
+ & \Rightarrow & \text{Ortonormalidade } (\psi_k) \\
\sum_{\gamma' \in \mathbb{Z}} a_j(\gamma') \overline{a_k(2\gamma + \gamma')} = 2\delta(j, k)\delta(\gamma, 0) & (2.24) & \\
\sum_{\gamma \in \mathbb{Z}} a_0(\gamma) = 2 & (2.25) &
\end{array}$$

Figura 3 – Condições sobre os coeficientes a_0 e a_1 para garantir a ortonormalidade das funções ψ_k

2.4 Transformada de Fourier e Wavelets

Devido à estrutura convolucional das propriedades demonstradas até agora, nesta seção será aplicada a Transformada de Fourier a tais equações que, devido ao teorema da Convolução, as transforma em simples multiplicações. O objetivo dessa etapa é construir os coeficientes da identidade Scaling, os quais são mais facilmente deduzidos no domínio da transformada de Fourier.

Primeiramente, serão enunciados teoremas e definições em $\mathbb{S}(\mathbb{R})$ para em seguida enunciar-los em $\mathbb{S}'(\mathbb{R})$, o espaço dual de Schwartz (REED; SIMON, 1980). Isso será necessário pois consideraremos a transformada de polinômios trigonométricos, que não estão em $\mathbb{S}(\mathbb{R})$ e nem em $L^2(\mathbb{R})$. As provas de tais teoremas podem ser encontrados em (REED; SIMON, 1980). A definição formal de $\mathbb{S}(\mathbb{R})$ e $\mathbb{S}'(\mathbb{R})$ é dada a seguir:

Definição 2.4.1. As funções de decaimento rápido $\mathbb{S}(\mathbb{R})$ é o conjunto das funções infinitamente diferenciáveis φ em \mathbb{R} tal que

$$\|\varphi\|_{\alpha, \beta} \equiv \sup_{x \in \mathbb{R}} |x^\alpha D^\beta \varphi(x)| < \infty \quad (2.29)$$

para todo $\alpha, \beta \in \mathbb{Z}_+$, sendo $D^\beta \varphi(x)$ a derivada de ordem β de $\varphi(x)$ em relação a x .

Definição 2.4.2. O espaço topológico dual de $\mathbb{S}(\mathbb{R})$, denotado por $\mathbb{S}'(\mathbb{R})$, é chamado de espaço de distribuições temperadas. Em outras palavras, consiste no espaço de funcionais lineares que levam funções em $\mathbb{S}(\mathbb{R})$ ao corpo \mathbb{C} .

Podemos agora definir a transformada de Fourier e sua inversa no espaço $\mathbb{S}(\mathbb{R})$:

Definição 2.4.3. A Transformada de Fourier $\hat{\varphi}$ (ou $\mathcal{F}\{\varphi\}$) de uma função φ em $\mathbb{S}(\mathbb{R})$ será igual à vista em (STRICHARTZ, 1993):

$$\hat{\varphi}(\xi) = \int e^{2\pi i x \xi} \varphi(x) dx$$

sendo a Transformada Inversa de Fourier ($\mathcal{F}^{-1}\{\varphi\}$) dada por:

$$\varphi(x) = \int e^{-2\pi i x \xi} \hat{\varphi}(\xi) d\xi$$

que também pode ser escrita por:

$$\hat{\varphi}(x) = \varphi(-x)$$

Note que a transformada e sua inversa estão bem definidas, pois as funções de $\mathcal{S}(\mathbb{R})$ tem decaimento rápido.

Um resultado importante da transformada de Fourier é a chamada fórmula de Plancherel, que diz que a norma da transformada de uma função em $\mathcal{S}(\mathbb{R})$ é preservada pela transformada de Fourier:

Teorema 2.4.1. Suponha que $\varphi \in \mathcal{S}(\mathbb{R})$. Então vale que

$$\|\varphi\|_2 = \|\hat{\varphi}\|_2 \tag{2.30}$$

Precisaremos utilizar a transformada não apenas em $\mathcal{S}(\mathbb{R})$, mas também em polinômios trigonométricos, que pertencem a $\mathcal{S}'(\mathbb{R})$. Sendo assim, é necessário definir a transformada de Fourier no dual de $\mathcal{S}(\mathbb{R})$:

Definição 2.4.4. Seja $T \in \mathcal{S}'(\mathbb{R})$. Então a transformada de Fourier de T , denotada por \hat{T} , é a distribuição temperada definida por $\hat{T}(\varphi) = T(\hat{\varphi})$

Desejamos também estender a transformada de Fourier a todo o $L^2(\mathbb{R})$, para que possamos utilizá-la para as funções Scaling e Wavelet, que estão definidas em $L^2(\mathbb{R})$. Para tal, temos o seguinte teorema, demonstrado em (REED; SIMON, 1980):

Teorema 2.4.2. A transformada de Fourier se estende unicamente a um operador unitário de $L^2(\mathbb{R})$ em $L^2(\mathbb{R})$. A transformada inversa se estende unicamente para seu adjunto.

2.4.1 Propriedades no domínio da Transformada de Fourier

Nas seções anteriores foram introduzidos lemas e teoremas sobre as funções ψ_k e seus coeficientes. Agora, iremos obter resultados relacionados com as transformadas de Fourier das funções ψ_k .

Lema 2.4.3. O conjunto $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal em $L^2(\mathbb{R})$ se e somente se

$$\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma)|^2 = 1 \text{ para todo } \xi \in \mathbb{R}$$

Demonstração. Considere $\gamma_1, \gamma_2 \in \mathbb{Z}$. Pela propriedade do deslocamento, temos que:

$$\mathcal{F}\{\varphi(x - \gamma_1)\}(\xi) = e^{2\pi i \xi \gamma_1} \hat{\varphi}(\xi)$$

$$\mathcal{F}\{\varphi(x - \gamma_2)\}(\xi) = e^{2\pi i \xi \gamma_2} \hat{\varphi}(\xi)$$

Pela Fórmula de Plancherel (2.30) a Transformada de Fourier preserva a norma, e portanto o produto interno. Sendo assim, $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal se e somente se:

$$\langle e^{2\pi i \xi \gamma_1} \hat{\varphi}(\xi), e^{2\pi i \xi \gamma_2} \hat{\varphi}(\xi) \rangle = \delta(\gamma_1 - \gamma_2, 0) \quad (2.31)$$

Denotando $\gamma_1 - \gamma_2 = \gamma$, a equação (2.31) é equivalente a

$$\int e^{2\pi i \xi \gamma} |\hat{\varphi}(\xi)|^2 d\xi = \delta(\gamma, 0)$$

Porém, a integral em \mathbb{R} pode ser quebrada em uma integral sobre o intervalo $[0, 1]$, com o adicional de uma soma em \mathbb{Z} :

$$\int_0^1 \sum_{\gamma' \in \mathbb{Z}} e^{2\pi i (\xi + \gamma') \gamma} |\hat{\varphi}(\xi + \gamma')|^2 d\xi = \delta(\gamma, 0)$$

Devido à periodicidade da função $e^{2\pi i \xi \gamma}$, ela pode ser retirada do somatório, resultando em

$$\int_0^1 e^{2\pi i \xi \gamma} \sum_{\gamma' \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma')|^2 d\xi = \delta(\gamma, 0)$$

Isso significa que os coeficientes de Fourier da função $\sum_{\gamma' \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma')|^2$ no intervalo $[0, 1]$ são $\delta(\gamma, 0)$, o que implica que $\sum_{\gamma' \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma')|^2$ deve ser a função constante 1. ■

Algumas das expressões obtidas nas seções anteriores terão uma forma diferente se considerarmos sua Transformada de Fourier. A identidade escala vista em (2.23), ao

aplicar a Transformada de Fourier resulta em

$$\begin{aligned}
 \hat{\psi}_k(\xi) &= \mathcal{F} \left\{ \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) \varphi(2 \cdot -\gamma) \right\} \\
 &= \left(\sum_{\gamma \in \mathbb{Z}} a_k(\gamma) \right) \mathcal{F} \{ \varphi(2 \cdot -\gamma) \} \\
 &= \left(\frac{1}{2} \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) e^{\pi i \xi \gamma} \right) \hat{\varphi} \left(\frac{1}{2} \xi \right) \\
 &= A_k \left(\frac{1}{2} \xi \right) \hat{\varphi} \left(\frac{1}{2} \xi \right)
 \end{aligned} \tag{2.32}$$

onde

$$A_k(\xi) = \frac{1}{2} \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) e^{2\pi i \xi \gamma} \tag{2.33}$$

Para a geração de Wavelets (e portanto no restante do capítulo) será considerado que existe apenas um conjunto finito de inteiros γ para os quais $a_k(\gamma)$ é diferente de zero, o que justifica a troca entre somatórios e integrais. Além disso, esse número finito de γ garante que A_k é periódico, condição necessária mais à frente. Logo, pela equação (2.33), é possível notar que A_k é suave e periódica. Pela equação (2.25), obtemos que

$$A_0(0) = 1 \tag{2.34}$$

e utilizando a equação (2.19), também temos que

$$\hat{\varphi}(0) = 1 \tag{2.35}$$

Como $\psi_0 = \varphi$, a equação (2.32) pode ser iterada, já que resulta em

$$\hat{\varphi}(\xi) = A_0 \left(\frac{1}{2} \xi \right) \hat{\varphi} \left(\frac{1}{2} \xi \right) \tag{2.36}$$

Substituindo sucessivamente no lado direito da equação $\hat{\varphi} \left(\frac{1}{2^j} \xi \right)$ por $A_0 \left(\frac{1}{2^{j+1}} \xi \right) \hat{\varphi} \left(\frac{1}{2^{j+1}} \xi \right)$, obtemos

$$\hat{\varphi}(\xi) = \left[\prod_{j=1}^k A_0 \left(\frac{1}{2^j} \xi \right) \right] \hat{\varphi} \left(\frac{1}{2^k} \xi \right) \tag{2.37}$$

Com base em (2.37), é de se esperar que valha a equação

$$\hat{\varphi}(\xi) = \prod_{j=1}^{\infty} A_0 \left(\frac{1}{2^j} \xi \right) \tag{2.38}$$

Essa equação é de extrema importância, pois diz que a função $\hat{\varphi}$ pode ser unicamente determinada a partir do coeficiente A_0 . Uma vez provada (2.38), devido a (2.32), podemos afirmar também que:

$$\hat{\psi}_k(\xi) = A_k \left(\frac{1}{2} \xi \right) \prod_{j=2}^{\infty} A_0 \left(\frac{1}{2^j} \xi \right) \quad (2.39)$$

Para provar (2.38), será necessário introduzir o operador $S : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ dado por

$$(Sf)(x) = \sum_{\gamma} a_0(\gamma) f(2x - \gamma)$$

Além disso, considere as funções η_l ($l \in \mathbb{N}$) dadas por

$$\begin{aligned} \eta_0(x) &= \chi_{[-1/2, 1/2)}(x) \\ \eta_l(x) &= (S\eta_{l-1})(x) \end{aligned}$$

e defina

$$\hat{\eta}_{\infty}(\xi) = \prod_{j=1}^{\infty} A_0 \left(\frac{1}{2^j} \xi \right) \quad (2.40)$$

A intenção é mostrar que $\eta_l \rightarrow \varphi$ na norma $L^2(\mathbb{R})$, o que será obtido nos dois próximos lemas e no teorema subsequente, cujas provas encontram-se em (DAUBECHIES, 1988). O próximo lema diz que o produtório em (2.40) converge pontualmente:

Lema 2.4.4. Suponha que, para algum $\epsilon > 0$,

$$\sum_{\gamma} |a_0(\gamma)| |\gamma|^{\epsilon} < \infty$$

Então (2.40) converge pontualmente, para todo $\xi \in \mathbb{R}$. A convergência é uniforme em conjuntos compactos.

Demonstração. Como $\sum_{\gamma} a_0(\gamma) = 2$, temos que

$$A_0(\xi) = 1 + \frac{1}{2} \sum_{\gamma} a(\gamma) (e^{2\pi i \xi} - 1)$$

logo

$$|A_0(\xi) - 1| \leq \sum_{\gamma} |a(\gamma)| |\operatorname{sen}(\pi \xi \gamma)|$$

Como para todo $0 < \sigma \leq 1$ existe $C_\sigma > 0$ tal que, para todo $\alpha \in \mathbb{R}$, $|\text{sen}(\alpha)| \leq C_\sigma |\alpha|^\sigma$, temos que

$$|A_0(\xi) - 1| \leq C_1 \left[\sum_{\gamma} |a(\gamma)| |\gamma|^{\min(1,\epsilon)} \right] |\xi|^{\min(1,\epsilon)}$$

Portanto, temos que

$$|A_0(2^{-j}\xi) - 1| \leq C_2 \lambda^{-j} |\xi|^{\min(1,\epsilon)} \quad (2.41)$$

onde $\lambda = 2^{\min(1,\epsilon)} > 1$. A condição necessária e suficiente para o produtório convergir é que a série

$$\sum_{j=1}^{\infty} \log [A_0(2^{-j}\xi)] \quad (2.42)$$

seja convergente. Escrevendo $A_0(2^{-j}\xi) = k_{\{j,\xi\}} e^{b_{\{j,\xi\}}i}$, com $k_{\{j,\xi\}} > 0$ e $b_{\{j,\xi\}} \in [0, 2\pi)$, o somatório (2.42) pode ser reescrito como

$$\sum_{j=1}^{\infty} (\log |k_{\{j,\xi\}}| + b_{\{j,\xi\}}i)$$

Claramente, a série $\sum_{j=1}^{\infty} \log |k_{\{j,\xi\}}|$ converge, pois, para $j^* \in \mathbb{N}$:

$$\begin{aligned} \sum_{j=j^*}^{\infty} \log |k_{\{j,\xi\}}| &= \sum_{j=j^*}^{\infty} \log |A_0(2^{-j}\xi)| \\ &\leq \sum_{j=j^*}^{\infty} \log (1 + C_2 \lambda^{-j} |\xi|^{\min(1,\epsilon)}) \\ &\leq \sum_{j=j^*}^{\infty} \log e^{C_2 \lambda^{-j} |\xi|^{\min(1,\epsilon)}} \\ &= \sum_{j=j^*}^{\infty} C_2 \lambda^{-j} |\xi|^{\min(1,\epsilon)} \\ &= \frac{C_2 |\xi|^{\min(1,\epsilon)} \lambda^{-j^*+1}}{\lambda - 1} \end{aligned}$$

Considere agora j^* tal que $|b_{\{j,\xi\}}| \leq \frac{1}{2}$. Temos que

$$\begin{aligned} \sum_{j=j^*}^{\infty} b_{\{j,\xi\}}i &\leq \sum_{j=j^*}^{\infty} 2C_2 \lambda^{-j} |\xi|^{\min(1,\epsilon)} i \\ &= \frac{2C_2 |\xi|^{\min(1,\epsilon)} \lambda^{-j^*+1}}{\lambda - 1} i \end{aligned}$$

Portanto, a série (2.42) converge, o que implica que (2.40) converge pontualmente. Além disso, devido a (2.41), o produtório (2.40) converge uniformemente em conjuntos compactos. ■

Uma vez bem definida a função $\hat{\eta}_\infty$, é automático notar que ela satisfaz a equação (2.36) e sendo assim, iremos tomar η_∞ como a função Scaling φ . Resta agora mostrar que a sequência de funções η_l converge para η_∞ . Antes disso, precisaremos mostrar mais um resultado no próximo lema, que irá impor algumas restrições sobre A_0 , como a de que ele seja divisível por $(1 + e^{2\pi i \xi})^N$, para algum $N > 0$, o qual é um dos critérios para garantir a continuidade de η_∞ (DAUBECHIES, 1988).

Lema 2.4.5. Se $A_0(\xi) = \left[\frac{1}{2} (1 + e^{2\pi i \xi})\right]^N F(\xi)$, onde $N \in \mathbb{N}$ e $F(\xi) = \sum_\gamma f(\gamma) e^{2\pi \gamma \xi i}$ satisfaz

$$\sum_\gamma |f(\gamma)| |\gamma|^\epsilon < \infty$$

para algum $\epsilon > 0$, e

$$\sup_{\xi \in \mathbb{R}} |F(\xi)| = B$$

então existe $C > 0$ tal que, para todo $\xi \in \mathbb{R}$,

$$\left| \prod_{j=1}^{\infty} A_0\left(\frac{1}{2^j} \xi\right) \right| \leq C (1 + |\xi|)^{-N + (\log B)/(\log 2)}$$

Demonstração. Como $\prod_{j=1}^{\infty} \cos(2^{-j} x) = x^{-1} \text{sen}(x)$, temos que

$$\begin{aligned} \prod_{j=1}^{\infty} A_0\left(\frac{1}{2^j} \xi\right) &= \prod_{j=1}^{\infty} \left[\frac{1}{2} (1 + e^{2\pi i 2^{-j} \xi}) \right]^N F(2^{-j} \xi) \\ &= \prod_{j=1}^{\infty} \left[\frac{e^{\pi i 2^{-j} \xi}}{2} (e^{-\pi i 2^{-j} \xi} + e^{+\pi i 2^{-j} \xi}) \right]^N F(2^{-j} \xi) \\ &= e^{N\pi \xi i} \left[\prod_{j=1}^{\infty} \cos(\pi 2^{-j} \xi) \right]^N \prod_{j=1}^{\infty} F(2^{-j} \xi) \\ &= e^{N\pi \xi i} \left[\frac{\text{sen}(\pi \xi)}{\pi \xi} \right]^N \prod_{j=1}^{\infty} F(2^{-j} \xi) \end{aligned} \tag{2.43}$$

Devido ao lema 2.4.4, o lado direito da equação (2.43) converge uniformemente em compactos. Sendo assim, existe $C_1 > 0$ tal que, para todo $|\xi| \leq 1$:

$$\left| \prod_{j=1}^{\infty} A_0\left(\frac{1}{2^j} \xi\right) \right| \leq C_1 \tag{2.44}$$

Agora considere $|\xi| > 1$ e determine um $j_0 \in \mathbb{N}$ tal que

$$2^{-j_0} |\xi| < 1 \leq 2^{-j_0+1} |\xi|$$

ou seja, tal que

$$\log |\xi| / \log 2 < j_0 \leq 1 + \log |\xi| / \log 2$$

Então, reescrevendo o produtório de $F(\xi)$ em (2.43):

$$\begin{aligned} \left| \prod_{j=1}^{\infty} F(2^{-j}\xi) \right| &= \left| \prod_{j=1}^{j_0} F(2^{-j}\xi) \right| \left| \prod_{j=1}^{\infty} F(2^{-j}2^{-j_0}\xi) \right| \\ &\leq B^{j_0} \left| \prod_{j=1}^{\infty} F(2^{-j}2^{-j_0}\xi) \right| \end{aligned}$$

Note que como $F(0) = A_0(0)$, então $\sum_{\gamma} f(\gamma) = 1$, de modo que podemos realizar o mesmo desenvolvimento feito no lema 2.4.4, notando que $2^{-j_0}|\xi| < 1$, e logo obter que

$$\begin{aligned} \left| \prod_{j=1}^{\infty} F(2^{-j}\xi) \right| &\leq B^{j_0} \prod_{j=1}^{\infty} \left(1 + C^* |2^{-j_0}\xi|^{\min(1,\epsilon)} |2^{-j\epsilon}| \sum_{\gamma} |f(\gamma)| |\gamma|^{\epsilon} \right) \\ &\leq B^{j_0} \prod_{j=1}^{\infty} \left(1 + C^* |2^{-j\epsilon}| \sum_{\gamma} |f(\gamma)| |\gamma|^{\epsilon} \right) \\ &\leq CB^{j_0} \\ &= Ce^{j_0 \log B} \end{aligned}$$

Agora, para $B < 1$ temos

$$\begin{aligned} Ce^{j_0 \log B} &\leq Ce^{\log |\xi| \log B / \log 2} \\ &= C|\xi|^{\log B / \log 2} \end{aligned}$$

e para $B > 1$ temos

$$\begin{aligned} Ce^{j_0 \log B} &\leq Ce^{(1 + \log |\xi| / \log 2) \log B} \\ &\leq Ce^{\log B} |\xi|^{\log B / \log 2} \\ &= C' |\xi|^{\log B / \log 2} \end{aligned}$$

de modo que podemos escrever em geral para $|\xi| > 1$ e $B > 0$, existe algum $C > 0$ tal que

$$\left| \prod_{j=1}^{\infty} F(2^{-j}\xi) \right| \leq C |\xi|^{\log B / \log 2} \leq C(1 + |\xi|)^{\log B / \log 2}$$

Agora, note que para $|\xi| \leq 1$ existe $C > 0$ tal que

$$\left| \prod_{j=1}^{\infty} A_0 \left(\frac{1}{2^j} \xi \right) \right| \leq C(1 + |\xi|)^{-N+(\log B)/(\log 2)}$$

já que $(1 + |\xi|)^{-N+(\log B)/(\log 2)}$ é limitado inferiormente por uma constante positiva não nula quando ξ varia nesse intervalo. Já para $|\xi| > 1$, note que

$$|\xi|^{-N} \leq \left(\frac{1 + |\xi|}{2} \right)^{-N}$$

de modo que existe $C > 0$ tal que, para $|\xi| > 1$, vale

$$\begin{aligned} \left| \prod_{j=1}^{\infty} A_0 \left(\frac{1}{2^j} \xi \right) \right| &\leq \left| \frac{\text{sen}(\pi\xi)}{\pi\xi} \right|^N \left| \prod_{j=1}^{\infty} F(2^{-j}\xi) \right| \\ &\leq C_2 |\xi|^{-N} (1 + |\xi|)^{\log B / \log 2} \\ &\leq C(1 + |\xi|)^{-N+\log B / \log 2} \end{aligned}$$

Portanto, existe $C > 0$ tal que essa inequação é válida para todo $\xi \in \mathbb{R}$. ■

Agora, é possível mostrar que, de fato, $\eta_l \rightarrow \eta_\infty$ na norma $L^2(\mathbb{R})$. Isso será importante na próxima seção pois definirá uma situação na qual podemos dizer que uma φ tem suporte compacto.

Teorema 2.4.6. Seja $A_0(\xi)$ como em (2.33), sendo $\sum_\gamma a_0 = 2$. Suponha ainda que

$$A_0(\xi) = \left[\frac{1}{2} \left(1 + e^{2\pi i \xi} \right) \right]^N F(\xi)$$

para algum $N \in \mathbb{N}$ e $F(\xi) = \sum_\gamma f(\gamma) e^{2\pi i \gamma \xi}$ tal que

$$\sum_\gamma |f(\gamma)| |\gamma|^\epsilon < \infty$$

para algum $\epsilon > 0$ e ainda

$$\sup_{\xi \in \mathbb{R}} |F(\xi)| = B < 2^{N-\frac{1}{2}}$$

Então a função constante por partes η_l , definida recursivamente por

$$\eta_l(x) = \sum_\gamma a(\gamma) \eta_{l-1}(2x - \gamma)$$

com

$$\eta_0(x) = \chi_{[-1/2, 1/2)}(x)$$

converge pontualmente para a função contínua definida por

$$\eta_\infty(x) = \prod_{j=1}^{\infty} A_0(2^{-j}\xi)$$

Demonstração. Tomando a transformada de Fourier de η_l obtemos:

$$\hat{\eta}_l(\xi) = \left[\prod_{j=1}^l A_0(2^{-j}\xi) \right] \left[\frac{\text{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right]$$

Pelo lema 2.4.4, $\hat{\eta}_l$ converge uniformemente para $\hat{\eta}_\infty$ em conjuntos compactos. Isso significa que, para todo $\delta > 0$ e para todo $R > 0$, é possível achar l_0 tal que, para $l \geq l_0$

$$\int_{|\xi| \leq R} |\hat{\eta}_l(\xi) - \hat{\eta}_\infty(\xi)|^2 d\xi \leq \delta$$

Por outro lado, como $B < 2^{N-\frac{1}{2}}$, temos que $\hat{\eta}_\infty \in L^2(\mathbb{R})$, já que utilizando o lema 2.4.5

$$\begin{aligned} \int_{|\xi| > 1} |\hat{\eta}_\infty|^2 d\xi &\leq \int_{|\xi| > 1} \left| \frac{\text{sen}(\pi\xi)}{\pi\xi} \right|^{2N} \left| \prod_{j=1}^{\infty} F(2^{-j}\xi) \right|^2 d\xi \\ &\leq \int_{|\xi| > 1} C |\xi|^{-2N} |\xi|^{2 \log B / \log 2} d\xi \\ &\leq 2C \int_1^{\infty} |\xi|^\theta d\xi \\ &\leq C_2 \end{aligned} \tag{2.45}$$

sendo que $\theta = -2N + 2 \log B / \log 2 < -1$ e portanto a integral converge para uma constante C_2 . Portanto, para todo $\delta > 0$ existe $R > 0$ tal que

$$\int_{|\xi| \geq R} |\hat{\eta}_\infty|^2 d\xi \leq \delta$$

Sendo assim, como

$$\int_{|\xi| \geq R} |\hat{\eta}_l(\xi) - \hat{\eta}_\infty(\xi)|^2 d\xi \leq \int_{|\xi| \geq R} |\hat{\eta}_l(\xi)|^2 d\xi + \int_{|\xi| \geq R} |\hat{\eta}_\infty(\xi)|^2 d\xi \tag{2.46}$$

basta agora provar que, para todo $\delta > 0$, existe $R > 0$ e l_0 grandes o bastante tais que, para $l \geq l_0$

$$\int_{|\xi| \geq R} |\hat{\eta}_l(\xi)|^2 d\xi \leq \delta$$

Defina a função P_l como

$$P_l(\xi) = \prod_{j=1}^l A_0(2^{-j}\xi)$$

Então precisamos calcular a integral

$$\int_{|\xi| \geq R} |P_l(\xi)|^2 \left| \frac{\text{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right|^2 d\xi \quad (2.47)$$

Para tal, iremos separá-la em duas partes: uma para $|\xi| \geq 2^{l-1}$ e outra para $R \leq |\xi| \leq 2^{l-1}$. É também importante ressaltar algumas propriedades de P_l :

$$|P_l(\xi)| \leq 1, \text{ pois } |A_0(\xi)| \leq 1 \quad (2.48)$$

Devido ao lema 2.4.5, obtemos outra propriedade:

$$\begin{aligned} |P_l(\xi)| &\leq \left| \prod_{j=1}^l \cos(\pi 2^{-j}\xi) \right|^N \prod_{j=1}^l |F(2^{-j}\xi)| \\ &\leq C \left| \frac{2^{-l}\text{sen}(\pi\xi)}{\text{sen}(2^{-l}\pi\xi)} \right|^N \prod_{j=1}^l |F(2^{-j}\xi)| \\ &\leq C \left| \frac{2^{-l}\text{sen}(\pi\xi)}{\text{sen}(2^{-l}\pi\xi)} \right|^N (1 + |\xi|)^\beta \end{aligned} \quad (2.49)$$

sendo $\beta = \log B / \log 2$. Além disso, por (2.33), segue que P_l é periódica, com período 2^l .

Primeiramente, calcularemos a integral para $|\xi| \geq 2^{l-1}$. Utilizando a periodicidade de P_l , temos que:

$$\begin{aligned} \int_{|\xi| \geq 2^{l-1}} |P_l(\xi)|^2 \left| \frac{\text{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right|^2 d\xi &= \sum_{k \neq 0} \int_{|\xi| \leq 2^{l-1}} |P_l(\xi + 2^l k)|^2 \left| \frac{\text{sen}[2^{-l}\pi(\xi + 2^l k)]}{2^{-l}\pi(\xi + 2^l k)} \right|^2 d\xi \\ &= \sum_{k \neq 0} \int_{|\xi| \leq 2^{l-1}} |P_l(\xi)|^2 \left| \frac{\text{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi + \pi k} \right|^2 d\xi \\ &\leq \sum_{k \neq 0} \left[\pi \left(k - \frac{1}{2} \right) \right]^{-2} \int_{|\xi| \leq 2^{l-1}} |P_l(\xi)|^2 \left| \text{sen}(2^{-l}\pi\xi) \right|^2 d\xi \\ &\leq C \int_{|\xi| \leq 2^{l-1}} |P_l(\xi)|^2 \left| \text{sen}(2^{-l}\pi\xi) \right|^2 d\xi \end{aligned}$$

Considere agora $\lambda = 2^{-\alpha(l-1)}$, sendo que $\alpha \in (0, 1)$ será fixado posteriormente. Então

$$\begin{aligned}
& \int_{|\xi| \leq 2^{l-1}} |P_l(\xi)|^2 |\text{sen}(2^{-l}\pi\xi)|^2 d\xi \\
& \leq \int_{|\xi| \leq 2^{l-1}\lambda} |P_l(\xi)|^2 |\text{sen}(2^{-l}\pi\xi)|^2 d\xi + \int_{2^{l-1}\lambda \leq |\xi| \leq 2^{l-1}} |P_l(\xi)|^2 |\text{sen}(2^{-l}\pi\xi)|^2 d\xi \\
& \leq \frac{\pi^2 \lambda^2}{4} \int_{|\xi| \leq 2^{l-1}\lambda} |P_l(\xi)|^2 d\xi + \int_{2^{l-1}\lambda \leq |\xi| \leq 2^{l-1}} |P_l(\xi)|^2 d\xi
\end{aligned} \tag{2.50}$$

Agora, calculando a integral do primeiro termo em (2.50), temos, utilizando a propriedade vista em (2.49):

$$\begin{aligned}
\int_{|\xi| \leq 2^{l-1}\lambda} |P_l(\xi)|^2 d\xi & \leq \int_{|\xi| \leq 1} |P_l(\xi)|^2 d\xi + C \int_{1 \leq |\xi| \leq 2^{l-1}\lambda} (1 + |\xi|)^{2\beta} |2^l \text{sen}(2^{-l}\pi\xi)|^{-2N} d\xi \\
& \leq 1 + C \int_{1 \leq |\xi| \leq 2^{l-1}\lambda} (1 + |\xi|)^{2\beta} \left| \frac{\text{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right|^{-2N} |\pi\xi|^{-2N} d\xi \\
& \leq 1 + C |\pi|^{-2N} \int_{1 \leq |\xi| \leq 2^{l-1}\lambda} (1 + |\xi|)^{2\beta} |\xi|^{-2N} d\xi \\
& \leq 1 + C |\pi|^{-2N} \int_1^\infty (1 + x)^{2\beta} x^{-2N} dx \\
& \leq C_p
\end{aligned}$$

sendo que C_p é finito pois $2(\beta - N) < -1$. Calculando a integral do segundo termo em (2.50), considerando que $|\text{sen}(x)| \geq 2|x|/\pi$ para $|x| \leq \frac{\pi}{2}$, obtemos

$$\begin{aligned}
\int_{2^{l-1}\lambda \leq |\xi| \leq 2^{l-1}} |P_l(\xi)|^2 d\xi & \leq C \int_{2^{l-1}\lambda \leq |\xi| \leq 2^{l-1}} (1 + |\xi|)^{2\beta} |2^l \text{sen}(2^{-l}\pi\xi)|^{-2N} d\xi \\
& \leq 2^{-2Nl+1} 2^{l-1} C \int_\lambda^1 (1 + 2^{l-1}x)^{2\beta} \left| \text{sen}\left(\frac{\pi x}{2}\right) \right|^{-2N} dx \\
& \leq 2^{-2Nl+1} (1 + 2^{l-1})^{2\beta} 2^{l-1} C \int_\lambda^1 \left| \text{sen}\left(\frac{\pi x}{2}\right) \right|^{-2N} dx \\
& \leq 2^{-2Nl+1} (1 + 2^{l-1})^{2\beta} 2^{l-1} C \int_\lambda^1 \left| \frac{\pi x}{2} \right|^{-2N} dx \\
& \leq 2^{l(1+2\beta-2N)} C_2 \int_\lambda^1 |x|^{-2N} dx \\
& \leq C_2 2^{l(1+2\beta-2N)} \lambda^{-2N}
\end{aligned}$$

Substituindo os dois últimos resultados em (2.50) obtemos:

$$\begin{aligned}
\int_{|\xi| \leq 2^{l-1}} |P_l(\xi)|^2 |\text{sen}(2^{-l}\pi\xi)|^2 d\xi & \leq C_3 \lambda^2 + C_2 2^{l(1+2\beta-2N)} \lambda^{-2N} \\
& = C_3 2^{-2\alpha(l-1)} + C_2 2^{l(1+2\beta-2N)} 2^{2\alpha N(l-1)} \\
& = C_3 2^{-2\alpha(l-1)} + C_4 2^{l(1+2\beta-2N+2\alpha N)}
\end{aligned}$$

Escolhendo α como sendo

$$\alpha = \frac{2N - 2\beta - 1}{2N + 2}$$

implica que (2.50) tende a zero quando $l \rightarrow \infty$.

Por fim, calculando a integral (2.47) para $R \leq |\xi| \leq 2^{l-1}$, considerando novamente que $|\operatorname{sen}(x)| \geq 2|x|/\pi$ para $|x| \leq \frac{\pi}{2}$ e utilizando a propriedade vista em (2.49) temos

$$\begin{aligned} & \int_{R \leq |\xi| \leq 2^{l-1}} |P_l(\xi)|^2 \left| \frac{\operatorname{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right|^2 d\xi \\ & \leq \int_{R \leq |\xi| \leq 2^{l-1}} C \left| \frac{2^{-l}\operatorname{sen}(\pi\xi)}{\operatorname{sen}(2^{-l}\pi\xi)} \right|^{2N} (1 + |\xi|)^{2\beta} \left| \frac{\operatorname{sen}(2^{-l}\pi\xi)}{2^{-l}\pi\xi} \right|^2 d\xi \\ & \leq \int_{R \leq |\xi| \leq 2^{l-1}} C \left| \frac{2^{-l}\operatorname{sen}(\pi\xi)}{\operatorname{sen}(2^{-l}\pi\xi)} \right|^{2N} (1 + |\xi|)^{2\beta} d\xi \\ & \leq C_2 \int_{R \leq |\xi| \leq 2^{l-1}} |\xi|^{-2N} (1 + |\xi|)^{2\beta} d\xi \\ & \leq 2C_2 \int_R^\infty x^{-2N} (1 + x)^{2\beta} d\xi \end{aligned}$$

Como $2\beta - 2N < -1$, a integral tende a zero quando $R \rightarrow \infty$, uniformemente em l . Isso mostra que, para todo $\delta > 0$, escolhendo l e R suficientemente grandes, vale que

$$\int_{|\xi| \geq R} |\hat{\eta}_l(\xi)|^2 d\xi \leq \delta$$

e portanto, por (2.46), $\|\hat{\eta}_l - \hat{\eta}_\infty\|_2 \rightarrow 0$ quando $l \rightarrow \infty$. ■

Note que, com uma pequena modificação no teorema, fazendo $B < 2^{N-1}$, conseguimos garantir que a função η_∞ seja limitada pois, como $\eta_\infty(x) = \int e^{-2\pi x\xi} \hat{\eta}_\infty(\xi) d\xi$ e de acordo com a equação (2.45), temos:

$$\begin{aligned}
|\eta_\infty(x)| &= \left| \int e^{-2\pi x \xi i} \hat{\eta}_\infty(\xi) \, d\xi \right| \\
&\leq \int |e^{-2\pi x \xi i} \hat{\eta}_\infty(\xi)| \, d\xi \\
&\leq \int_{|\xi|>1} |\hat{\eta}_\infty(\xi)| \, d\xi \\
&\leq \int_{|\xi|>1} \left| \frac{\text{sen}(\pi \xi)}{\pi \xi} \right|^N \left| \prod_{j=1}^{\infty} F(2^{-j} \xi) \right| \, d\xi \\
&\leq \int_{|\xi|>1} C |\xi|^{-N} |\xi|^{\log B / \log 2} \, d\xi \\
&\leq 2C \int_1^{\infty} |\xi|^{\theta_2} \, d\xi \\
&\leq C_4
\end{aligned}$$

onde $\theta_2 = -N + \log B / \log 2 < -1$ e portanto a integral converge.

Uma vez bem definida $\hat{\eta}_\infty$, precisaremos de mais algumas informações sobre os coeficientes a_k e A_k . Uma delas é a generalização do lema 2.4.3, que será feita com o intuito de facilitar provas subsequentes:

Lema 2.4.7. Considerando válidas (2.24) e (2.25), ou seja:

$$\sum_{\gamma' \in \mathbb{Z}} a_j(\gamma') \overline{a_k(2\gamma + \gamma')} = 2\delta(j, k)\delta(\gamma, 0) \quad (2.51)$$

$$\sum_{\gamma \in \mathbb{Z}} a_0(\gamma) = 2 \quad (2.52)$$

então pelo lema 2.3.1

$$\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$$

é uma base ortonormal e

$$\sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma) \overline{\hat{\psi}_j(\xi + \gamma)} = \delta(j, k) \quad (2.53)$$

Demonstração. A prova é análoga à do lema 2.4.3. Pela ortonormalidade, pela propriedade do deslocamento e pela preservação da norma através da Fórmula de Plancherel, vale que

$$\langle e^{2\pi i \xi \gamma_1} \hat{\psi}_k(\xi), e^{2\pi i \xi \gamma_2} \hat{\psi}_j(\xi) \rangle = \delta(\gamma_1 - \gamma_2, 0) \delta(j, k)$$

Repetindo os passos feitos no lema 2.4.3, sendo $\gamma_1 - \gamma_2 = \gamma$, quebrando a equação em \mathbb{R} obtemos:

$$\begin{aligned}
\langle e^{2\pi i \xi \gamma_1} \hat{\psi}_k(\xi), e^{2\pi i \xi \gamma_2} \hat{\psi}_j(\xi) \rangle &= \int e^{2\pi i \xi \gamma} \hat{\psi}_k(\xi) \overline{\hat{\psi}_j(\xi)} d\xi \\
&= \int_0^1 e^{2\pi i \xi \gamma} \sum_{\gamma' \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma') \overline{\hat{\psi}_j(\xi + \gamma')} d\xi \\
&= \delta(\gamma, 0) \delta(j, k)
\end{aligned}$$

Novamente, os coeficientes de Fourier da função $\sum_{\gamma' \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma') \overline{\hat{\psi}_j(\xi + \gamma')}$ no intervalo $[0, 1]$ são $\delta(\gamma, 0) \delta(j, k)$, o que significa que $\sum_{\gamma' \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma') \overline{\hat{\psi}_j(\xi + \gamma')}$ deve ser a função delta de Kronecker $\delta(j, k)$. ■

Agora considere $\eta_1 = 0$ e $\eta_2 = 1/2$. Então cada elemento de \mathbb{Z} pode ser escrito na forma $2(\gamma + \eta_p)$, com γ variando em \mathbb{Z} e $p \in \{1, 2\}$. Substituindo essa notação em (2.53):

$$\sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma) \overline{\hat{\psi}_j(\xi + \gamma)} = \sum_{p=1}^2 \sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + 2(\gamma + \eta_p)) \overline{\hat{\psi}_j(\xi + 2(\gamma + \eta_p))}$$

Substituindo (2.32) obtemos

$$\sum_{p=1}^2 \sum_{\gamma \in \mathbb{Z}} A_k \left(\frac{1}{2} \xi + \gamma + \eta_p \right) \overline{A_j \left(\frac{1}{2} \xi + \gamma + \eta_p \right)} \left| \hat{\varphi} \left(\frac{1}{2} \xi + \gamma + \eta_p \right) \right|^2$$

Considerando que A_k é periódico, temos

$$\sum_{p=1}^2 A_k \left(\frac{1}{2} \xi + \eta_p \right) \overline{A_j \left(\frac{1}{2} \xi + \eta_p \right)} \sum_{\gamma \in \mathbb{Z}} \left| \hat{\varphi} \left(\frac{1}{2} \xi + \gamma + \eta_p \right) \right|^2$$

Pelo lema 2.4.3, o somatório mais interno vale 1 e pela equação (2.53) temos

$$\sum_{p=1}^2 A_k(\xi + \eta_p) \overline{A_j(\xi + \eta_p)} = \delta(j, k) \tag{2.54}$$

que representa a Transformada de Fourier da condição (2.24).

Além disso, se fizermos $j = k$ em (2.54), obtemos

$$|A_k(\xi)|^2 + \left| A_k \left(\xi + \frac{1}{2} \right) \right|^2 = 1 \text{ para todo } \xi \in \mathbb{R} \tag{2.55}$$

o que implica em

$$|A_k(\xi)| \leq 1 \text{ para todo } \xi \in \mathbb{R} \tag{2.56}$$

Sendo assim, em vista da equação (2.39), a equação (2.56) diz que a função $\hat{\psi}_k$ é limitada.

Com a teoria desenvolvida até aqui, já é possível provar o lema 2.3.2:

Lema 2.3.2: A equação

$$\sum_{k=0,1} \sum_{\gamma'' \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} + 2\gamma''')} a_k(2\gamma''' + \gamma'') = 2\delta(\tilde{\gamma}, \gamma'') \quad (2.57)$$

sempre vale, portanto $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$ é uma base ortonormal para V_1 .

Demonstração. Considere a matriz M_A dada por

$$M_A = \begin{pmatrix} A_0(\xi + \eta_1) & A_0(\xi + \eta_2) \\ A_1(\xi + \eta_1) & A_1(\xi + \eta_2) \end{pmatrix}$$

A equação (2.54) diz que a matriz M_A é unitária, pois:

$$\begin{aligned} M_A M_A^* &= \begin{pmatrix} A_0(\xi + \eta_1) & A_0(\xi + \eta_2) \\ A_1(\xi + \eta_1) & A_1(\xi + \eta_2) \end{pmatrix} \begin{pmatrix} \overline{A_0(\xi + \eta_1)} & \overline{A_1(\xi + \eta_1)} \\ \overline{A_0(\xi + \eta_2)} & \overline{A_1(\xi + \eta_2)} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{p=1}^2 |A_0(\xi + \eta_p)|^2 & \sum_{p=1}^2 A_1(\xi + \eta_p) \overline{A_0(\xi + \eta_p)} \\ \sum_{p=1}^2 A_0(\xi + \eta_p) \overline{A_1(\xi + \eta_p)} & \sum_{p=1}^2 |A_1(\xi + \eta_p)|^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Isso também significa que $M_A^* M_A = I$, o que resulta na seguinte condição:

$$\sum_{k=0}^1 A_k(\xi + \eta_p) \overline{A_k(\xi + \eta_q)} = \delta(p, q) \quad (2.58)$$

Substituindo (2.33) em (2.58):

$$\begin{aligned} &\sum_{k=0}^1 \frac{1}{4} \left(\sum_{\gamma \in \mathbb{Z}} a_k(\gamma) e^{2\pi i(\xi + \eta_p)\gamma} \right) \left(\sum_{\gamma' \in \mathbb{Z}} \overline{a_k(\gamma')} e^{-2\pi i(\xi + \eta_q)\gamma'} \right) \\ &= \sum_{\gamma \in \mathbb{Z}} \left(\frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma) \overline{a_k(\gamma')} e^{-2\pi i(\xi + \eta_q)\gamma'} e^{2\pi i(\xi + \eta_p)\gamma} \right) \\ &\stackrel{\gamma \rightarrow \gamma + \gamma'}{=} \sum_{\gamma \in \mathbb{Z}} \left(\frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{-2\pi i(\xi + \eta_q)\gamma'} e^{2\pi i(\xi + \eta_p)(\gamma + \gamma')} \right) \quad (2.59) \\ &= \sum_{\gamma \in \mathbb{Z}} \left(\frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{2\pi i\gamma\eta_p} e^{2\pi i\gamma'(\eta_p - \eta_q)} \right) e^{2\pi i\gamma\xi} \\ &= \delta(p, q) \end{aligned}$$

Considerando a equação (2.59) como uma expansão em Série de Fourier, significa que a Série de Fourier do termo entre parênteses é constante e igual a 1 quando $p = q$. Mas a função cuja Série de Fourier é uma constante é a $\delta(\gamma, 0)$. Portanto:

$$\frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{2\pi i \gamma \eta_p} e^{2\pi i \gamma' (\eta_p - \eta_q)} = \delta(p, q) \delta(\gamma, 0) \quad (2.60)$$

Considerando $\eta_p = 0$ e somando em relação a q :

$$\sum_{q=1}^2 \frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{-2\pi i \gamma' \eta_q} = \frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} \sum_{q=1}^2 e^{-2\pi i \gamma' \eta_q} \quad (2.61)$$

Com isso temos

$$\sum_{q=1}^2 e^{-2\pi i \gamma' \eta_q} = \begin{cases} 2 & \text{se } \gamma' \in 2\mathbb{Z} \\ 0 & \text{caso contrário} \end{cases} \quad (2.62)$$

Substituindo (2.62) em (2.61):

$$\frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} \sum_{q=1}^2 e^{-2\pi i \gamma' \eta_q} \stackrel{\gamma' \rightarrow 2\gamma'}{=} \frac{1}{2} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(2\gamma' + \gamma) \overline{a_k(2\gamma')}$$

Portanto

$$\sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(2\gamma' + \gamma) \overline{a_k(2\gamma')} = 2\delta(\gamma, 0)$$

que é a equação (2.57) para $\tilde{\gamma} = 0$.

Para $\tilde{\gamma} = 1$, considere $\eta_p = 1/2$. Multiplicando os dois lados de (2.60) por $e^{2\pi i \eta_q}$ e somando em relação a q :

$$\begin{aligned} \sum_{q=1}^2 \frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{2\pi i \gamma \eta_p} e^{2\pi i \gamma' (\eta_p - \eta_q)} e^{2\pi i \eta_q} &= \sum_{q=1}^2 \delta(p, q) \delta(\gamma, 0) e^{2\pi i \eta_q} \\ \sum_{q=1}^2 \frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma') \overline{a_k(\gamma')} e^{\pi i (\gamma' + \gamma)} e^{2\pi i \eta_q (1 - \gamma')} &= -\delta(\gamma, 0) \\ \stackrel{\gamma' \rightarrow \gamma' + 1}{=} \frac{1}{4} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + \gamma' + 1) \overline{a_k(\gamma' + 1)} e^{\pi i (\gamma' + \gamma + 1)} \sum_{q=1}^2 e^{-2\pi i \eta_q \gamma'} &= -\delta(\gamma, 0) \end{aligned}$$

Novamente, utilizando (2.62) e fazendo a mudança de variáveis de γ' para $2\gamma'$:

$$\frac{1}{2} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + 2\gamma' + 1) \overline{a_k(2\gamma' + 1)} e^{\pi i (2\gamma' + \gamma + 1)} = -\delta(\gamma, 0)$$

Fazendo a mudança de variáveis de $\gamma - 1$ para γ :

$$\frac{1}{2} \sum_{k=0}^1 \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma + 2\gamma') \overline{a_k(2\gamma' + 1)} e^{\pi i(2\gamma' + \gamma)} = -\delta(\gamma, 1)$$

Para $\gamma = 1$, a expressão $e^{\pi i(2\gamma' + \gamma)}$ é igual a -1 , resultando na equação (2.57) para $\tilde{\gamma} = 1$. Para qualquer valor de $\tilde{\gamma}$ diferente de 0 ou 1 basta fazer uma mudança de variáveis em relação a γ''' para recair em um desses dois casos. Sendo assim, provamos que a equação (2.57) sempre vale. ■

2.4.2 Ortonormalidade do conjunto $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$

Os lemas que demonstravam a ortonormalidade do conjunto $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}; k \in \{0,1\}}$ partiam da suposição de que $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ era ortonormal. No entanto, para provar a ortonormalidade do conjunto $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ diretamente, será necessário estabelecer uma condição sobre os coeficientes A_0 .

Teorema 2.4.8. Suponha que

$$A_0(\xi) \neq 0 \text{ para } |\xi| \leq \frac{1}{4} \quad (2.63)$$

Então $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal.

Demonstração. A prova consistirá na construção de funções φ_j tais que $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é ortonormal e $\varphi_j \rightarrow \varphi$ na norma L^2 quando $j \rightarrow \infty$. Isso basta para provar o teorema pois

$$\langle \lim_{j \rightarrow \infty} \varphi_j(x - \gamma_1), \lim_{j \rightarrow \infty} \varphi_j(x - \gamma_2) \rangle = \lim_{j \rightarrow \infty} \langle \varphi_j(x - \gamma_1), \varphi_j(x - \gamma_2) \rangle$$

devido à continuidade do produto interno. Sendo assim, se cada uma das φ_j gera uma base ortonormal $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$, então a φ também gera uma base ortonormal $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$, já que $\lim_{j \rightarrow \infty} \varphi_j = \varphi$.

Para a função inicial φ_0 , considere $\hat{\varphi}_0(\xi) = \chi_{[-\frac{1}{2}, \frac{1}{2})}(\xi)$. Então, pelo lema 2.4.3, $\{\varphi_0(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é ortonormal. As funções φ_j serão definidas indutivamente por meio de suas transformadas de Fourier $\hat{\varphi}_j$, utilizando o coeficiente A_0 (conforme a equação (2.36)) de $\hat{\varphi}$:

$$\hat{\varphi}_j(\xi) = A_0\left(\frac{1}{2}\xi\right) \hat{\varphi}_{j-1}\left(\frac{1}{2}\xi\right) \quad (2.64)$$

O conjunto $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ também é ortonormal. De fato, suponha por hipótese de indução que o conjunto $\{\varphi_{j-1}(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ seja ortonormal. Considere a equação obtida a partir de (2.64):

$$\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}_j(\xi + \gamma)|^2 = \sum_{\gamma \in \mathbb{Z}} \left| A_0 \left(\frac{1}{2}(\xi + \gamma) \right) \right|^2 \left| \hat{\varphi}_{j-1} \left(\frac{1}{2}(\xi + \gamma) \right) \right|^2$$

Desmenbrando o somatório:

$$\sum_{\gamma \in \mathbb{Z}} \left| A_0 \left(\frac{1}{2}\xi + \gamma \right) \right|^2 \left| \hat{\varphi}_{j-1} \left(\frac{1}{2}\xi + \gamma \right) \right|^2 + \sum_{\gamma' \in \mathbb{Z}} \left| A_0 \left(\frac{1}{2}\xi + \gamma' - \frac{1}{2} \right) \right|^2 \left| \hat{\varphi}_{j-1} \left(\frac{1}{2}\xi + \gamma' - \frac{1}{2} \right) \right|^2$$

Segue de (2.55) que:

$$|A_k(\xi)|^2 = |A_k(\xi + 1)|^2$$

Colocando os termos A_0 em evidência:

$$\left| A_0 \left(\frac{1}{2}\xi \right) \right|^2 \sum_{\gamma \in \mathbb{Z}} \left| \hat{\varphi}_{j-1} \left(\frac{1}{2}\xi + \gamma \right) \right|^2 + \left| A_0 \left(\frac{1}{2}\xi + -\frac{1}{2} \right) \right|^2 \sum_{\gamma' \in \mathbb{Z}} \left| \hat{\varphi}_{j-1} \left(\frac{1}{2}\xi + \gamma' - \frac{1}{2} \right) \right|^2$$

Aplicando a condição de ortonormalidade em $\{\varphi_{j-1}(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ vista no lema 2.4.3 e a equação (2.55) em seguida:

$$\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}_j(\xi + \gamma)|^2 = \left| A_0 \left(\frac{1}{2}\xi \right) \right|^2 + \left| A_0 \left(\frac{1}{2}\xi + -\frac{1}{2} \right) \right|^2 = 1$$

Portanto, segue que

$$\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}_j(\xi + \gamma)|^2 = 1$$

Logo, pelo lema 2.4.3, $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal.

A função $\hat{\varphi}_j$, devido a (2.64) também pode ser escrita como

$$\hat{\varphi}_j(\xi) = \left[\prod_{k=1}^j A_0(2^{-k}\xi) \right] \chi_{[-2^{j-1}, 2^j]}(\xi) \quad (2.65)$$

forma na qual fica explícito que $\hat{\varphi}_j \rightarrow \hat{\varphi}$ pontualmente, já que φ pode ser escrita na forma da equação (2.38). Agora, basta mostrar que $\varphi_j \rightarrow \varphi$ na norma $L^2(\mathbb{R})$. Pela Fórmula de Plancherel (2.30), é suficiente mostrar que $\hat{\varphi}_j \rightarrow \hat{\varphi}$ na norma $L^2(\mathbb{R})$. Para tal, serão utilizados dois teoremas da literatura:

Teorema 2.4.9. Lema de Fatou: (ROYDEN; FITZPATRICK, 2010) Seja $\{f_n\}$ uma sequência de funções não-negativas em um espaço de medida (E, Σ, μ) . Se $\{f_n\} \rightarrow f$ pontualmente em quase todo o espaço E , então

$$\int_E f \, d\mu \leq \liminf_{n \rightarrow \infty} \int_E f_n \, d\mu$$

Teorema 2.4.10. Teorema da Convergência Dominada: (ROYDEN; FITZPATRICK, 2010) Seja $\{f_n\}$ uma sequência de funções mensuráveis em um espaço de medida (E, Σ, μ) . Suponha que exista uma função g que é integrável em E e domina $\{f_n\}$ em E , ou seja, $|f_n| < g$ em E para todo n . Se $\{f_n\}$ converge a f pontualmente μ q.t.p., então f é integrável em E e

$$\lim_{n \rightarrow \infty} \int_E |f_n - f| \, d\mu = 0$$

o que implica em

$$\lim_{n \rightarrow \infty} \int_E f_n \, d\mu = \int_E f \, d\mu$$

Pelo lema de Fatou, considerando o espaço $L^2(\mathbb{R})$ e as funções $|\hat{\varphi}_j|^2$ e $|\hat{\varphi}|^2$ (que são não-negativas), segue que $\|\hat{\varphi}\|_2 \leq 1$ e logo $\hat{\varphi} \in L^2(\mathbb{R})$, já que $\|\hat{\varphi}_j\|_2 = 1$ para todo j . Agora, para aplicar o teorema da Convergência Dominada, será considerado um múltiplo de $\hat{\varphi}$ como função dominante.

Combinando as equações (2.65) e (2.38), segue que

$$\hat{\varphi}_j(\xi) = \begin{cases} \frac{\hat{\varphi}(\xi)}{\hat{\varphi}(2^{-j}\xi)} & \text{se } |\xi| \leq 2^{j-1} \\ 0 & \text{caso contrário} \end{cases} \quad (2.66)$$

A divisão é possível graças à condição (2.63). Porém, $\hat{\varphi}$ é limitada inferiormente (em módulo) em $[-\frac{1}{2}, \frac{1}{2}]$, já que é contínua e $A_0(2^{-j}\xi) \neq 0$ para $|\xi| \leq 1/2$ devido à equação (2.63). Então existe $c > 0$ em \mathbb{R} tal que $c = (\inf_{[-1/2, 1/2]} |\hat{\varphi}|)^{-1}$. Sendo assim, por (2.66) segue que $|\hat{\varphi}_j(\xi)| \leq c|\hat{\varphi}(\xi)|$ para todo ξ . Logo, aplicando o teorema da Convergência Dominada (considerando a sequência de funções $|\hat{\varphi}_j - \hat{\varphi}|^2$ e a função dominante $[2(c+1)|\hat{\varphi}|^2]$), segue que $\hat{\varphi}_j \rightarrow \hat{\varphi}$ na norma $L^2(\mathbb{R})$ e logo $\{\varphi(x-\gamma)\}_{\gamma \in \mathbb{Z}}$ é um conjunto ortonormal. ■

Resumindo os principais resultados obtidos até aqui para a Wavelet e a Scaling no domínio de Fourier, podemos notar que as condições para gerar uma análise multiresolucional foram atendidas:

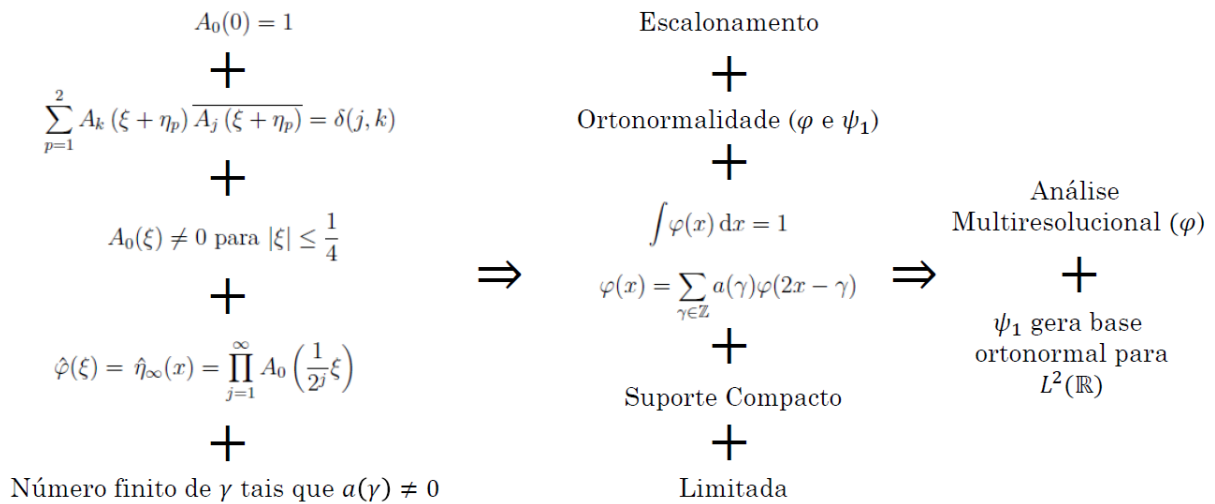


Figura 4 – Condições sobre os coeficientes A_0 e A_1 para obtenção de uma análise multiresolucional

2.5 A construção de uma Wavelet

Para construir a função Wavelet, será necessário encontrar coeficientes a_k ou A_k que satisfaçam as condições estabelecidas nas seções anteriores, a saber (2.34), (2.55) e (2.63).

$$A_0(0) = 1 \tag{2.34}$$

$$|A_k(\xi)|^2 + \left|A_k\left(\xi + \frac{1}{2}\right)\right|^2 = 1 \text{ para todo } \xi \in \mathbb{R} \tag{2.55}$$

$$A_0(\xi) \neq 0 \text{ para } |\xi| \leq \frac{1}{4} \tag{2.63}$$

Além disso, as soluções procuradas aqui são as que têm um número finito de $a_k(\gamma)$ diferentes de zero, o que significa (pela equação (2.33)) que $A_k(\xi)$ são polinômios trigonométricos. Tal fato implica que as funções φ e ψ_1 possuem suporte compacto, o que será provado no seguinte teorema:

Teorema 2.5.1. Se $a_k(\gamma)$ é diferente de 0 para um número finito de $\gamma \in \mathbb{Z}$, então ψ_k possui suporte compacto.

Demonstração. Suponha que $a_0(\gamma)$ é não-nulo para $N_- \leq \gamma \leq N_+$. De acordo com a construção feita no teorema 2.4.6, o suporte de cada η_l está contido no intervalo $[N_{l,-}, N_{l,+}]$, o qual dado por

$$N_{l,-} = \frac{1}{2}(N_{l-1,-} + N_-), \quad N_{l,+} = \frac{1}{2}(N_{l-1,+} + N_+)$$

Isso significa que, como o suporte de η_0 é dado por $[-1/2, 1/2)$, para $l \geq 1$ temos:

$$N_{l,-} = -\left(\frac{1}{2}\right)^{l+2} + \sum_{k=1}^l \left(\frac{1}{2}\right)^k N_-$$

$$N_{l,+} = \left(\frac{1}{2}\right)^{l+2} + \sum_{k=1}^l \left(\frac{1}{2}\right)^k N_+$$

e portanto, fazendo $l \rightarrow \infty$, o suporte de η_∞ está contido em $[N_-, N_+]$, e portanto φ tem suporte compacto. Além disso, considerando que $a_1(\gamma)$ é não-nulo para um número finito de $\gamma \in \mathbb{Z}$, como

$$\psi_1(x) = \sum_{\gamma} a_1(\gamma) \varphi(2x - \gamma)$$

então ψ_1 também possui suporte compacto. ■

Retornando à construção da Wavelet, primeiro será determinado o coeficiente A_0 , que deve satisfazer às seguintes condições:

$$A_0(0) = 1 \tag{2.67}$$

$$|A_0(\xi)|^2 + \left|A_0\left(\xi + \frac{1}{2}\right)\right|^2 = 1 \tag{2.68}$$

$$A_0(\xi) \neq 0 \text{ para } |\xi| \leq \frac{1}{4} \tag{2.69}$$

além de ser da forma

$$A_0(\xi) = \frac{1}{2} \sum_{\gamma \in \mathbb{Z}} a_0(\gamma) e^{2\pi i \gamma \xi} \text{ (soma finita)}$$

Note também que $|A_0(\xi)|^2$ pode ser escrito como

$$|A_0(\xi)|^2 = \frac{1}{4} \sum_{\gamma \in \mathbb{Z}} (\Re\{a(\gamma)\} \cos(2\pi\gamma\xi) - \Im\{a(\gamma)\} \text{sen}(2\pi\gamma\xi))^2$$

$$+ \frac{1}{4} \sum_{\gamma \in \mathbb{Z}} (\Re\{a(\gamma)\} \text{sen}(2\pi\gamma\xi) + \Im\{a(\gamma)\} \cos(2\pi\gamma\xi))^2$$

onde $\Re\{c\}$ é a parte real de um número complexo c e $\Im\{c\}$ a parte imaginária de x .

Uma solução possível seria

$$A_0(\xi) = \frac{1}{2}(1 + e^{2\pi i\xi}) = e^{\pi i\xi} \cos(\pi\xi)$$

que resulta na Wavelets Haar. Porém é interessante também obter Wavelets contínuas, e tal Wavelet é descontínua. Para conseguir atender a continuidade e as condições acima, considere a identidade trigonométrica $\cos^2 \pi\xi + \sin^2 \pi\xi = 1$. Elevando-a à quinta potência, obtemos

$$\begin{aligned} 1 &= (\cos^2 \pi\xi + \sin^2 \pi\xi)^5 \\ &= \cos^{10} \pi\xi + 5 \cos^8 \pi\xi \sin^2 \pi\xi + 10 \cos^6 \pi\xi \sin^4 \pi\xi \\ &\quad + 10 \cos^4 \pi\xi \sin^6 \pi\xi + 5 \cos^2 \pi\xi \sin^8 \pi\xi + \sin^{10} \pi\xi \end{aligned} \quad (2.70)$$

Considere a primeira metade dos termos como $|A_0(\xi)|^2$:

$$|A_0(\xi)|^2 = \cos^{10} \pi\xi + 5 \cos^8 \pi\xi \sin^2 \pi\xi + 10 \cos^6 \pi\xi \sin^4 \pi\xi \quad (2.71)$$

Como $\cos^2 \left(\pi\xi + \frac{1}{2}\right) = \sin^2 \pi\xi$ e $\sin^2 \left(\pi\xi + \frac{1}{2}\right) = \cos^2 \pi\xi$, segue que

$$\left|A_0\left(\xi + \frac{1}{2}\right)\right|^2 = 10 \cos^4 \pi\xi \sin^6 \pi\xi + 5 \cos^2 \pi\xi \sin^8 \pi\xi + \sin^{10} \pi\xi$$

que é a segunda metade de (2.70). Logo, vale que

$$|A_0(\xi)|^2 + \left|A_0\left(\xi + \frac{1}{2}\right)\right|^2 = 1$$

Além disso, $A_0(0) = 1$ e sendo A_0 a soma de três termos não-negativos, segue que $A_0(\xi) \neq 0$ para $|\xi| < \frac{1}{4}$, pois $\cos^{10} \pi\xi$ é diferente de 0 para $|\xi| < \frac{1}{4}$.

Para obter A_0 de fato, é necessário obter uma função cujo módulo ao quadrado seja (2.71). Como

$$\begin{aligned} |A_0(\xi)|^2 &= \cos^6 \pi\xi \left(\cos^4 \pi\xi + 5 \cos^2 \pi\xi \sin^2 \pi\xi + 10 \sin^4 \pi\xi \right) \\ &= \cos^6 \pi\xi \left[\left(\cos^2 \pi\xi - \sqrt{10} \sin^2 \pi\xi \right)^2 + \left(5 + 2\sqrt{10} \right) \cos^2 \pi\xi \sin^2 \pi\xi \right] \end{aligned}$$

podemos tomar A_0 como

$$\begin{aligned} A_0(\xi) &= \left(e^{\pi i\xi} \cos \pi\xi \right)^3 \left(\cos^2 \pi\xi - \sqrt{10} \sin^2 \pi\xi + i\sqrt{5 + 2\sqrt{10}} \cos \pi\xi \sin \pi\xi \right) \\ &= \frac{1}{8} \left(e^{2\pi i\xi} + 1 \right)^3 \left[\frac{1 - \sqrt{10}}{2} + \frac{1 + \sqrt{10}}{4} \left(e^{2\pi i\xi} + e^{-2\pi i\xi} \right) \right. \\ &\quad \left. + \frac{1}{4} \sqrt{5 + 2\sqrt{10}} \left(e^{2\pi i\xi} - e^{-2\pi i\xi} \right) \right] \end{aligned}$$

Comparando com (2.33), notamos que $a_0(\gamma)$ é real e $a_0(\gamma) \neq 0$ somente se $-1 \leq \gamma \leq 4$.

Agora, basta encontrar A_1 tal que:

$$|A_1(\xi)|^2 + \left| A_1 \left(\xi + \frac{1}{2} \right) \right|^2 = 1$$

e

$$A_0(\xi) \overline{A_1(\xi)} + A_0 \left(\xi + \frac{1}{2} \right) \overline{A_1 \left(\xi + \frac{1}{2} \right)} = 0$$

Para tal, basta tomar

$$A_1(\xi) = -e^{2\pi i \xi} \overline{A_0 \left(\xi + \frac{1}{2} \right)} \quad (2.72)$$

já que neste caso

$$\begin{aligned} |A_1(\xi)|^2 + \left| A_1 \left(\xi + \frac{1}{2} \right) \right|^2 &= \left| -e^{2\pi i \xi} \overline{A_0 \left(\xi + \frac{1}{2} \right)} \right|^2 + \left| e^{2\pi i \xi} \overline{A_0(\xi + 1)} \right|^2 \\ &= \left| A_0 \left(\xi + \frac{1}{2} \right) \right|^2 + |A_0(\xi + 1)|^2 \\ &= 1 \end{aligned}$$

e

$$\begin{aligned} &A_0(\xi) \overline{A_1(\xi)} + A_0 \left(\xi + \frac{1}{2} \right) \overline{A_1 \left(\xi + \frac{1}{2} \right)} \\ &= -A_0(\xi) e^{-2\pi i \xi} A_0 \left(\xi + \frac{1}{2} \right) + A_0 \left(\xi + \frac{1}{2} \right) e^{-2\pi i \xi} A_0(\xi + 1) \\ &= -A_0(\xi) e^{-2\pi i \xi} A_0 \left(\xi + \frac{1}{2} \right) + A_0 \left(\xi + \frac{1}{2} \right) e^{-2\pi i \xi} A_0(\xi) \\ &= 0 \end{aligned}$$

Sendo assim, por (2.39), $\hat{\psi}_k$ está definida e para obter ψ_k basta realizar a Transformada Inversa de Fourier de $\hat{\psi}_k$.

As ψ_k desenvolvidas podem ser vistas na figura 5, que representam a Wavelet Daubechies 3 (DAUBECHIES, 1988).

Para se obter outras Wavelets com tal método podemos tomar diferentes potências da identidade trigonométrica. Por exemplo, tomando a segunda potência, obtemos:

$$\begin{aligned} 1 &= \left(\cos^2(\xi\pi) + \sin^2(\xi\pi) \right)^2 \\ &= \cos^4(\xi\pi) + 2 \cos^2(\xi\pi) \sin^2(\xi\pi) + \sin^4(\xi\pi) \end{aligned}$$

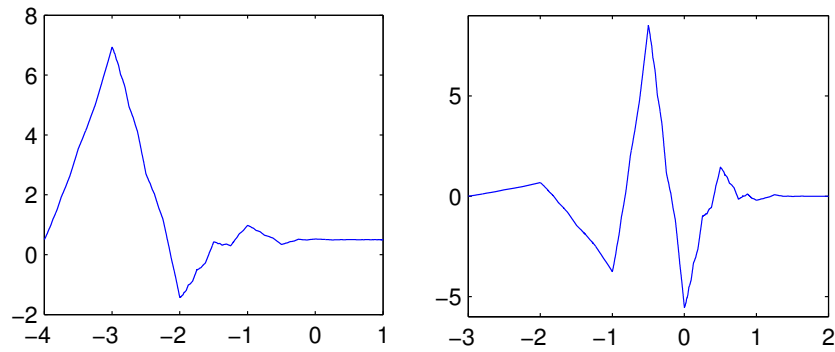


Figura 5 – Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a quinta potência da identidade trigonométrica

Fazendo

$$\begin{aligned} |A_0(\xi)|^2 &= \cos^4(\xi\pi) + \cos^2(\xi\pi)\sin^2(\xi\pi) \\ &= \cos^2(\xi) (\cos^2(\xi\pi) + \sin^2(\xi\pi)) \\ &= \cos^2(\xi) \end{aligned}$$

vale (2.67), (2.68) e (2.69). Podemos então definir

$$A_0(\xi) = e^{\pi i \xi} \cos(\xi\pi)$$

e tomar $A_1(\xi)$ como feito em (2.72). Com esse par de coeficientes a Wavelet gerada é justamente a Haar, como apresentado anteriormente.

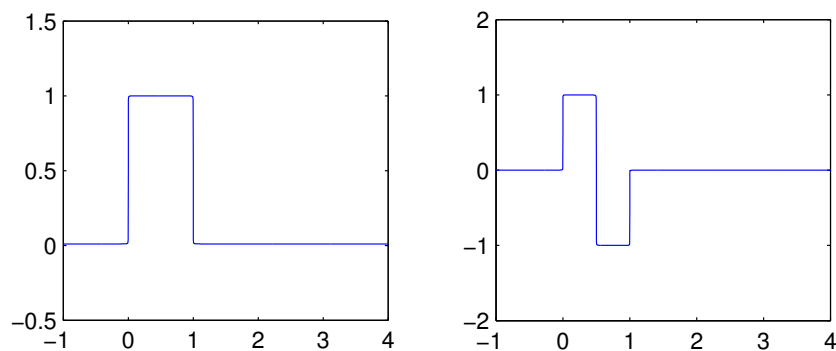


Figura 6 – Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a segunda potência da identidade trigonométrica

O gráfico aparece contínuo devido à impossibilidade computacional de realizar as infinitas iterações presentes em (2.39) e devido à discretização, já que computacionalmente um ponto é ligado ao seguinte.

Por fim, considerando agora a terceira potência da identidade trigonométrica, temos:

$$\begin{aligned}
 1 &= (\cos^2(\xi\pi) + \text{sen}^2(\xi\pi))^3 \\
 &= \cos^6(\xi\pi) + 3\cos^4(\xi\pi)\text{sen}^2(\xi\pi) + 3\cos^2(\xi\pi)\text{sen}^4(\xi\pi) + \text{sen}^6(\xi\pi)
 \end{aligned}$$

Tomando $|A_0|^2$ como

$$\begin{aligned}
 |A_0(\xi)|^2 &= \cos^6(\xi\pi) + 3\cos^4(\xi\pi)\text{sen}^2(\xi\pi) \\
 &= \cos^4(\xi) (\cos^2(\xi\pi) + 3\text{sen}^2(\xi\pi))
 \end{aligned}$$

também vale (2.67), (2.68) e (2.69). Então, fazendo

$$A_0(\xi) = (e^{\pi i \xi} \cos(\pi \xi))^2 (\cos^2(\pi \xi) + i3\text{sen}^2(\pi \xi))$$

e tomando $A_1(\xi)$ como o negativo do feito em (2.72), obtemos a Wavelet Daubechies 2 (DAUBECHIES, 1988), como visto na figura 7.

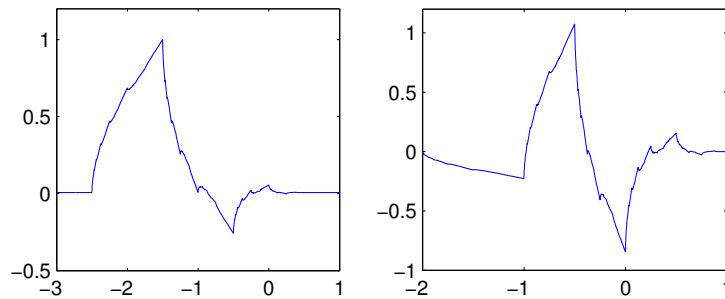


Figura 7 – Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) quando tomada a terceira potência da identidade trigonométrica

As Wavelets utilizadas para o preditor de propriedades do petróleo foram as Daubechies 3 e 4. A construção da Daubechies 4 (figura 8) utilizando o método proposto pode ser complexa algebricamente e portanto não foi realizada neste trabalho. Mais detalhes podem ser vistos em (DAUBECHIES, 1988).

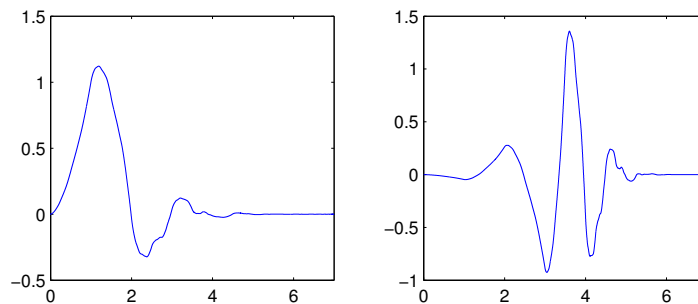


Figura 8 – Gráfico da Scaling φ (esquerda) e da Wavelet ψ_1 (direita) Daubechies 4

3 Método de Obtenção de Propriedades do Petróleo

O método proposto tem como objetivo obter as propriedades de amostras de petróleo a partir do seu espectro, obtido pela emissão de infravermelho.

Basicamente, o espectro é um conjunto de pontos que descreve uma curva, como a seguinte:

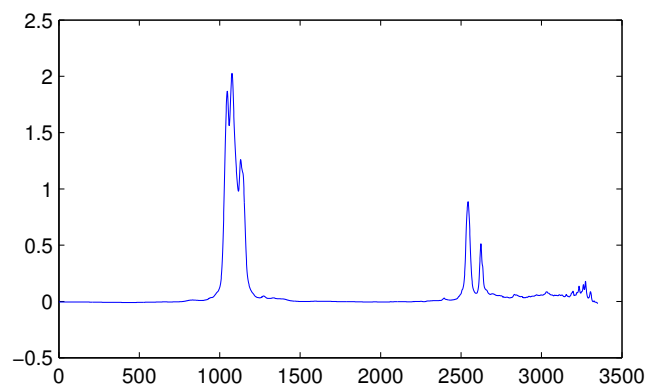


Figura 9 – Gráfico do Espectro da Amostra 1

A partir dele, serão obtidas duas propriedades do petróleo: a **densidade relativa a 20/4 °C** (por simplicidade, mencionada apenas como densidade) e a **TIAC** (Temperatura de Início de Aparecimento de Cristais). Os valores reais dessas propriedades relativos a cada amostra (medidos experimentalmente) foram fornecidos pelo LabPetro.

A amostra 1, por exemplo, possui os seguintes valores relativos a essas propriedades:

Amostra 1	
Densidade	TIAC (°C)
0.7586	23.3500

Tabela 1 – Densidade e TIAC da Amostra 1

Porém, considere duas outras amostras, cujos espectros são mostrados na figura 10. Note que o formato de todos os três espectros são bem parecidos, o que vale para todas as outras amostras. Porém, o valor das propriedades de cada uma é bastante diferente uma da outra, como pode ser visto na tabela 2.

Sendo assim, inspirado pela Análise Multiresolucional apresentada nesta dissertação, propomos nesse método que se utilize Wavelets para tentar extrair de cada espectro

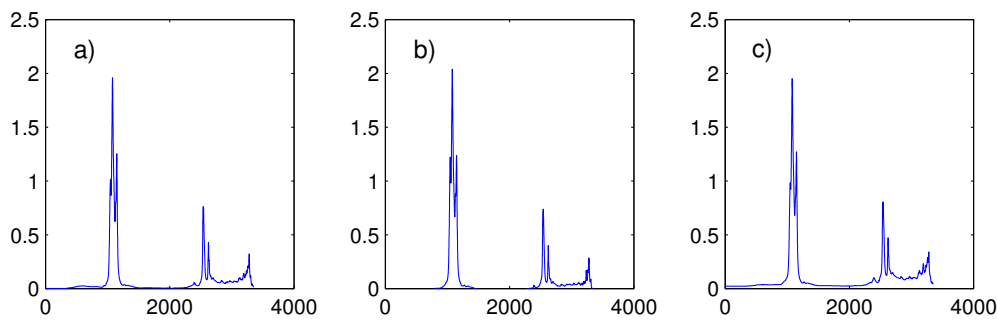


Figura 10 – a) - Gráfico do Espectro da Amostra 1. b) - Gráfico do Espectro da Amostra 47. c) - Gráfico do Espectro da Amostra 57.

	Amostra 1	Amostra 47	Amostra 57
Densidade	0.7586	0.8844	0.8744
TIAC ($^{\circ}C$)	23.3500	8.32	36.53

Tabela 2 – Densidade e TIAC das Amostras 1, 47 e 57

a informação relevante a cada propriedade. Isso será feito obtendo coeficientes dados pelo produto interno das Wavelets em uma dada resolução com as duas principais regiões do espectro:

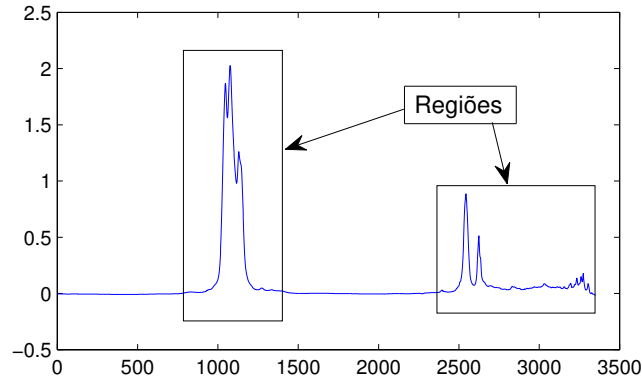


Figura 11 – Regiões do Espectro utilizadas para obtenção das Propriedades

Tais regiões podem variar em relação ao seu ponto de origem e término. Esses 4 pontos (origem e término de cada região) são parâmetros que necessitam de otimização, realizada pelo algoritmo genético.

Os coeficientes então extraídos dos espectros serão fornecidos como entrada de um sistema Fuzzy modelo Takagi-Sugeno. Tal modelo Fuzzy será gerado a partir da Clustering via Método da Montanha mencionada na seção 3.7, sendo as entradas do método dadas pelos coeficientes e a saída uma das duas propriedades. Porém, os parâmetros do método, a saber **Raio de Influência**, **Fator de Esmagamento**, **Taxa de Aceitação** e **Taxa de Rejeição** influenciam enormemente o desempenho do sistema. Sendo assim, eles são outros 4 parâmetros que necessitam de otimização.

Os 8 parâmetros mencionados que carecem de otimização constituirão um indivíduo de um Algoritmo Genético. Esse terá uma população de 200 indivíduos, que sofrerão mutação, *cross-over* e possivelmente serão eliminados caso tenham baixo desempenho. Sempre que um indivíduo é gerado, um sistema Fuzzy é construído com base nas entradas fornecidas quando se tem as regiões cujos pontos estão definidos no indivíduo e nos parâmetros da clusterização também definidos no indivíduo. Esse sistema é então testado com amostras de teste, não utilizadas na clusterização, e com base na proximidade de sua saída com o valor real da propriedade é estabelecido o desempenho do indivíduo.

A figura 12 mostra um fluxograma do método.

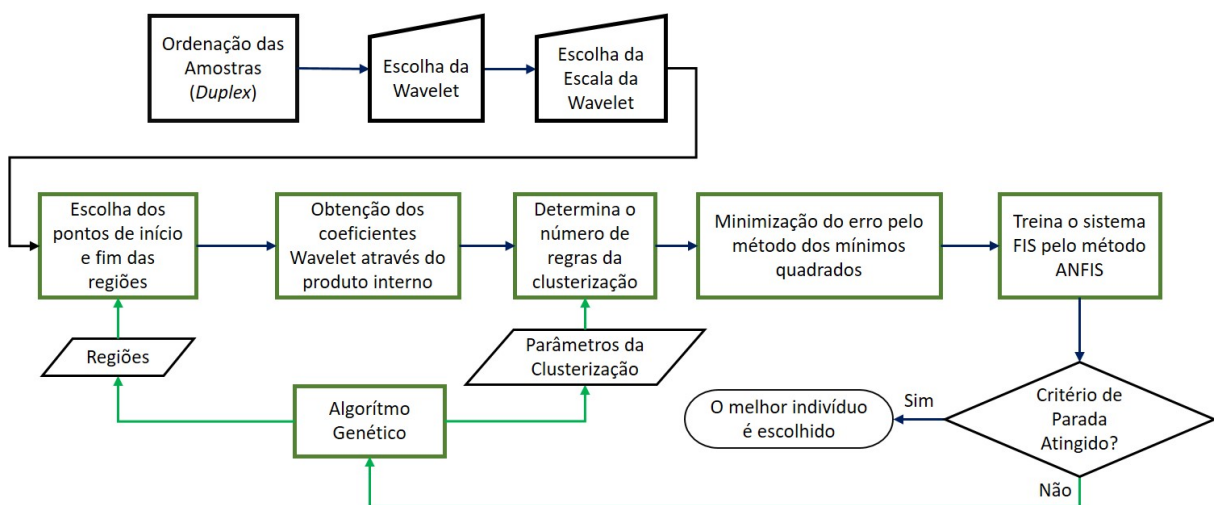


Figura 12 – Fluxograma do Método

3.1 Ordenação das amostras

Antes de aplicar o método, é necessário definir quais amostras serão utilizadas para treino (no método em si) e quais serão utilizadas para teste (verificação da validade do sistema). Fazemos essa ordenação porque é importante que no treino haja várias amostras de características distintas, mas que ainda sobre um conjunto diversificado para validação. Isso permite que o modelo contemple um grande número de tipos de espectro, e ainda prove sua validade em boa parte deles. Neste trabalho foram consideradas 70% das amostras para treino e 30% para teste.

O método utilizado para ordenar as amostras foi o *Duplex* (SNEE, 1977), e se baseia em tomar cada ponto dos 3351 pontos do espectro como uma dimensão euclidiana para calcular distâncias entre dois espectros pela norma euclidiana. Os dois espectros que tenham a maior distância entre si serão colocados no grupo de treino. Dos espectros restantes, os dois espectros que tenham a maior distância entre si serão colocados no grupo de teste.

A partir dessa etapa, dentre os espectros não classificados em teste ou treino, relacionamos cada um deles ao espectro de treino mais próximo de cada um, e aquele que tiver a maior distância ao seu respectivo espectro de treino mais próximo será colocado no grupo de treino. O mesmo é feito com o grupo de teste considerando os espectros ainda não classificados. Esse processo se repete até que a porcentagem de espectros de teste com relação ao número total de espectros seja alcançada. Os que sobraram sem classificação são então colocados no grupo de treino.

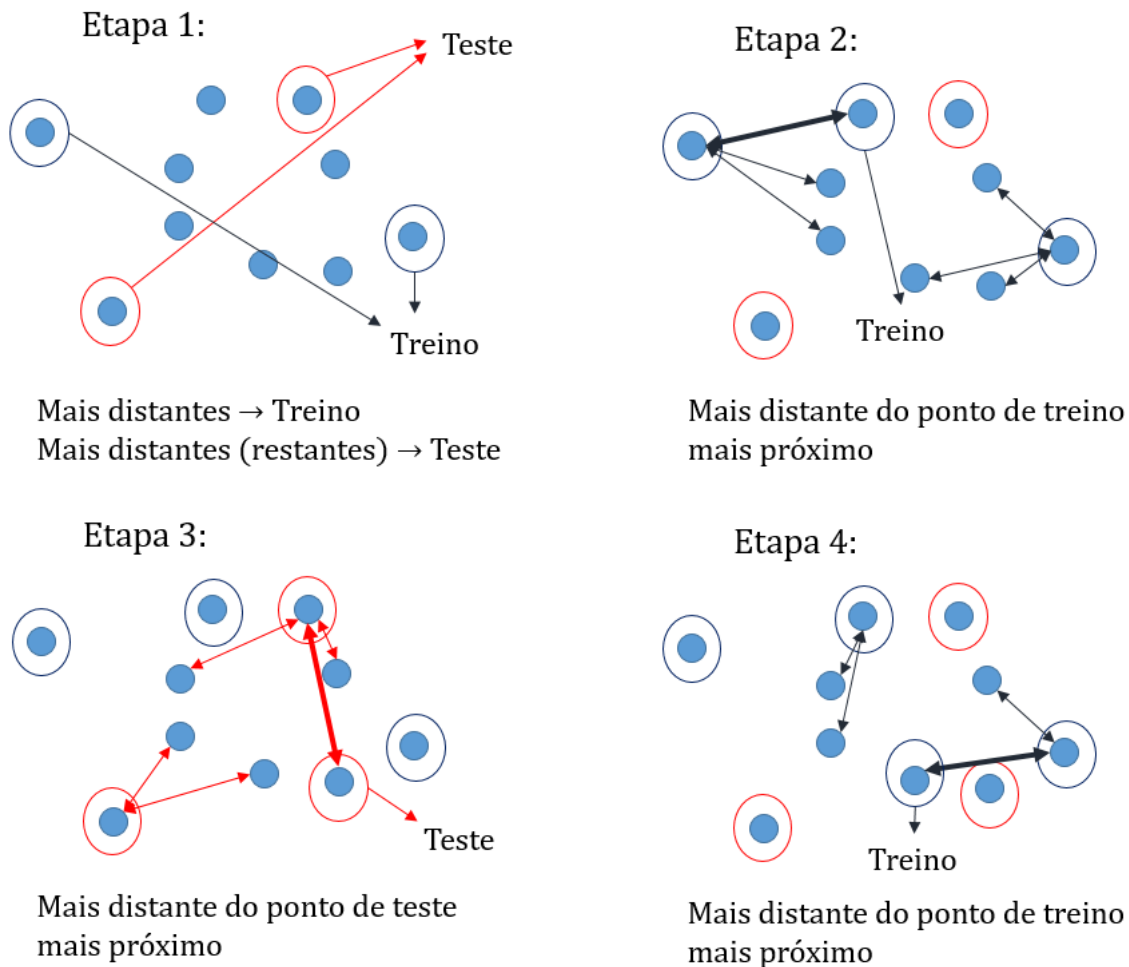


Figura 13 – Exemplo bidimensional do método Duplex (até à quarta etapa apenas)

3.2 Escolha da Wavelet

Como é feito o produto interno da Wavelet com parte do espectro, cada Wavelet irá produzir coeficientes diferentes. As Wavelets em geral satisfazem $\int \psi_1(x) dx = 0$, o que significa que elas "filtram" constantes. Isso é desejado, pois os espectros geralmente possuem um ruído que pode ser considerado constante, o qual deve ser eliminado para melhores resultados.

Porém, existem Wavelets que filtram componentes lineares, quadráticas, e muitas outras. Sendo assim, uma Wavelet pode ser melhor para obter uma certa propriedade por esse método do que outra. Não há como dizer previamente qual é a melhor, por isso é necessário testá-las até obter um bom resultado.

A Wavelet que mostrou melhor resultado para a densidade foi a Daubechies 3:

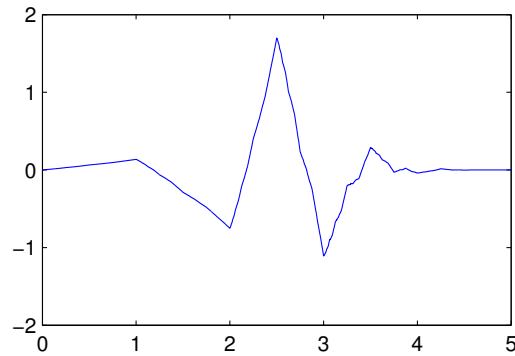


Figura 14 – Wavelet Daubechies 3

Já para a propriedade TIAC, a Wavelet com melhor resultado foi a Daubechies 4:

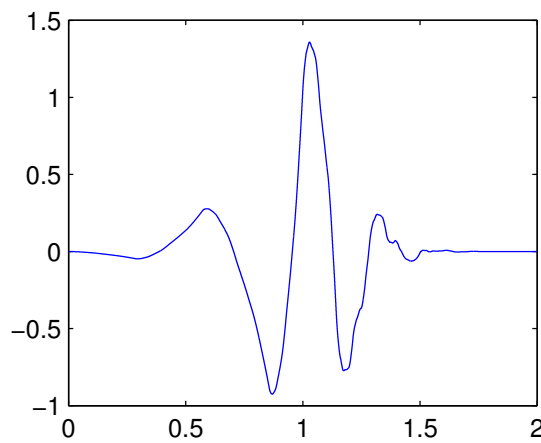


Figura 15 – Wavelet Daubechies 4

3.3 Escolha da escala da Wavelet

A escala da Wavelet é basicamente o número de k partes de mesmo tamanho em que cada região do espectro será dividida. É então realizado o produto interno de cada uma dessas $2k$ partes (pois há 2 regiões) com uma Wavelet (em todo o seu suporte), que é ajustada de modo a seu suporte coincidir com o tamanho da região. Mais dessa operação serão discutidos na Seção 3.5, que trata do produto interno.

Ambas as propriedades apresentaram melhor resultado quando a escala da Wavelet vale 5. A figura 16 exemplifica tal escala aplicada em uma região:

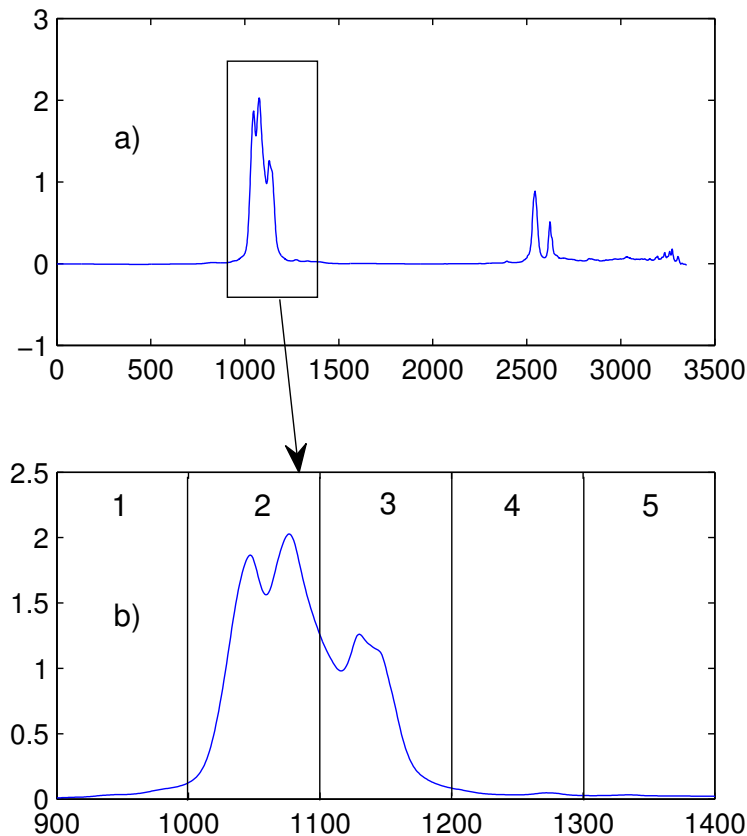


Figura 16 – a) Região selecionada no espectro. b) Divisão da região selecionada em 5 partes iguais (subregiões), devido à escala 5 escolhida.

3.4 Escolha dos pontos de início e fim das regiões

As informações mais visíveis no espectro estão em seus picos. Porém, existe uma região em volta de cada pico que pode conter informações necessárias para a determinação do valor da propriedade. Além disso, mudando o tamanho e centro da região, mudarão também as subregiões criadas pela escala, e portanto o valor dos coeficientes obtidos pelo produto interno delas com a Wavelet.

Portanto, como não é possível dizer qual das possíveis subregiões resultará em um coeficiente que forneça o melhor resultado, foi necessário colocar os pontos de início e término das regiões como parâmetros do indivíduo no Algoritmo Genético.

Ainda assim, os possíveis pontos de início e fim de cada região são limitados, de modo a não ficarem muito distantes dos picos, mas também impedindo que um ponto inicial seja maior que um ponto final para uma mesma região.

Para a primeira região, o ponto inicial pode variar de 800 a 1100 e o ponto final de 1101 a 1400.

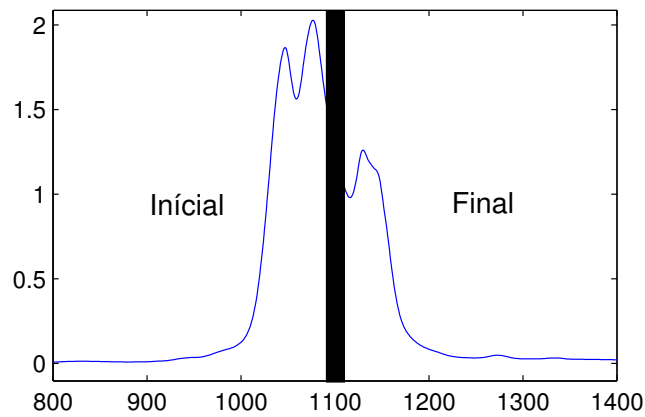


Figura 17 – Possíveis pontos iniciais e finais da primeira região

Para a segunda região, o ponto inicial pode variar de 2000 a 2651 e o ponto final de 2652 a 3351.

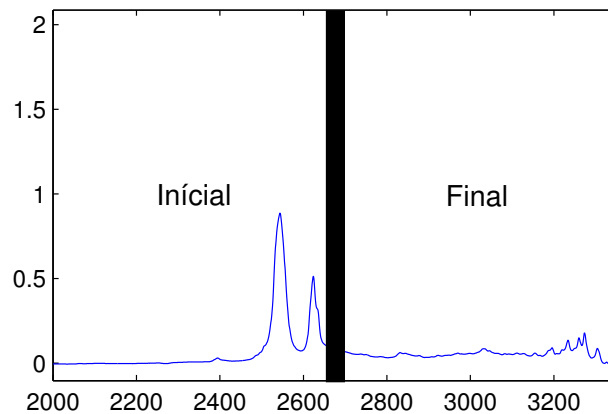


Figura 18 – Possíveis pontos iniciais e finais da segunda região

3.5 Produto interno

Cada Wavelet terá um certo número fixo de pontos N_1 , gerados computacionalmente. Cada subregião gerada na escala terá um número de pontos N_2 , bem menor que N_1 . Para fazer o produto interno da Wavelet com a subregião, foi necessário adaptar o produto interno entre vetores, para que tal operação pudesse contemplar dois vetores com tamanhos distintos.

Para tal, simplesmente dividimos o número de pontos da subregião em N_1 pontos, de modo que o primeiro ponto, original da subregião, coincida com o primeiro ponto da Wavelet e também o último. Os pontos fictícios criados na região (já que $N_1 > N_2$) são simplesmente a interpolação entre dois pontos existentes na representação original de N_1 pontos. Por exemplo, o ponto n ($1 \leq n \leq N_2$) criado na subregião terá valor

$$v_{N_2}(n) = [v_{N_1}(k+1) - v_{N_1}(k)]\alpha + v_{N_1}(k)$$

onde v_{N_2} representa o valor do ponto na nova divisão da região por N_2 pontos e v_{N_1} o valor na representação original da região por N_1 pontos. Os valores de k e α são dados por

$$k = \left\lfloor (N_1 - 1) \frac{(n - 1)}{(N_2 - 1)} \right\rfloor + 1 \quad (3.1)$$

$$\alpha = (N_1 - 1) \left(\frac{n - 1}{N_2 - 1} - \frac{k - 1}{N_2 - 1} \right)$$

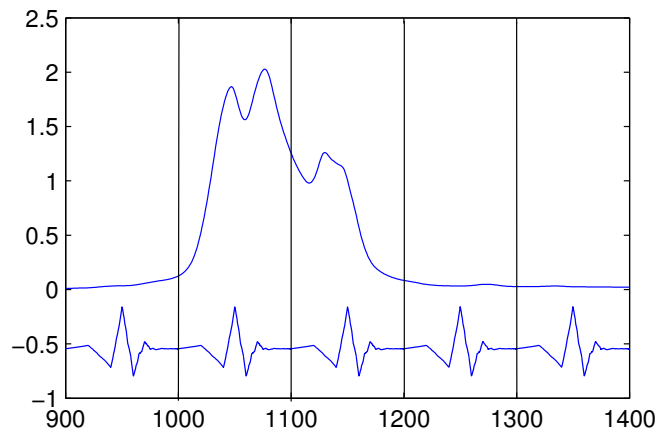


Figura 19 – Wavelets ajustadas ao tamanho de cada subregião para uma escala de 5 na primeira região com ponto inicial de 900 e final de 1400

Sendo assim, dada uma escala s , obtemos através de tal produto interno das sub-regiões com a Wavelet escolhida os $2s$ coeficientes Wavelets, que serão utilizados como entrada para o sistema Fuzzy.

3.6 Modelagem Fuzzy

A Teoria Fuzzy é um método eficaz de lidar com sistemas não lineares de forma prática, utilizando informação heurística (PASSINO; YURKOVICH, 1998). Através do Fuzzy, comportamentos de sistemas que envolvem modelagens complexas podem ser traduzidos em simples regras, que podem até mesmo ser adicionadas devido à própria observação de especialistas no processo.

Seu uso é de grande interesse em controladores, e conseqüentemente em indústrias, mas também utilizado em eletrônicos através de microcontroladores (DRIANKOV; HELLENDORRN; REINFRANK, 1996). Sua teoria é bastante difundida no Japão, por

exemplo, onde um sistema que utilize Fuzzy é referenciado como moderno e de alta qualidade (DRIANKOV; HELLENDORN; REINFRANK, 1996). A Figura 20 representa basicamente o mecanismo de um sistema Fuzzy:

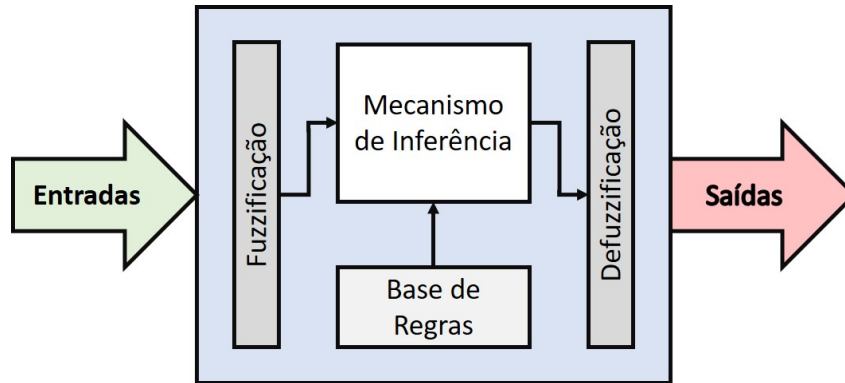


Figura 20 – Representação Genérica de um Sistema Fuzzy

Seu funcionamento pode ser descrito da seguinte maneira:

- Entradas de um processo são fornecidas ao sistema;
- A informação das entradas é então traduzida em termos da lógica Fuzzy, para serem processados através de variáveis Fuzzy;
- Um Mecanismo de Inferência associa as variáveis Fuzzy obtidas com as regras que constam na Base de Regras, gerando outras variáveis que estão relacionadas à resposta do sistema às entradas;
- As variáveis de saída do Mecanismo de Inferência são então "defuzzyficadas", traduzidas na resposta do sistema àquelas entradas.

Seu comportamento é definido matematicamente, e será descrito nas seguintes seções deste capítulo, de acordo com a teoria encontrada em (DRIANKOV; HELLENDORN; REINFRANK, 1996).

3.6.1 Teoria Fuzzy

Nesta seção, será abordada a teoria Fuzzy, necessária para desenvolver o método de predição de propriedades do petróleo, complementada nas próximas subseções pela definição do funcionamento do sistema Fuzzy.

Definição 3.6.1. Seja U um conjunto qualquer. Considere $C \subseteq U$. Então a função característica ν_C em U é dada por

$$\chi_C : U \mapsto \{0, 1\}$$

$$\chi_C(u) = \begin{cases} 1 & \text{se } u \in C \\ 0 & \text{se } u \notin C \end{cases}$$

A teoria Fuzzy trabalha com funções denominadas funções de pertinência. Elas, assim como a função característica, determinam se um elemento pertence ou não a um conjunto. A diferença é que as funções de pertinência podem assumir qualquer valor em $[0, 1]$, associando a um elemento o grau de pertinência dele ao conjunto em questão.

Definição 3.6.2. Seja U um conjunto qualquer. Considere $C \subseteq U$. Então a função de pertinência μ_C em U é uma função qualquer tal que

$$\mu_C : U \mapsto [0, 1]$$

Agora, podemos definir um conjunto Fuzzy como:

Definição 3.6.3. Um conjunto Fuzzy F com respectiva função de pertinência μ_F definida em U (também chamado de conjunto universo) é dado por

$$F = \{(u, \mu_F(u)) : u \in U\}$$

Um exemplo do gráfico de uma função de pertinência pode ser visto na Figura 21.



Figura 21 – Exemplo de uma função de pertinência definida no intervalo $[0, 10]$

Exemplo 3.6.1. Considere o conjunto $U = \{0, 1, 2, \dots, 9, 10\}$. Desejamos definir uma função de pertinência que relacione o quão próximo de 5 um elemento de U está. Podemos denotá-la por μ_{p5} e fazer

Podemos então criar um conjunto Fuzzy $P5$, dado por

u	$\mu_{p5}(u)$
0	0
1	0
2	0.1
3	0.4
4	0.8
5	1
6	0.8
7	0.4
8	0.1
9	0
10	0

Tabela 3 – Valores da Função de Pertinência μ_{p5}

$$P5 = \{(0, 0), (1, 0), (2, 0.1), (3, 0.4), (4, 0.8), (5, 1), (6, 0.8), (7, 0.4), (8, 0.1), (9, 0), (10, 0)\}$$

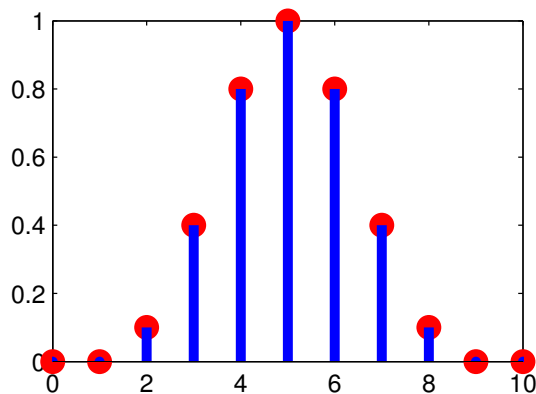


Figura 22 – Pertinência da proximidade com o número 5 dos números de 1 a 10

As próximas definições consideram um conjunto Fuzzy F definido num conjunto universo U .

Definição 3.6.4. O suporte de um conjunto Fuzzy F é dado por

$$S(F) = \{u \in U : \mu_F(u) > 0\}$$

Definição 3.6.5. O núcleo de um conjunto Fuzzy F é dado por

$$N(F) = \{u \in U : \mu_F(u) = 1\}$$

Definição 3.6.6. A altura de um conjunto Fuzzy F , denotada por $\text{alt}(F)$ é dado por

$$\text{alt}(F) = \sup_{u \in U} \mu_F(u)$$

e um conjunto Fuzzy F é chamado **normal** se $\text{alt}(F) = 1$ e **subnormal** se $\text{alt}(F) < 1$.

Definição 3.6.7. Um conjunto Fuzzy F é convexo se, e somente se:

$$\forall u, v \in U, \forall \lambda \in [0, 1] : \mu_F(\lambda u + (1 - \lambda)v) \geq \min\{\mu_F(u), \mu_F(v)\}$$

A noção de conjunto Fuzzy convexo é que ele não possui "vales" em sua função de pertinência. Por exemplo, a figura 23 representa a função de pertinência de um conjunto Fuzzy não convexo, enquanto a figura 24 representa a função de pertinência de um conjunto Fuzzy convexo.

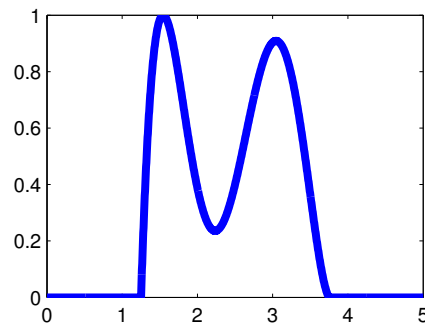


Figura 23 – Função de pertinência de um conjunto não-convexo

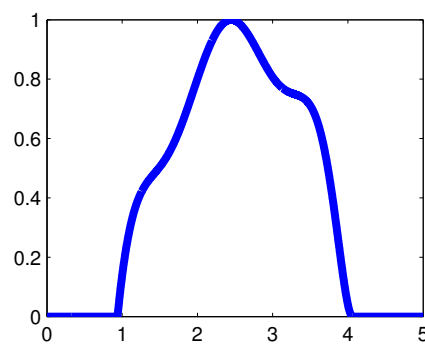


Figura 24 – Função de pertinência de um conjunto convexo

Definição 3.6.8. Dois conjuntos Fuzzy A e B (definidos no mesmo conjunto universo U) são iguais se, e somente se

$$\mu_A(u) = \mu_B(u) \quad \forall u \in U$$

Definição 3.6.9. O conjunto Fuzzy A é um subconjunto do conjunto Fuzzy B ($A \subseteq B$) se, e somente se

$$\mu_A(u) \leq \mu_B(u) \quad \forall u \in U$$

Podemos definir ainda que $A \subsetneq B$ se, e somente se

$$\mu_A(u) < \mu_B(u) \quad \forall u \in U$$

Pretendemos agora realizar operações entre conjuntos Fuzzy, como interseção, união e complemento. Existem vários modos de definir tais operações, porém elas seguem algumas regras gerais, definidas a seguir.

Definição 3.6.10. Uma norma triangular (T-norma) $\widehat{\star}$ é uma classe de funções binárias que podem representar a operação interseção. Ela satisfaz as seguintes propriedades, para $a, b \in [0, 1]$:

1. $a \widehat{\star} b = b \widehat{\star} a$ (comutativa)
2. $(a \widehat{\star} b) \widehat{\star} c = a \widehat{\star} (b \widehat{\star} c)$ (associativa)
3. $a \leq c$ e $b \leq d$ implica que $a \widehat{\star} b \leq c \widehat{\star} d$
4. $a \widehat{\star} 1 = a$

Definição 3.6.11. Uma T-norma é dita Arquimediana se satisfaz todas as propriedades de uma T-norma e

$$a \widehat{\star} a < a \quad \forall a \in (0, 1)$$

Definição 3.6.12. Uma co-norma triangular, ou S-norma $\underbrace{\star}$, é uma classe de funções binárias que podem representar a operação união. Ela satisfaz os 3 primeiros critérios da T-norma, porém seu quarto critério é

$$a \underbrace{\star} 0 = a$$

Definição 3.6.13. A operação complemento c deve satisfazer pelo menos as seguintes propriedades:

1. $c(0) = 1$
2. $a < b \implies c(a) > c(b)$

$$3. c(c(a)) = a$$

Sendo assim, dados dois conjuntos Fuzzy A e B podemos definir as operações interseção ($A \cap B$), união ($A \cup B$) e complemento (A^c) através de suas funções de pertinência, respectivamente:

$$\mu_{A \cap B}(u) = \mu_A(u) \overset{\star}{\wedge} \mu_B(u) \quad \forall u \in U$$

$$\mu_{A \cup B}(u) = \mu_A(u) \overset{\star}{\vee} \mu_B(u) \quad \forall u \in U$$

$$\mu_{A^c}(u) = c(\mu_A(u)) \quad \forall u \in U$$

A operação união são também chamadas da operação lógica **O**U. Já a operação interseção é chamada da operação lógica **E**. Note que tais operações atuam no valor real da função de pertinência, e logo podem ser utilizadas em quaisquer outros números reais pertencentes a $[0, 1]$ (não precisam necessariamente fazerem parte de uma função de pertinência).

Exemplo 3.6.2. A função mínimo representa uma T-norma. Considere dois conjuntos Fuzzy A e B no conjunto universo $[0, 10]$ definidos pelas suas funções de pertinência como na figura 25.

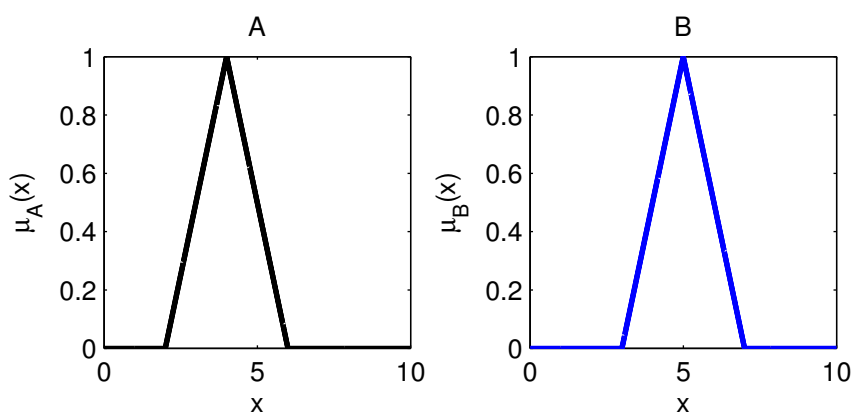
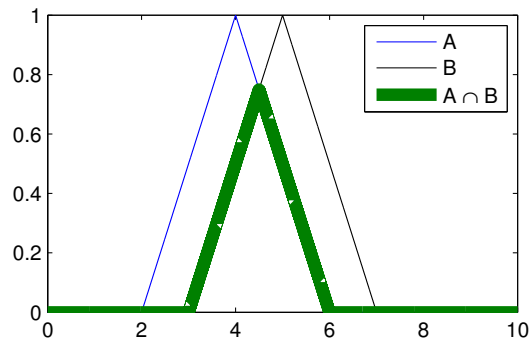


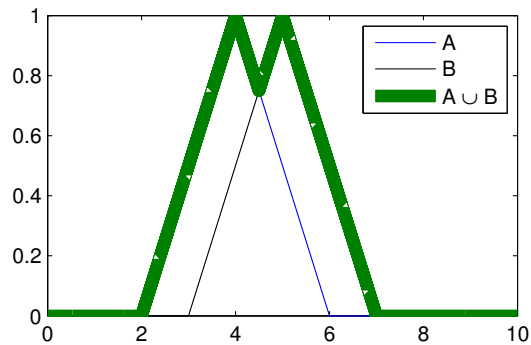
Figura 25 – Funções de pertinência de A e B

Sendo assim, o conjunto Fuzzy $A \cap B$, utilizando a T-norma "mínimo", possui a função de pertinência como na figura 26.

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}$$

Figura 26 – Função de pertinência de $A \cap B$

Exemplo 3.6.3. A função máximo representa uma S-norma. Considere os dois conjuntos Fuzzy A e B do exemplo 3.6.2. O conjunto Fuzzy $A \cup B$, utilizando a S-norma "máximo", possui a função de pertinência como na figura 27.

Figura 27 – Função de pertinência de $A \cup B$

$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}$$

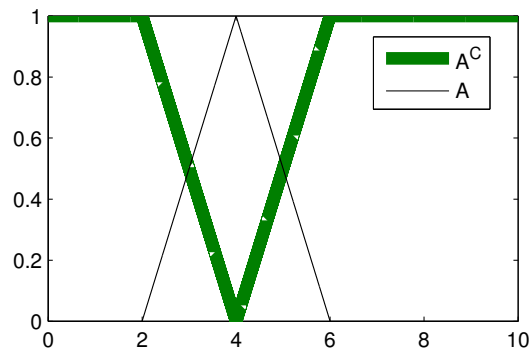
Exemplo 3.6.4. A função $f(x) = 1 - x$, $f : [0, 1] \rightarrow [0, 1]$ representa uma função complemento. Considere o conjunto Fuzzy A do exemplo 3.6.2. O conjunto Fuzzy A^c , utilizando a função f , possui a função de pertinência como na figura 28.

$$\mu_{A^c}(x) = 1 - \mu_A(x)$$

Definição 3.6.14. (Relação Fuzzy) Considere N conjuntos universo A_1, A_2, \dots, A_N . Considere uma função $\mu_R : A_1 \times A_2 \times \dots \times A_N \rightarrow [0, 1]$, onde \times representa o produto cartesiano. Então o conjunto R dado por

$$R = \{(u, \mu_R(u)) : u \in A_1 \times A_2 \times \dots \times A_N\}$$

é denominado uma relação Fuzzy.

Figura 28 – Função de pertinência de A^c

Exemplo 3.6.5. Considere o conjunto

$$A = \{\text{João, Maria, José}\}$$

e o conjunto

$$B = \{\text{Pedro, Paula, Pietra}\}$$

Podemos estabelecer uma relação R entre A e B que represente o quanto uma pessoa conhece a outra, sendo $\mu_R = 1$ para conhece há muito tempo e $\mu_R = 0$ para não conhece:

	Pedro	Paula	Pietra
João	0.9	0.3	1
Maria	0.2	0.8	0.5
José	1	0.4	0.3

Tabela 4 – Pertinência em uma Relação entre Dois Conjuntos de Três Pessoas

3.6.2 Funcionamento de um sistema Fuzzy

Nas próximas subseções será abordado o funcionamento de um sistema que utiliza lógica Fuzzy, desde da etapa da fuzzyficação até a desfuzzyficação, quando o sistema devolve os valores de saída para um conjunto de entradas.

Para tal, será necessário introduzir o conceito de MIMO e MISO. MIMO (*Multiple Input Multiple Output*) é um sistema que possui mais de uma entrada e mais de uma saída. MISO (*Multiple Input Single Output*) é um sistema com mais de uma entrada e uma única saída. Como sistemas MIMO podem ser decompostos em sistemas MISO neste caso, será considerado nas seguintes subseções um sistema MISO.

Considere agora um Sistema Fuzzy MISO. Ele é basicamente composto por conjuntos Fuzzy para cada entrada, um conjunto regras e em sua saída podem ou existir

conjuntos Fuzzy de saída ou existir funções de saída, como em um modelo Takagi-Sugeno, que será discutido posteriormente.

3.6.3 Fuzzyficação

Os valores de entrada de um sistema Fuzzy necessitam de ser transformados em conjuntos Fuzzy para aplicá-los no sistema. Tal processo é denominado fuzzyficação.

Sendo assim, seja U um conjunto universo de entrada. Seja \mathbb{F} o conjunto de todos os conjuntos Fuzzy. Então o operador fuzzyficação \mathfrak{F} é dado por

$$\mathfrak{F} : U \rightarrow \mathbb{F}$$

$$\mathfrak{F}(u) = \dot{F}, \quad u \in U, \dot{F} \in \mathbb{F}$$

Exemplo 3.6.6. Uma fábrica produz resistências de 5 Ohms, porém existem a chance do valor da resistência produzida ser um pouco diferente da projetada, com densidade de probabilidade D dada por

$$D(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-c)^2}{2\sigma^2}}$$

onde o centro é em $c = 5$, sua variância dada por $\sigma^2 = 0.1$ e logo seu valor máximo igual a $M = \frac{10}{\sqrt{2\pi}}$.

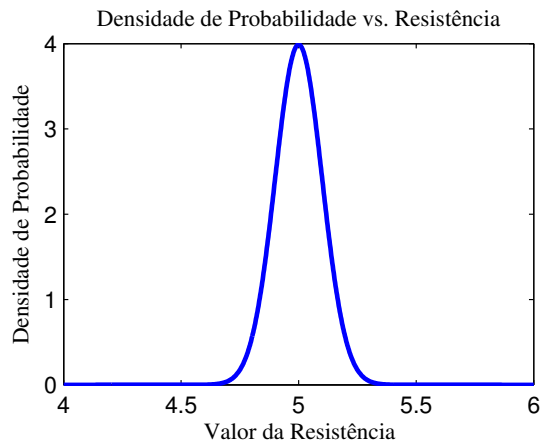


Figura 29 – Densidade de probabilidade D dos valores que uma resistência produzida pode assumir

Um indivíduo que comprar uma dessas resistências pode assumir em seu projeto que tal elemento possui uma resistência igual ao conjunto Fuzzy R dado por

$$R = \{(r, \mu_R(r)) : r \in \mathbb{R}_+\}$$

onde

$$\mu_R(x) = e^{-\frac{(x-c)^2}{2\sigma^2}}$$

que representa a mesma função de densidade de probabilidade, porém normalizada em relação a seu valor máximo.

Note que ocorreu uma fuzzyficação do valor da resistência de 5 Ohms: agora ela é representada por um valor de resistência e a pertinência daquele valor.

Exemplo 3.6.7. Uma das fuzzyficações mais utilizadas é a *singleton*. Ela se baseia em transformar um valor u^* dentro do conjunto universo U em um conjunto Fuzzy F cuja função de pertinência é o delta de Kronecker em u^*

$$F = \{(u, \mu_F(u)) : u \in U\}$$

$$\mu_F(u) = \delta(u, u^*)$$

Para um conjunto universo $U = [3, 5]$ e um valor de entrada $u^* = 4$, temos a função de pertinência de F (a fuzzyficação de u^*) representada na figura 30.

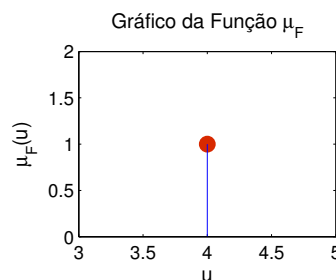


Figura 30 – Gráfico da fuzzyficação do valor 4 em um *singleton*

3.6.4 Afirmação Fuzzy

Uma Afirmação Fuzzy de uma regra é uma afirmação na lógica Fuzzy entre uma entrada e um conjunto Fuzzy (chamada de **premissa**) ou entre uma saída e um conjunto Fuzzy (chamada de **consequência**). Por exemplo, considere uma entrada E , a qual assume valores num conjunto universo U , e um conjunto Fuzzy A , cujo conjunto universo também é U e função de pertinência μ_A . Então podemos estabelecer a premissa entre a entrada E e o conjunto Fuzzy A da seguinte forma:

$$E \text{ é } A$$

O significado dessa premissa pode ser descrito da seguinte maneira: considere que o valor da entrada E é e . A premissa " E é A " diz que, quando a entrada E assume o

valor e , ela pode ser representada pelo conjunto Fuzzy A . Porém, é necessário considerar a pertinência dessa afirmação, que está em $[0, 1]$ (se fosse apenas uma lógica booleana, ou a afirmação seria verdadeira $\mathbf{1}$ ou seria falsa $\mathbf{0}$).

Considere que a fuzzyficação de e seja o conjunto Fuzzy \dot{E} , cuja função de pertinência é $\mu_{\dot{E}}$. Seja \mathbb{M} o conjunto de funções de pertinência. Então definimos uma $f : \mathbb{M} \times \mathbb{M} \rightarrow [0, 1]$, a qual será utilizada para definir a pertinência da premissa com base nas funções de pertinência de \dot{E} e A .

Logo, a premissa " E é A " possui pertinência $\alpha \in [0, 1]$, dada por

$$\alpha = f(\mu_{\dot{E}}, \mu_A)$$

Quando o valor de pertinência da premissa é 0, assume-se que a premissa é falsa e logo não é considerada.

Cada entrada e cada saída podem ter vários conjuntos Fuzzy possíveis de se estabelecer uma premissa. Tais conjuntos são específicos da entrada/saída, o que significa que uma entrada/saída só pode estabelecer premissas com os conjuntos Fuzzy relativos a ela (tal fato será melhor discutido no Exemplo 3.6.10).

Exemplo 3.6.8. Considere os conjuntos \dot{E}, A descritos acima, bem como o valor de entrada e . Na fuzzyficação por *singleton* temos $\mu_{\dot{E}}(u) = \delta(u, e)$, e a função f utilizada é simplesmente o valor da pertinência de e na função μ_A , ou seja:

$$f(\mu_{\dot{E}}, \mu_A) = \mu_A(e) \tag{3.2}$$

(note que o valor e está associado ao conjunto \dot{E}). Porém, como podem existir diferentes tipos de fuzzyficação, podemos realizar uma generalização de f , dada por:

$$f(\mu_1, \mu_2) = \frac{\int_U \mu_1(x) \widehat{\star} \mu_2(x) dx}{\int_U \mu_1(x') dx'} \tag{3.3}$$

para alguma T-norma $\widehat{\star}$. Neste caso, o produto seria uma interessante escolha de T-norma, pois f resultaria em uma média ponderada dos valores de μ_2 em relação a μ_1 .

Note que, considerando para efeito de cálculo um *singleton* não mais como um delta de Kronecker, mas como um delta de Dirac, ou seja, uma distribuição, a função f descrita em (3.3) generaliza (3.2) pois, considerando $\mu_{\dot{E}}$ descrita acima e lembrando que $a \widehat{\star} 1 = a$:

$$\begin{aligned}
f(\mu_{\dot{E}}, \mu_A) &= \frac{\int_U \mu_{\dot{E}}(x) \widehat{\star} \mu_A(x) dx}{\int_U \mu_{\dot{E}}(x') dx'} \\
&= \int_U \mu_{\dot{E}}(x) \widehat{\star} \mu_A(x) dx \\
&= \mu_A(e)
\end{aligned}$$

Note que foi realizado um abuso de notação para a distribuição delta de Dirac δ_e em $e \in \mathbb{R}$:

$$g(e) = \delta_e(g) \triangleq \int_U \delta_e(x) \widehat{\star} g(x) dx$$

Exemplo 3.6.9. Considere $U = \{-3, -2, -1, 0, 1, 2, 3\}$, relacionado à entrada E , a fuzzyficação por *singleton* e o conjunto Fuzzy PZ definido em U que representa quando um número é praticamente zero, cuja função de pertinência é dada por

u	-3	-2	-1	0	1	2	3
$\mu_{\text{PZ}}(u)$	0	0	0.5	1	0.5	0	0

Tabela 5 – Pertinência do conjunto PZ

Considerando a função f definida em (3.3), temos que quando E assume valor de entrada $u^* = -1$, podemos dizer que

E é PZ (com pertinência 0.5)

3.6.5 Regras

Uma **Regra** na lógica Fuzzy é uma sentença com as seguintes componentes:

- Premissas
- Conectivos Lógicos entre as premissas (**OU/E**)
- Consequências das premissas (ou uma função no caso Takagi-Sugeno)
- Conectivos Lógicos entre as Consequências (se houver mais de uma)

(Obs.: O caso Takagi-Sugeno será explicado nas subseções seguintes)

Para aplicar uma regra, primeiro é necessário saber se ela é verdadeira. Por isso, as sentenças das regras são escritas condicionalmente:

se (Premissa 1) **E/OU** (Premissa 2) **E/OU** ... **E/OU** (Premissa N) **então**
 (Consequência 1) **E/OU** (Consequência 2) **E/OU** ... **E/OU** (Consequência M)

Em um sistema MISO apenas é necessária uma consequência, pois o objetivo é saber o efeito das entradas na única saída.

Exemplo 3.6.10. Deseja-se controlar a temperatura em uma sala através de um ar-condicionado que utiliza um sistema Fuzzy. Para tal, era possível considerar duas entradas:

- n : Número de pessoas na sala
- T_a : A temperatura no ambiente externo

A saída do sistema é a potência fornecida ao ar-condicionado:

- P : Potência fornecida ao ar-condicionado (quanto maior, mais refrigeração)

A entrada n está definida no conjunto universo $N = [0, 30]$, T_a está definida em $T = [25^\circ, 45^\circ]$ e P em $[300 \text{ W}, 3000 \text{ W}]$. A cada entrada e saída foram relacionados três conjuntos Fuzzy definidos nos respectivos conjuntos universo. Com relação à entrada n , os conjuntos Fuzzy são ilustrados na figura 31:

- **NP**: Número Pequeno
- **NM**: Número Médio
- **NG**: Número Grande

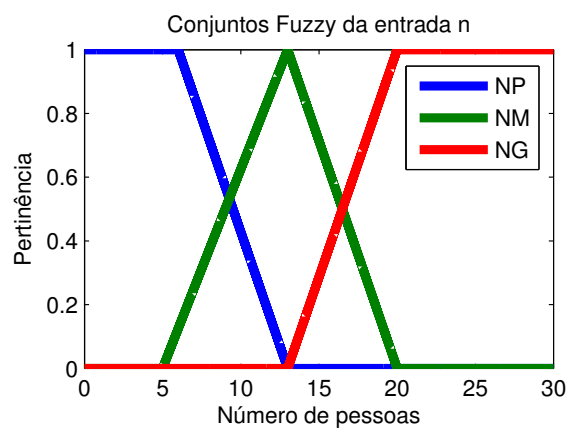


Figura 31 – Conjuntos Fuzzy respectivos à entrada n

Com relação à entrada T_a , os conjuntos Fuzzy são ilustrados na figura 32:

- **TB**: Temperatura Baixa
- **TM**: Temperatura Média

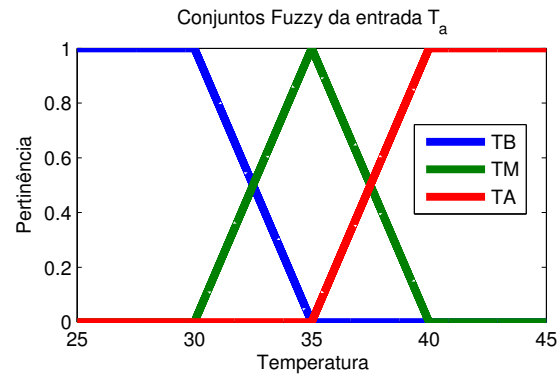


Figura 32 – Conjuntos Fuzzy respectivos à entrada T_a

- **TA:** Temperatura Alta

Com relação à saída P , os conjuntos Fuzzy são ilustrados na figura 33:

- **PB:** Potência Baixa
- **PM:** Potência Média
- **PA:** Potência Alta

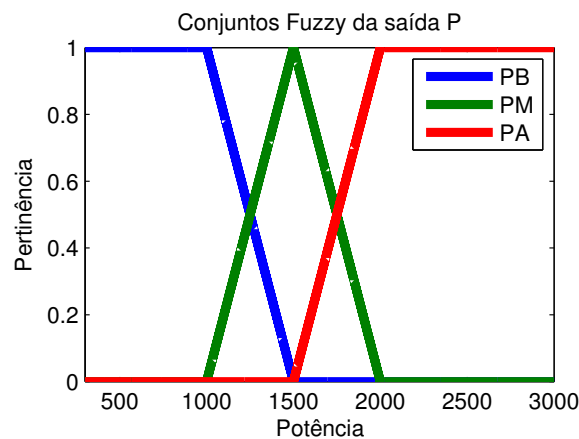


Figura 33 – Conjuntos Fuzzy respectivos à saída P

Note que cada conjunto Fuzzy é exclusivo de sua respectiva entrada/saída.

Podemos agora estabelecer um conjunto de regras para esse sistema Fuzzy com base nos conjuntos Fuzzy. Um exemplo de regra é:

se n é NG E T_a é TA então P é PA

Essa regra se traduz na seguinte lógica: se a sala possui um alto número de pessoas e a temperatura externa é muito alta, então deve ser fornecida uma potência alta ao ar-condicionado para que ele seja capaz de resfriar a sala. Neste caso em que há 2 entradas

e 1 saída, podemos montar uma tabela de regras. Na seguinte tabela de regras todas as relações entre as duas entradas são através do conectivo lógico **E** e o resultado da combinação das duas como a consequência de saída (**se** n é F_1 **E** T_a é F_2 **então** P é F_3).

E	NP	NM	NG
TB	PB	PB	PM
TM	PB	PM	PA
TA	PM	PA	PA

Tabela 6 – Consequência de Saída (Potência PX) para as Combinações de Número de Pessoas (NX) e Temperatura (TX)

3.6.6 Operação Fuzzy implicação

Na subseção anterior foi definido o significado da expressão " A é B " no sentido da lógica Fuzzy. Nesta subseção será definido o significado de $p \implies c$, ou em outras palavras: "se p então c ", sendo que neste trabalho será abordada a inferência baseada em cada regra individualmente.

Esta "operação" de implicar (\implies) é utilizada em um conjunto de premissas para implicar uma função de saída (no caso de um modelo Takagi-Sugeno) ou simplesmente uma consequência. Sendo assim, a implicação será dividida em dois casos: o caso Takagi-Sugeno e o caso consequência.

Takagi-Sugeno. Em um modelo Takagi-Sugeno, as premissas implicam em uma função da saída f em relação às entradas. Por exemplo, considere um sistema com N entradas E_1, \dots, E_N , com respectivos conjuntos universo U_1, \dots, U_N , que tenha N conjuntos Fuzzy A_1, \dots, A_N , onde A_i é relativo a E_i , e suponha que valha a premissa

$$E_1 \text{ é } A_1 \text{ E } E_2 \text{ é } A_2 \text{ E } \dots \text{ E } E_N \text{ é } A_N$$

(note que foi introduzida a operação lógica **E** entre as premissas. Outra condição lógica, como **OU**, poderia ser utilizada). Considere agora a saída S do sistema, cujo conjunto universo é V , e uma função $f^* : U_1 \times \dots \times U_N \rightarrow V$. Então a implicação em uma regra no modelo Takagi-Sugeno é

$$\text{se } E_1 \text{ é } A_1 \text{ E } E_2 \text{ é } A_2 \text{ E } \dots \text{ E } E_N \text{ é } A_N \text{ ENTÃO } f = f^*$$

Cada regra no modelo Takagi-Sugeno possui uma função f^* específica. Portanto, se existirem N regras, existem N funções diferentes, uma para cada regra. Essa função, assim como a pertinência de cada premissa, será utilizada no processo para determinar o valor de saída, que será visto numa subseção posterior. O modelo Takagi-Sugeno originalmente possui a função f^* como uma combinação linear das entradas somada com uma constante.

Consequência de saída. Considere um sistema MISO com as entradas descritas acima e que possui uma saída S , definida no conjunto universo V , e que possua M conjuntos Fuzzy relacionados a ela. Então uma implicação em uma regra sobre o conjunto Fuzzy de saída B_k (com função de pertinência μ_{B_k} e definido em V) é do tipo

$$\text{se } E_1 \text{ é } A_1 \text{ E } E_2 \text{ é } A_2 \text{ E } \cdots \text{ E } E_N \text{ é } A_N \text{ ENTÃO } S \text{ é } B_k$$

Como visto, cada premissa i (relativa a A_i) possui um grau de pertinência μ_i , e logo podemos determinar a pertinência do conjunto de premissas:

$$\mu_1 \underbrace{\star}_{\star} \mu_2 \underbrace{\star}_{\star} \cdots \underbrace{\star}_{\star} \mu_N = \underbrace{\star}_{k=1}^N \mu_k = \mu_T \quad (3.4)$$

(caso sejam utilizados conectivos **OU**, devemos realizar a operação $\underbrace{\star}_{\star}$). Portanto, a pertinência total de todas as premissas vale μ_T . Seja P a premissa equivalente a tal conjunto de premissas. Logo, temos que $P \implies S \text{ é } B_K$ com a pertinência de P sendo μ_T .

Agora é necessário definir o significado da implicação. De maneira geral, a implicação \implies estabelece uma função $f \implies$ entre o conjunto Fuzzy B_k em função da pertinência da premissa P . Denotando \mathbb{F} como o conjunto dos conjuntos Fuzzy, temos

$$f \implies : \{[0, 1], \mathbb{F}\} \rightarrow \mathbb{F}$$

$$f \implies (\mu_T, B_k) = \tilde{B}_k$$

$$\tilde{B}_k = \{(v, \mu_{\tilde{B}_k}(v)), v \in V\}$$

Portanto, a implicação altera o conjunto Fuzzy de saída B_k em função da pertinência total das premissas μ_T . Esse conjunto Fuzzy resultante irá, em conjunto com os resultantes das outras regras, determinar o valor de saída de S , como será mostrado na subseção posterior.

Exemplo 3.6.11. Implicação Mandani. Na implicação Mandani, a função $f \implies$ é simplesmente gerar um conjunto Fuzzy cuja função de pertinência é o mínimo entre μ_T e μ_{B_k} .

$$\text{se } P \text{ então } S \text{ é } B_K$$

Resultado:

$$f \implies (\mu_T, B_k) = \{(v, \mu_{\tilde{B}_k}(v)), v \in V\}$$

sendo

$$\mu_{\tilde{B}_k}(v) = \min(\mu_T, \mu_{B_k}(v))$$

Tal implicação equivale a "ceifar" a função de pertinência de B_k no valor μ_T . Para ilustrar tal implicação, considere $\mu_T = 0.5$ e B_k com domínio $[0, 10]$ e função de pertinência como na figura 34.

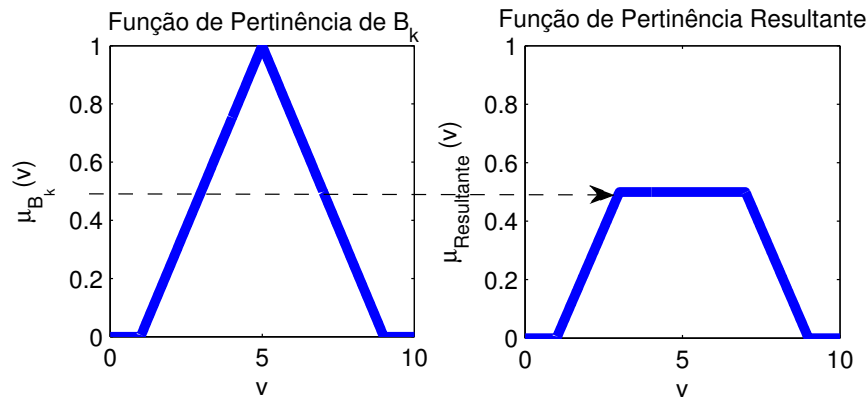


Figura 34 – Função de Pertinência $\mu_{\tilde{B}_k}$ via implicação Mandani em comparação com μ_{B_k}

Exemplo 3.6.12. Implicação Gödel. Na implicação Gödel, a função $f \implies$ aplicada a μ_T e B_k resulta no seguinte conjunto Fuzzy:

$$\tilde{B}_k = \{(v, \mu_{\tilde{B}_k}(v)), v \in V\}$$

$$\mu_{\tilde{B}_k}(v) = \begin{cases} 1 & \text{se } \mu_T \leq \mu_{B_k}(v) \\ \mu_{B_k}(v) & \text{caso contrário} \end{cases}$$

Para ilustrar tal implicação, considere $\mu_T = 0.5$ e B_k com domínio $[0, 10]$, como no exemplo 3.6.11 e função de pertinência como na figura 35.

3.6.7 Agregação

Num modelo Fuzzy existem várias regras e o resultado de saída final deve considerar todas as regras ativas. Como há diferenças entre o modelo Consequência de Saída e o modelo Takagi-Sugeno, novamente será feita uma separação entre os dois. Porém, a agregação no modelo Takagi-Sugeno será deixada para a próxima subseção, pois está extremamente ligada ao processo de Defuzzyficação.

Consequência de Saída. Considere o seguinte sistema Fuzzy:

- N entradas E_1, \dots, E_N

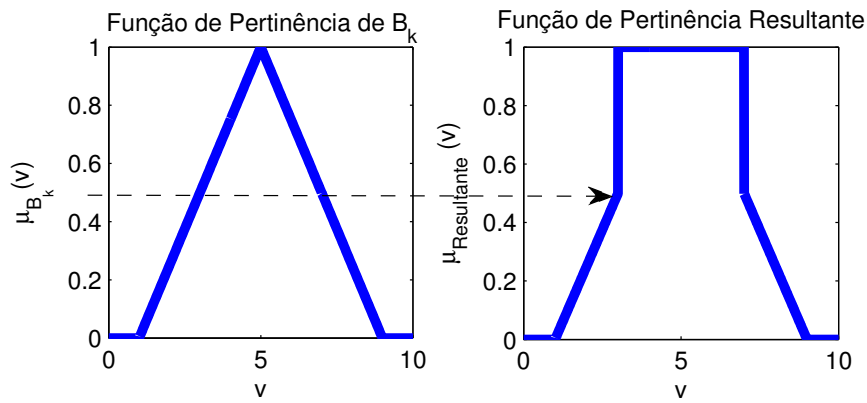


Figura 35 – Função de Pertinência $\mu_{\tilde{B}_k}$ via implicação Gödel em comparação com μ_{B_k}

- 1 saída S
- R regras
- $N \cdot R$ conjuntos Fuzzy relacionados a cada entrada, denotando por A_i^j o conjunto relativo à regra j e entrada i , com $1 \leq i \leq N$ e $1 \leq j \leq R$.
- M conjuntos Fuzzy relacionados à saída S , denotando B^j o conjunto Fuzzy de saída relativo à regra j , com $1 \leq j \leq R$

O conjunto de regras pode ser escrito como

$$\text{se } E_1 \text{ é } A_1^1 \text{ E } \dots \text{ E } E_N \text{ é } A_N^1 \text{ ENTÃO } S \text{ é } B^1$$

$$\text{se } E_1 \text{ é } A_1^2 \text{ E } \dots \text{ E } E_N \text{ é } A_N^2 \text{ ENTÃO } S \text{ é } B^2$$

⋮

$$\text{se } E_1 \text{ é } A_1^R \text{ E } \dots \text{ E } E_N \text{ é } A_N^R \text{ ENTÃO } S \text{ é } B^R$$

Considere agora os conjuntos Fuzzy resultantes das implicações de cada regra, denotando \tilde{B}^j o conjunto Fuzzy resultante da implicação na regra j . A agregação consiste em utilizar a informação proveniente de todos os \tilde{B}^j resultantes para criar um último conjunto Fuzzy, proveniente da agregação de todos eles. Portanto, é necessário definir uma função f_A para realizar tal operação. Denotando por \mathbb{F} o conjunto de conjuntos Fuzzy:

$$f_A : \mathbb{F}^R \rightarrow \mathbb{F}$$

onde \mathbb{F}^R é o produto cartesiano de R (número de regras) conjuntos \mathbb{F} . O conjunto Fuzzy gerado por f_A será denotado \tilde{B}_T .

Exemplo 3.6.13. Agregação Mandani. A agregação Mandani utiliza a união dos conjuntos \tilde{B}^j para agregá-los:

$$\tilde{B}_T = f_A(\tilde{B}^1, \dots, \tilde{B}^R) = \bigcup_{k=1}^R \tilde{B}^k = \underbrace{\star}_{k=1}^R \tilde{B}^k$$

Geralmente utiliza-se a S-norma máximo, que é ilustrada na Figura 36 para dois conjuntos \tilde{B}^1 e \tilde{B}^2 :

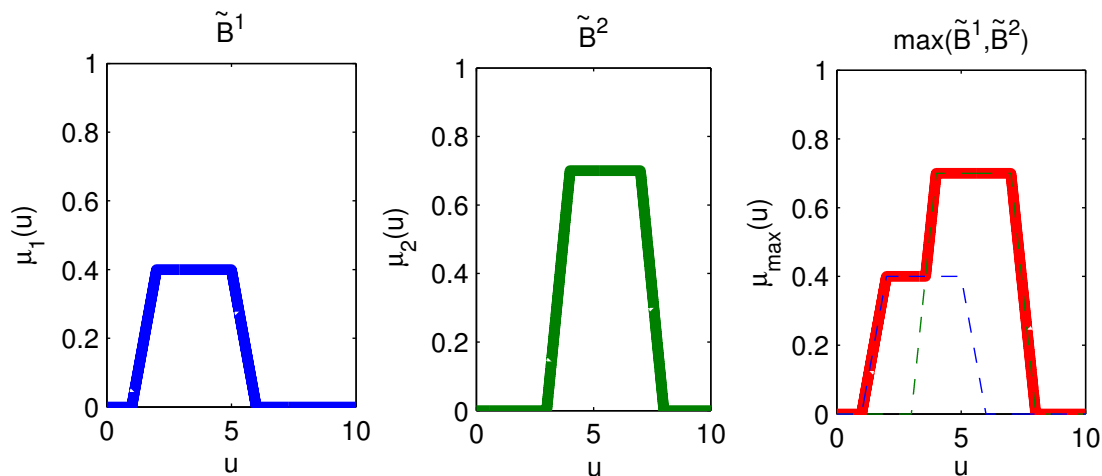


Figura 36 – Função de agregação pelo máximo

Exemplo 3.6.14. Agregação Gödel (inferência feita individualmente em cada regra). Na agregação Gödel, utiliza-se a interseção dos conjuntos \tilde{B}^j para agregá-los:

$$\tilde{B}_T = f_A(\tilde{B}^1, \dots, \tilde{B}^R) = \bigcap_{k=1}^R \tilde{B}^k = \underbrace{\star}_{k=1}^R \tilde{B}^k$$

Normalmente utiliza-se a T-norma mínimo, que é ilustrada na Figura 37 para dois conjuntos \tilde{B}^1 e \tilde{B}^2 :

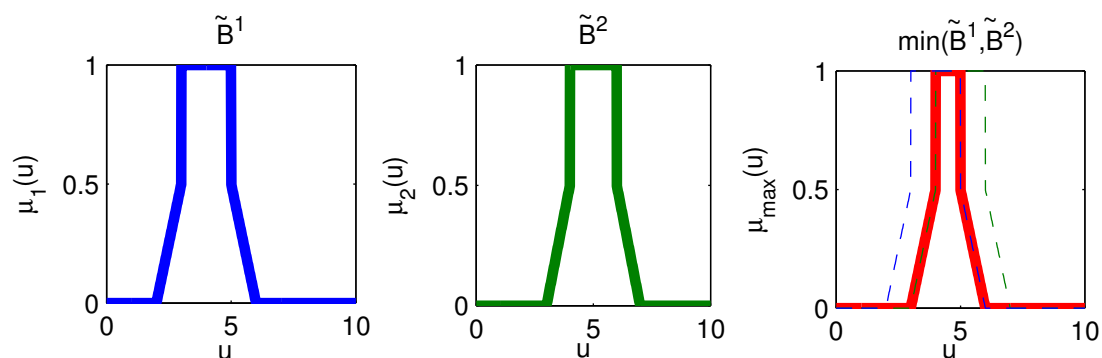


Figura 37 – Função de agregação pelo mínimo

3.6.8 Defuzzyficação

Após o estágio de agregação, temos como resultado das entradas um conjunto Fuzzy de saída. Porém, é necessário extrair de tal conjunto o valor de saída desejado no conjunto universo da saída. Para tal, é realizada a defuzzyficação do conjunto. Sendo assim, para um modelo Consequência de Saída com uma saída S , definida no conjunto universo V , podemos definir a defuzzyficação como uma função f_D tal que

$$f_D : \mathbb{F} \rightarrow V$$

Já para um modelo Takagi-Sugeno, a defuzzyficação é aplicada nas funções de saída, como pode ser visto no próximo exemplo.

Defuzzyficação Takagi-Sugeno. Como visto, no modelo Takagi-Sugeno, a cada regra está associada uma função f^* . Considere então o seguinte sistema:

- N entradas E_1, \dots, E_N , definidas em U_1, \dots, U_N respectivamente
- 1 saída S
- R regras
- $N \cdot R$ conjuntos Fuzzy relacionados a cada entrada, denotando por A_i^j o conjunto relativo à regra j e entrada i , com $1 \leq i \leq N$ e $1 \leq j \leq R$.
- R funções de saída, sendo f_j^* relacionada à regra j , e ainda $f_j^* : U_1 \times \dots \times U_N \rightarrow V$ para todo j .

Considere $U_T = U_1 \times \dots \times U_N$. A defuzzyficação das regras é dada por uma função f_D tal que

$$f_D : U_1 \times \dots \times U_N \rightarrow V$$

$$f_D(u) = \frac{\sum_{k=1}^R f_k^*(u) \mu_T^k}{\sum_{k=1}^R \mu_T^k}, \quad u \in U_T$$

onde μ_T^j é a pertinência do conjunto de premissas da regra j , como visto na equação (3.4).

Defuzzyficação do modelo Consequência de Saída. Como visto, no modelo Consequência de Saída, é gerado um conjunto Fuzzy de saída \tilde{B}_T . Existem vários métodos de Defuzzyficação nesse caso (várias f_D), e aqui serão apresentados dois deles.

- Centro de Área/Gravidade

O método de defuzzyficação pelo Centro de Área é um dos mais conhecidos na literatura (DRIANKOV; HELLENDORRN; REINFRANK, 1996). De fato, ele consiste em tomar a coordenada $v \in V$ do centro da área da função de pertinência de \tilde{B}_T como a saída do sistema:

$$v_s = \frac{\int_V v \mu_{\tilde{B}_T}(v) dv}{\int_V \mu_{\tilde{B}_T}(v') dv'}$$

onde v_s é o valor da saída S do sistema e V o conjunto universo de S .

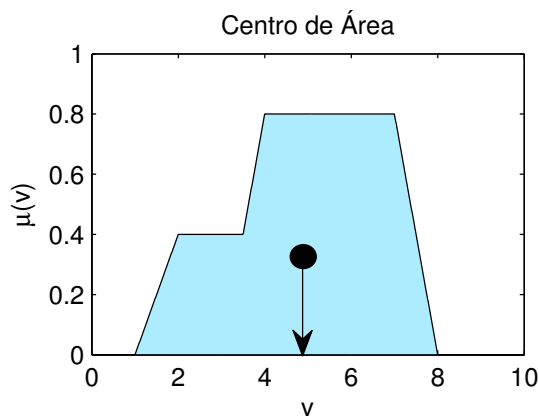


Figura 38 – Obtenção da coordenada v do centro de área da função de pertinência de \tilde{B}_T

- **Primeiro Ponto de Máximo / Último Ponto de Máximo**

Considere o conjunto universo ordenado V de \tilde{B}_T e então tome o conjunto $H \subset V$ tal que

$$H = \{v \in V : \mu_{\tilde{B}_T}(v) = \text{alt}(\tilde{B}_T) = \sup_{v \in V} \mu_{\tilde{B}_T}(v)\}$$

A Defuzzyficação por Primeiro Ponto de Máximo consiste em apenas tomar

$$v_s = \inf H$$

Já a Defuzzyficação por Último Ponto de Máximo consiste em tomar

$$v_s = \sup H$$

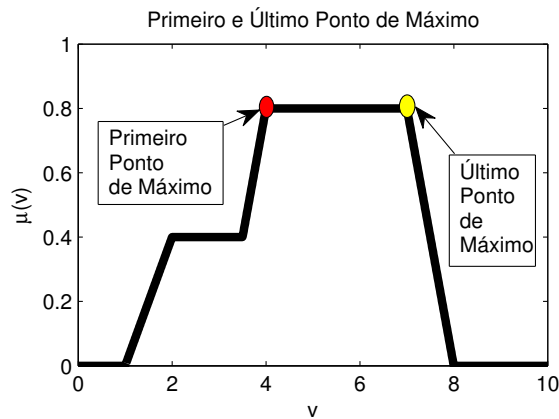


Figura 39 – Representação do Primeiro e Último Ponto de Máximo da função de pertinência de \tilde{B}_T

3.7 Clusterização via Método da Montanha

Na seção anterior foi introduzida a Teoria Fuzzy, mostrando como funciona um sistema Fuzzy e sua lógica. Porém, existem várias possibilidades para se construir uma estrutura Fuzzy, desde a escolha das funções de pertinência até as regras entre elas.

Nesta seção será apresentado um método que consiste em gerar um sistema Fuzzy a partir de um conjunto de dados de entrada relacionados à sua respectiva saída. Tal método é uma extensão do Método da Montanha, proposto por Ronald R. Yager e estendido em (CHIU, 1994), sendo esta última a utilizada para compor esta seção.

Basicamente, o método funciona da seguinte maneira: um conjunto de dados de entrada e saída é processado e nele são definidos "*clusters centers*", aqui chamados de núcleos, os quais representam dados que possuem as características predominantes de tal conjunto. Através deles são gerados conjuntos Fuzzy e regras, que são otimizadas através do método dos mínimos quadrados.

3.7.1 Obtenção dos núcleos

Considere um sistema com M entradas Y_1, \dots, Y_M e N saídas Z_1, \dots, Z_N . Considere então a entrada como um vetor $y \in \mathbb{R}^M$, chamado de dado de entrada, e a saída como um vetor $z \in \mathbb{R}^N$, chamado de dado de saída, sendo cada dimensão i respectiva à entrada/saída i . Considere também um conjunto de n dados de entrada $Y = \{y_1, y_2, \dots, y_n\}$ e n dados de saída $Z = \{z_1, z_2, \dots, z_n\}$, sendo a saída z_i relativa à entrada y_i . Seja X o espaço formado pela concatenação de Y e Z , dado por

$$X = \{(y_k, z_k) : y_k \in Y, z_k \in Z\}, \quad X \subset \mathbb{R}^{M+N}$$

Sendo assim, concatenando os n dados de entrada com os n dados de saída, é gerado um conjunto de n dados $X = \{x_1, x_2, \dots, x_n\}$. Sem perda de generalidade, considere que os

dados estão normalizados em cada dimensão, ou seja

$$x_i = (a_{1i}, a_{2i}, \dots, a_{(M+N)i}) \quad : \quad |a_{ki}| \leq 1 \quad \forall i \in \{1, \dots, n\}, \quad \forall k \in \{1, \dots, M+N\}$$

Inicialmente, todo x_i é considerado um núcleo em potencial. Definimos então o potencial P_i de cada x_i , dado por

$$P_i = \sum_{j=1}^n e^{-\alpha \|x_i - x_j\|^2}$$

onde

$$\alpha = \frac{4}{r_a^2}$$

sendo $r_a \in \mathbb{R}, r_a > 0$. A constante r_a será chamada de **Raio de Influência**.

Tal função P consegue quantificar a distância de um ponto x_i a todos os outros. Sendo assim, os que tiverem uma maior aglomeração de pontos em torno deles terão um potencial P maior, como ilustrado na figura 40.

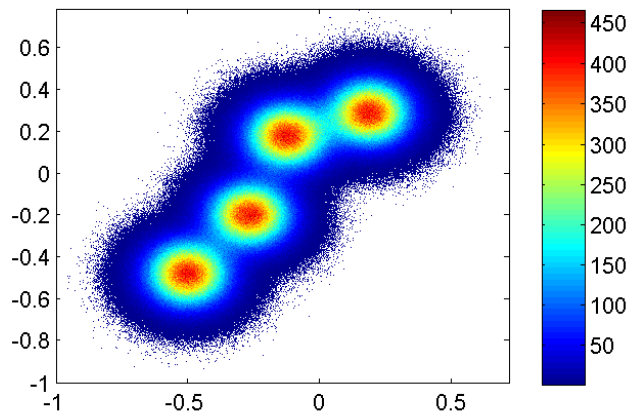


Figura 40 – Ilustração do potencial P gerada pelo histograma de um conjunto de pontos bidimensionais. Quanto maior o valor na barra (referente a uma cor), maior o potencial correspondente do ponto.

O valor da constante r_a é importante para definir o raio da vizinhança de um ponto, no sentido de que quanto maior o valor de r_a , maior a contribuição de pontos distantes ao potencial de um ponto.

Após obter o potencial de todos os pontos, escolhemos aquele com maior potencial para ser o primeiro núcleo, denotado x_1^* . Caso haja mais de um ponto, escolhemos qualquer um deles:

$$x_1^* \in \{x_k \in X : P_k = \max_{1 \leq j \leq n} P_j\}$$

sendo P_1^* o potencial de x_1^* .

Agora o objetivo é determinar outro núcleo longe do já escolhido, para conseguir generalizar com maior eficácia o conjunto de dados X por meio desses núcleos. Porém, é necessário atualizar os potenciais, fazendo com que aqueles pontos que estavam próximos a x_1^* percam potencial. Para tal, considere a atualização do potencial P_i de x_i de acordo com a expressão

$$P_i \Leftarrow P_i - P_1^* e^{-\beta \|x_i - x_1^*\|^2}$$

onde

$$\beta = \frac{4}{r_B^2}$$

sendo $r_B \in \mathbb{R}, r_B > 0$ uma constante. Denominamos a razão $\frac{r_b}{r_a}$ por **Fator de Esmagamento**.

O valor da constante r_B definirá o raio em volta de x_1^* que perderá um valor significativo de potencial: quanto maior o valor de r_B , mais pontos distantes de x_1^* perderão uma quantidade significativa de potencial. Note que o potencial de x_1^* passará a ser 0.

Com o potencial de todos os pontos atualizado, escolhemos o segundo núcleo x_2^* da mesma forma que foi escolhido o primeiro:

$$x_2^* \in \{x_k \in X : P_k = \max_{1 \leq j \leq n} P_j\}$$

sendo P_2^* o potencial de x_2^* . Novamente, recalculamos os potenciais de todos os pontos. Após repetir esse processo c vezes, obtemos o conjunto $C = \{x_1^*, x_2^*, \dots, x_c^*\}$. O número c será definido de acordo com um critério de parada, descrito a seguir.

3.7.2 Critério de parada

No critério de parada é preciso definir duas constantes: $\bar{\epsilon}$, a **Taxa de Aceitação**; e $\underline{\epsilon}$, a **Taxa de Rejeição**. A única propriedade entre os dois é $\underline{\epsilon} < \bar{\epsilon}$.

Enquanto

$$P_k^* > \bar{\epsilon} P_1^*$$

o ponto x_k^* será considerado um núcleo. Se

$$P_k^* < \underline{\epsilon}P_1^*$$

então rejeitamos x_k^* como núcleo e encerramos o processo, com $c = k - 1$. Na condição em que $\underline{\epsilon}P_1^* < P_k^* < \bar{\epsilon}P_1^*$, o ponto x_k^* poderá ou não ser aceito como um núcleo.

Para decidir se ele será aceito, considere $d_{\min} = \min_{1 \leq i \leq k-1} \|x_k^* - x_i^*\|$. Se

$$\frac{d_{\min}}{r_a} + \frac{P_k^*}{P_1^*} \geq 1$$

então x_k^* será aceito como um núcleo e o processo continuará. Caso contrário, x_k^* será rejeitado como um núcleo e seu potencial passará a ser 0 ($P_k^* = 0$). Ainda, ele deixará de ser denotado por x_k^* e o próximo ponto com maior potencial será o novo x_k^* . Então, testamos novamente o critério de parada.

3.7.3 Determinação do modelo Fuzzy

Com o conjunto $C = \{x_1^*, x_2^*, \dots, x_c^*\}$ definido, será necessário separá-lo nos dados de entrada e saída:

$$Y^* = \{y_k^* \in \mathbb{R}^M : x_k^* = (y_k^*, z_k^*), 1 \leq k \leq c\}$$

$$Z^* = \{z_k^* \in \mathbb{R}^N : x_k^* = (y_k^*, z_k^*), 1 \leq k \leq c\}$$

Cada $x_k^* = (y_k^*, z_k^*)$ comporá de uma regra no sistema Fuzzy, no sentido de que a entrada y_k^* implicará na saída z_k^* . O modelo utilizado será o Takagi-Sugeno. Portanto, haverá c regras ao todo, sendo que as entradas serão consideradas como *singletons*. Uma regra i qualquer será da forma

se Y_1 é A_{i1} **E** Y_2 é A_{i2} **E** \dots **E** Y_M é A_{iM} **então** Z_1 é B_{i1} **E** Z_2 é B_{i2} **E** \dots **E** Z_N é B_{iN}

onde Y_j é a entrada j do sistema ($1 \leq j \leq M$), Z_l a saída l do sistema ($1 \leq l \leq N$) e os termos B_{il} são funções (modelo Takagi-Sugeno). A função de pertinência de A_{ji} é dada por

$$A_{ij}(u) = e^{-\alpha(u-y_{ij}^*)^2}$$

onde y_{ij}^* é o termo pertencente a \mathbb{R} relativo à entrada j em y_i^* . A função B_{il} é dada simplesmente por uma constante

$$B_{il} = z_{il}^*$$

onde z_i^* é o termo pertencente a \mathbb{R} relativo à saída l em z_i^* . Além disso, o conectivo \mathbf{E} é realizado pela T-norma produto. A saída $z \in \mathbb{R}^N$ do sistema para uma entrada $y \in \mathbb{R}^M$ é então dada por

$$z = \frac{\sum_{i=1}^c \mu_i(y) z_i^*}{\sum_{i=1}^c \mu_i(y)} \quad (3.5)$$

onde

$$\mu_i(y) = e^{-\alpha \|y - y_i^*\|^2}$$

Nesta etapa, (CHIU, 1994) propôs um método para otimizar os parâmetros do sistema Fuzzy, o qual será descrito a seguir.

3.7.4 Otimização do Modelo Fuzzy

Considere na equação (3.5) que z_i^* é uma função linear das variáveis $y \in \mathbb{R}^M$ de entrada, dada por

$$z_i^* = G_i y + h_i \quad (3.6)$$

onde G_i é uma matriz constante $N \times M$ e h_i um vetor constante $N \times 1$. Considere também que

$$\rho_i = \frac{\mu_i}{\sum_{j=1}^c \mu_j} \quad (3.7)$$

Substituindo (3.7) e (3.6) em (3.5):

$$z = \sum_{i=1}^c \rho_i (G_i y + h_i) \quad (3.8)$$

Escrevendo (3.8) em forma matricial, denotando a transposta de uma matriz A por A^T :

$$z^T = [\rho_1 y^T \quad \rho_1 \quad \cdots \quad \rho_c y^T \quad \rho_c] \begin{bmatrix} G_1^T \\ h_1^T \\ \vdots \\ G_c^T \\ h_c^T \end{bmatrix}$$

Agora, considerando um conjunto o dados de entrada $Y = \{y_1, \dots, y_n\}$, é possível escrever

$$\begin{bmatrix} z_1^T \\ \vdots \\ z_n^T \end{bmatrix} = \begin{bmatrix} \rho_{11}y^T & \rho_{11} & \cdots & \rho_{c1}y^T & \rho_{c1} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ \rho_{1n}y^T & \rho_{1n} & \cdots & \rho_{cn}y^T & \rho_{cn} \end{bmatrix} \begin{bmatrix} G_1^T \\ h_1^T \\ \vdots \\ G_c^T \\ h_c^T \end{bmatrix} \quad (3.9)$$

onde ρ_{ij} representa ρ_i calculada para a entrada y_j . Denotando

$$B = \begin{bmatrix} z_1^T \\ \vdots \\ z_n^T \end{bmatrix}$$

$$A = \begin{bmatrix} \rho_{11}y^T & \rho_{11} & \cdots & \rho_{c1}y^T & \rho_{c1} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ \rho_{1n}y^T & \rho_{1n} & \cdots & \rho_{cn}y^T & \rho_{cn} \end{bmatrix}$$

$$X = \begin{bmatrix} G_1^T \\ h_1^T \\ \vdots \\ G_c^T \\ h_c^T \end{bmatrix}$$

note que a equação (3.9) é da forma

$$AX = B$$

passível de aplicar o método dos mínimos quadrados, já que A é uma matriz constante e X possui os parâmetros da linearização de z que desejamos encontrar. Portanto, desejamos encontrar um X_{\min} tal que

$$X_{\min} = \min_{X \in \mathbb{R}^{c(M+1)N}} \{ \|AX - B\|^2 \}$$

que pelo método dos mínimos quadrados resulta em

$$X_{\min} = (A^T A)^{-1} A^T B.$$

Computacionalmente, o cálculo de $(A^T A)^{-1}$ pode ser custoso. Por isso, (CHIU, 1994) cita um método iterativo de se calcular tal resultado, que pode ser encontrado no próprio (CHIU, 1994).

3.8 Algoritmo Genético

O Algoritmo Genético é um processo que visa otimizar parâmetros de um sistema. Ele possui esse nome pois, intuitivamente, se baseia na ideia de uma população de indivíduos que sofrem mutações e interagem entre si ao longo de gerações para que seu desempenho no sistema possa melhorar.

Tal algoritmo é de extrema importância no método desenvolvido nessa dissertação, pois nele existem vários parâmetros que mudam drasticamente quando outros são alterados. Sendo assim, executar uma otimização manualmente seria algo extremamente demorado, além de não ser intuitivo, pois alguns parâmetros apresentam mudanças bruscas no desempenho com pequenas variações (na prática, o sistema não é contínuo com relação a eles).

As próximas seções visam traduzir os conceitos linguísticos, como população e mutações, para os práticos utilizados em um sistema computacional.

3.8.1 Indivíduo

Um **indivíduo** em um algoritmo genético representa um conjunto ordenado I de n parâmetros α_k do sistema S que desejamos otimizar:

$$I = (\alpha_1, \alpha_2, \dots, \alpha_n) : \alpha_k \in S, 1 \leq k \leq n \quad (3.10)$$

Em (3.10) I é representado de uma maneira mais abstrata, podendo os parâmetros α_k terem naturezas diferentes: α_1 poderia ser um número real, α_2 um conjunto de letras, α_3 um número natural, e assim por diante. O sistema S seria então um conjunto de classes que representam cada parâmetro, onde $\alpha_k \in S$ significa que o parâmetro α_k é um dos parâmetros do sistema S .

É importante citar que cada parâmetro α_k deve estar definido em um conjunto Ω_k que geralmente é limitado, impondo restrições sobre os valores de cada parâmetro. Ainda assim, o número n de parâmetros escolhidos de S não pode ser exagerado, pois caso contrário, o número de indivíduos possíveis em $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$ seria elevado, o que resultaria num tempo alto computacionalmente.

Exemplo 3.8.1. Considere que o sistema S é uma empresa em construção que deseja minimizar os custos com ar-condicionado em relação à produtividade da unidade. Para isso, desejamos otimizar o tamanho das 5 salas a serem construídas, além do número de pessoas em cada sala. Poucas pessoas e salas menores reduziriam o consumo, porém também diminuiria a produtividade da unidade. Além disso, cada sala deve ter no mínimo 40 m² e no máximo 240 m², com cada sala contendo no mínimo 3 pessoas e no máximo 18.

Um indivíduo do algoritmo genético proposto para tal problema poderia ser da forma

Sala	1	2	3	4	5
Tamanho (m ²)	A_1	A_2	A_3	A_4	A_5
Número de Pessoas	n_1	n_2	n_3	n_4	n_5

Tabela 7 – Notação dos Parâmetros Tamanho e Número de Pessoas referentes a cada Sala

$$I = (A_1, A_2, A_3, A_4, A_5, n_1, n_2, n_3, n_4, n_5)$$

satisfazendo

$$40 \leq A_k \leq 240$$

$$3 \leq n_k \leq 18$$

Portanto, $\Omega_i = [40, 240]$ para $1 \leq i \leq 5$ e $\Omega_j = \{3, 4, \dots, 17, 18\}$ para $6 \leq j \leq 10$. Um possível exemplo numérico de indivíduo seria

$$I = (60.3, 49.7, 203, 104, 110, 5, 9, 13, 8, 6)$$

3.8.2 População

População é um conjunto ordenado (pela performance) de indivíduos I do sistema S . Todos os indivíduos devem ter os mesmos tipos de parâmetros para que possam interagir entre si. Geralmente, o número de indivíduos de uma população é fixo durante todo o processo. Sendo assim, podemos definir uma população P de m indivíduos por

$$P = (I_1, I_2, I_3, \dots, I_m)$$

Exemplo 3.8.2. Utilizando o mesmo sistema descrito no exemplo 3.8.1, uma possível população seria

$$P = (I_1, I_2, I_3)$$

$$I_1 = (60.3, 49, 80, 100.9, 110, 5, 7, 4, 8, 6)$$

$$I_2 = (40, 60, 89, 80, 230, 4, 8, 7, 5, 16)$$

$$I_3 = (45, 75, 50, 180, 50, 4, 8, 3, 5, 6)$$

3.8.3 Geração

Uma **geração** está relacionada ao número de vezes que uma população foi submetida às operações do Algoritmo Genético. A população inicial é definida como a população da primeira geração. Após submetê-la às operações de mutação, *cross-over* e eliminação (definidas nas próximas subseções), obtemos uma nova população, que pertence à segunda geração. Indutivamente, a população da geração j , após as operações do algoritmo, torna-se a população da geração $j + 1$.

O intuito é sempre obter populações cujos indivíduos possuam melhor performance em relação à geração anterior. Portanto, o melhor indivíduo é sempre salvo, para que no final do algoritmo seja fornecido o melhor resultado dos computados.

3.8.4 Operações em indivíduos

Nesta subseção serão descritas as operações utilizadas pelo Algoritmo Genético para modificar os indivíduos de uma dada população, com o intuito de obter novos indivíduos com melhor desempenho que os anteriores e eliminar aqueles que possuem baixo desempenho.

3.8.5 Mutação

A **mutação** é uma operação feita em cada indivíduo separadamente. Ela consiste em mudar um ou mais parâmetros do indivíduo para um elemento aleatório do respectivo conjunto Ω ao qual o parâmetro pertence. Quais parâmetros serão mudados também é uma escolha aleatória, tendo a possibilidade de nenhum ser mudado. É inspirada na mutação genética, estudada no ramo da biologia, na qual um dos nucleotídeos presentes no DNA é substituído por outro, alterando a sequência genética.

Nesta operação definimos uma probabilidade de mutação p_M . Sempre que a mutação é aplicada em um indivíduo, gera-se um número $\delta_k \in [0, 1]$ aleatório para cada parâmetro α_k , e se $\delta_k > p_M$ então aquele parâmetro será substituído por um novo α_k^R aleatório, escolhido em algum subconjunto de Ω_k . Neste trabalho, α_k^R pode ser qualquer valor em Ω_k .

Denotando por f_M a função de mutação e $\Lambda = \Omega_1 \times \cdots \times \Omega_n$ o conjunto de possíveis indivíduos, podemos representar a mutação da seguinte maneira:

$$f_M : \Lambda \rightarrow \Lambda$$

de modo que, dado $I = (\alpha_1, \dots, \alpha_n)$, temos

$$f_M(I) = (\alpha_1^M, \dots, \alpha_n^M)$$

onde

$$\alpha_k^M = \begin{cases} \alpha_k & \text{se } \delta_k > p_M \\ \alpha_k^R & \text{se } \delta_k \leq p_M \end{cases}$$

sendo α_k^R um elemento aleatório de $B(\alpha_k, \epsilon, \Omega_k) \subseteq \Omega_k$, onde $B(\alpha_k, \epsilon, \Omega_k)$ representa uma vizinhança de α_k de raio ϵ para algum $\epsilon > 0$ definido e alguma métrica $\|\cdot\|_{\Omega_k}$. O raio ϵ pode ser fixo ou variante de acordo com o número de gerações.

Exemplo 3.8.3. Novamente, no sistema S do exemplo 3.8.1, considere o indivíduo

$$I_1 = (60.3, 49.7, 73, 190, 110, 5, 7, 14, 8, 6).$$

Considere a probabilidade de mutação $p_M = 0.3$ e $B(\alpha_k, \epsilon, \Omega_k) = \Omega_k$, ou seja, α_k^R pode assumir qualquer valor em Ω_k quando ocorre mutação. Considere ainda que foram gerados os seguintes valores de δ_k :

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}
0.4	0.1	0.2	0.9	0.7	0.33	0.25	0.05	0.45	0.5

Tabela 8 – Valores de δ_k , relativos à Probabilidade de Mutação de cada Parâmetro

Ao aplicar a função f_M em I_1 , um indivíduo que pode ser gerado é:

$$f_M(I_1) = (60.3, 110, 200, 190, 110, 5, 8, 11, 8, 6)$$

Note que apenas para $k \in \{2, 3, 7, 8\}$ houve alteração no valor de α_k .

3.8.6 Cross-Over

A operação **cross-over** é aplicada a um par indivíduos. Também é inspirada em sua definição na biologia, na qual dois cromossomos trocam material genético, resultando em dois outros cromossomos constituídos da combinação dos anteriores.

O *cross-over* em um Algoritmo Genético consiste em substituir um certo número de parâmetros (não todos) de um indivíduo I_{2j} pelos respectivos parâmetros de um indivíduo I_{2j+1} (devido a isso, o tamanho da população precisa ser par). O indivíduo I_{2j+1} também substitui os mesmos parâmetros pelos do I_{2j} anteriores à troca. O número de parâmetros a serem trocados também é escolhido aleatoriamente. Além disso, o *cross-over* só é realizado em um par de indivíduos se um número $\delta_C \in [0, 1]$ gerado aleatoriamente no momento da operação for maior que a probabilidade de ocorrer *cross-over* p_C .

É importante ressaltar que, geralmente, define-se que os parâmetros trocados são escolhidos em sequência, começando do parâmetro α_1 , sendo sempre os t primeiros parâmetros, com $0 \leq t \leq n - 1$, onde n é o número de parâmetros.

Denotando por f_C^t a operação de *cross-over* entre dois indivíduos para uma troca de t parâmetros, temos

$$f_C^t : \Lambda \times \Lambda \rightarrow \Lambda \times \Lambda$$

Dados $I_1 = (\alpha_1^1, \dots, \alpha_n^1)$ e $I_2 = (\alpha_1^2, \dots, \alpha_n^2)$, temos

$$f_C^t(I_1, I_2) = \begin{cases} (I_1^{C,t}, I_2^{C,t}) & \text{se } \delta_C \leq p_C \\ (I_1, I_2) & \text{se } \delta_C > p_C \end{cases}$$

onde

$$I_1^{C,t} = (\alpha_1^2, \dots, \alpha_t^2, \alpha_{t+1}^1, \dots, \alpha_n^1)$$

$$I_2^{C,t} = (\alpha_1^1, \dots, \alpha_t^1, \alpha_{t+1}^2, \dots, \alpha_n^2)$$

Exemplo 3.8.4. Considere dois indivíduos do sistema S do exemplo 3.8.1:

$$I_1 = (60.3, 49, 80, 100.9, 110, 5, 7, 4, 8, 6)$$

$$I_2 = (40, 60, 89, 80, 230, 4, 8, 7, 5, 16)$$

Foi gerado aleatoriamente no conjunto $\{1, 2, \dots, 10\}$ o número $t = 3$, sendo que $p_C = 0.4$. Também gerado aleatoriamente, obteve-se $\delta_C = 0.2$. Portanto ocorrerá o *cross-over* entre I_1 e I_2 , resultando em

$$I_1^{C,3} = (40, 60, 89, 100.9, 110, 5, 7, 4, 8, 6)$$

$$I_2^{C,3} = (60.3, 49, 80, 80, 230, 4, 8, 7, 5, 16)$$

3.8.7 Eliminação

Em uma dada população, podem existir indivíduos cujo desempenho é extremamente baixo. Em tal situação, não é desejável que esses indivíduos continuem na população, pois além de estarem ocupando o lugar de outros que poderiam ter performance melhor, eles irão interagir com os outros indivíduos de alto desempenho em um *cross-over*, podendo gerar dois indivíduos de baixo desempenho.

Para tentar minimizar tal fato, define-se uma porcentagem de eliminação p_E , que define a porcentagem dos indivíduos da população que serão excluídos antes de aplicadas as operações. No lugar deles entrarão repetições de indivíduos presentes da porcentagem não excluída, de forma aleatória, porém dando prioridade aos com melhor desempenho.

Definindo por f_{E,p_E} a operação eliminação de porcentagem p_E e Ψ o conjunto de possíveis populações P de tamanho m quando definidos os indivíduos do sistema S , temos

$$f_{E,p_E} : \Psi \rightarrow \Psi$$

de modo que, dada $P = (I_1, \dots, I_m)$, então

$$f_{E,p_E}(P) = (I_1, I_2, \dots, I_e, I_{R_1}, I_{R_2}, \dots, I_{R_h})$$

sendo $e = \lfloor m(1 - p_E) \rfloor$, $h = m - e$ e

$$I_{R_k} \in \{I_1, I_2, \dots, I_e\}, \quad 1 \leq k \leq h$$

3.8.8 Função *fitness*

A função *fitness* é aquela que definirá o desempenho de um indivíduo. Será utilizada como uma figura de mérito, para comparar dois indivíduos quanto ao seu desempenho no sistema. Sendo assim, uma função *fitness* *fit* é basicamente definida por

$$fit : \Lambda \rightarrow \mathbb{R} \cup \{\infty\}$$

onde $fit(I_1) > fit(I_2)$ significa que o indivíduo I_1 tem melhor desempenho no sistema que I_2 . Caso $fit(I) = \infty$, significa que os parâmetros escolhidos para o indivíduo têm desempenho máximo no sistema.

Exemplo 3.8.5. Considere um sistema R no qual desejamos aproximar uma dada curva s em um intervalo $[a, b]$ por um polinômio de grau 7. Os 8 coeficientes a_0, a_1, \dots, a_7 do polinômio representariam os parâmetros do sistema, e a função *fit* poderia ser dada pelo inverso do erro quadrático total:

$$p(x) = a_7x^7 + a_6x^6 + \dots + a_1x + a_0$$

$$fit(I_1 = (a_0, a_1, \dots, a_7)) = \begin{cases} \frac{1}{\int_a^b [p(x) - s(x)]^2 dx} & \text{se } \int_a^b [p(x) - s(x)]^2 \neq 0 \\ +\infty & \text{caso contrário} \end{cases}$$

3.8.9 Elitismo

Elitismo simplesmente se baseia em manter o indivíduo com maior desempenho de uma geração k para a geração $k + 1$. Isso permite que tal indivíduo interaja com os novos gerados para a próxima geração, aumentando a chance de aprimorá-lo. Porém,

dependendo da aplicação, isso pode gerar populações muito homogêneas, fazendo com que a procura de uma solução ótima fique restrita a um mínimo local.

3.8.10 Critério de parada

Idealmente, o Algoritmo Genético pode ter qualquer número de gerações. Porém, um número elevado delas acarretaria em um grande tempo computacional. Por isso, é necessário estabelecer um critério de parada para o Algoritmo Genético.

Um exemplo de critério de parada é finalizar o algoritmo ao atingir um número fixo de gerações. Outra possibilidade é parar o algoritmo quando surgir um indivíduo de desempenho maior que um limiar desejável.

3.9 Sistema Fuzzy utilizado e desempenho

Uma vez obtidos os coeficientes Wavelets dos espectros de treino, esses são fornecidos como entrada ao método de Clusterização visto na seção 3.7, sendo as saídas o valor real da propriedade escolhida. Nesse método serão determinadas regras Fuzzy do modelo Takagi-Sugeno com base nos dados de entrada e saída, cujas funções de saída serão otimizadas pelo método dos mínimos quadrados. Haverá um sistema Fuzzy por indivíduo da população do Algoritmo Genético, pois é em seus indivíduos que residem os parâmetros da clusterização.

Um exemplo de sistema Fuzzy obtido quando se utilizam duas entradas apenas pode ser visto na figura 41, que mostra a fuzzyficação para uma entrada específica, sendo cada coluna respectiva às funções Fuzzy da respectiva entrada e a última coluna correspondente às funções Fuzzy de saída. Note que as funções de pertinência de entrada são gaussianas, já que o sistema Fuzzy foi gerado pelo método de clusterização aqui apresentado. A fuzzyficação pode ser notada pela limitação da gaussiana quando esta interage com o valor de entrada, gerando uma nova função de pertinência. Cada linha denota uma regra, composta por duas funções Fuzzy de entrada (uma para cada entrada) e uma de saída, como pode ser visto na figura 42. Por fim, como mostrado na figura 43, é realizada a defuzzyficação, considerando todas as funções de pertinência Fuzzy geradas pela interação dos valores de entrada com cada regra Fuzzy.

Após obter o sistema Fuzzy, este é submetido ao método ANFIS (JANG, 1993), que utiliza uma Rede Neural adaptativa para aprimorar os parâmetros das regras um modelo Takagi-Sugeno. No ANFIS existe um conceito chamado **número de épocas**, que representa quantas vezes os dados fornecidos foram re-treinados pelo ANFIS. Sempre que obtemos um novo sistema na nova época, verificamos se houve melhoras no sistema com a aplicação do método. Caso não haja melhoras, paramos de utilizar o ANFIS naquele sistema e é mantido aquele com melhor desempenho.

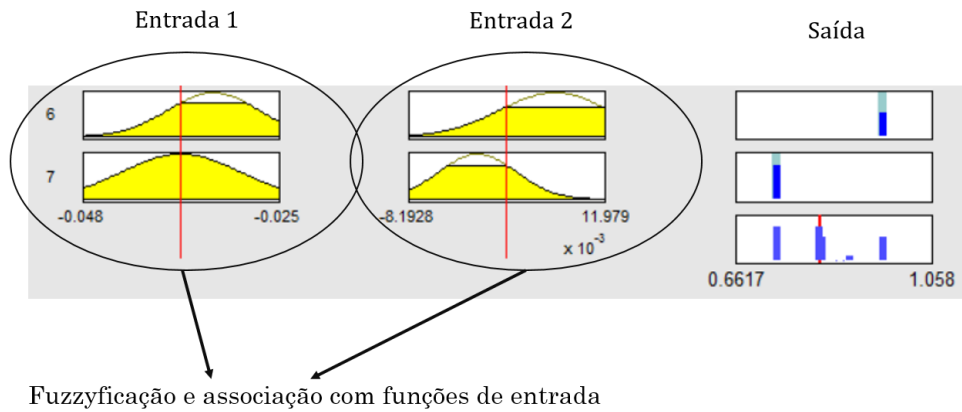


Figura 41 – Funções de pertinência relativas a cada entrada (para duas regras apenas)

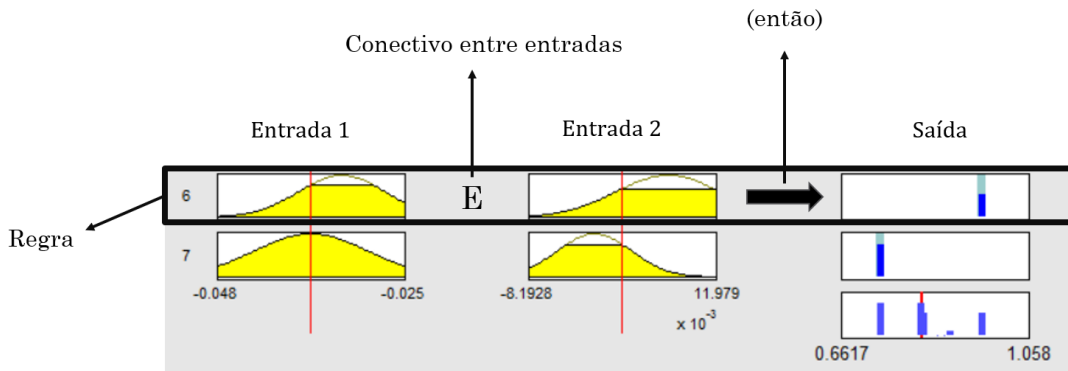


Figura 42 – Regra Fuzzy para um sistema com duas entradas

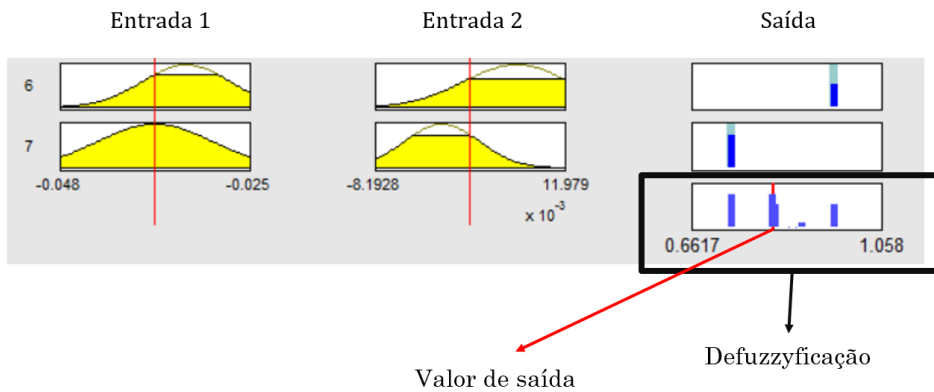


Figura 43 – Defuzzyficação das funções de pertinência de saída resultantes

Nessa etapa, os coeficientes Wavelets de teste são fornecidos ao sistema Fuzzy, gerando as saídas S_F , relativas à cada amostra. Tais saídas são comparadas com o valor real S das propriedades (obtido experimentalmente) por meio do coeficiente de correlação de Pearson ρ :

$$\rho(S, S_F) = \frac{\text{cov}(S, S_F)}{\sigma_S \sigma_{S_F}}$$

onde $\text{cov}(S, S_F)$ representa a covariância entre S e S_F . As variáveis σ_S e σ_{S_F} são o desvio padrão de S e S_F , respectivamente.

A covariância entre dois vetores X e Y de n elementos cada é obtida por

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

onde

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i \\ \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i. \end{aligned} \quad (3.11)$$

O desvio padrão σ_X de um vetor X de n elementos é obtido por:

$$\sigma_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

sendo \bar{x} dado por (3.11).

O coeficiente de correlação de Pearson é um número entre -1 e 1 . Quanto mais próximo de 1 , mais correlacionados estão os dois vetores. Isso significa que a saída fornecida pelo sistema está próxima da real, medida experimentalmente. De fato, suponha que nos vetores X e Y de n elementos tenhamos $x_i = y_i$ para todo $i = 1, 2, \dots, n$. Então

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \sigma_X^2$$

Logo

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sigma_X^2}{\sigma_X^2} = 1$$

Para determinar o desempenho do sistema, a função *fitness* utilizada foi o inverso do coeficiente de Pearson entre S e S_F :

$$fit = \frac{1}{\rho(S, S_F)} \quad (3.12)$$

A função (3.12) apenas fornece a informação do desempenho do sistema para as amostras de teste. Porém, em algumas situações, quando aplicada às amostras de treino, o desempenho do sistema foi ruim, mesmo sendo bom para as amostras de teste. Sendo assim, para essas ocasiões, a equação (3.12) precisou ser adaptada. Considerando

$$F = \frac{\rho(S^{tr}, S_F^{tr})}{\rho(S, S_F)} - 5[1 - \rho(S^{tr}, S_F^{tr})]$$

a nova fit é dada por

$$fit = \begin{cases} F & \text{se } F < 50 \\ F + \frac{1}{\rho(S^{tr}, S_F^{tr})} & \text{caso contrário} \end{cases}$$

onde S^{tr} e S_F^{tr} são, respectivamente, os valores reais da propriedade referente às amostras de treino e os valores obtidos pelo sistema Fuzzy a partir das amostras de treino.

3.10 Resultados

A seguintes tabelas fornecem os parâmetros e características do melhor sistema obtido para a propriedade densidade:

Densidade	
Característica	Valor/Tipo
Número de Amostras	160
Wavelet	Daubechies 3
Escala Wavelet	5
Épocas ANFIS	1

Tabela 9 – Características utilizadas para Geração do Sistema Fuzzy para a Propriedade Densidade

Densidade	
Parâmetro A.G.	Valor
Início da Região 1	979
Final da Região 1	1319
Início da Região 2	2514
Final da Região 2	3033
Raio de Influência	0.613
Fator de Esmagamento	0.86
Taxa de Aceitação	0.018
Taxa de Rejeição	0.01

Tabela 10 – Parâmetros do Melhor Indivíduo obtido pelo Algoritmo Genético (A.G.) para a Propriedade Densidade

Sendo 70% das amostras para treino e 30% para teste, o coeficiente de correlação de Pearson ρ entre a saída gerada por esse sistema e o valor real da densidade foi:

- Amostras de Teste: $\rho = 0.9823$
- Amostras de Treino: $\rho = 0.98804$
- Todas as Amostras: $\rho = 0.98561$

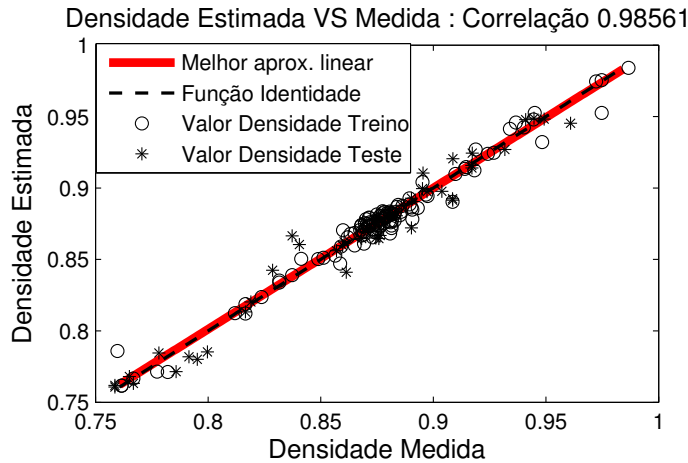


Figura 44 – Densidade Medida Vs. Densidade Estimada pelo Sistema Fuzzy para a Densidade

Na Figura 44, quanto mais próximo da função identidade os valores estiverem, melhor a predição feita pelo sistema Fuzzy, pois significa que a densidade estimada é próxima à medida. A figura 45 mostra o erro entre o valor estimado e o valor medido com relação ao número da amostra.

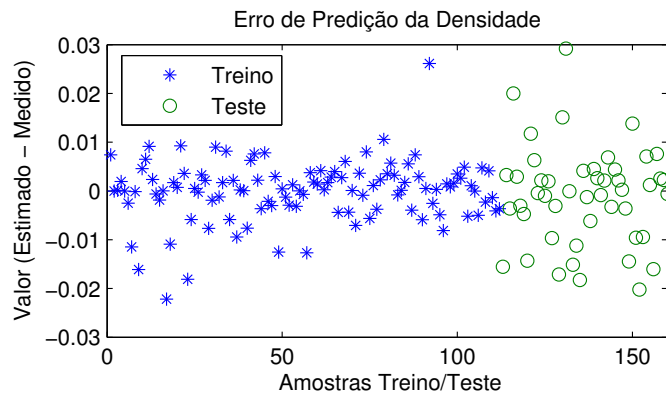


Figura 45 – Erro entre o Valor Estimado e o Medido para a Densidade

Densidade			
Tipo de Erro	Treino	Teste	Total
Máximo Absoluto	0.0261	0.0292	0.0292
Médio Absoluto	0.0043	0.0074	0.0052
RMS	0.0062	0.0100	0.0075

Tabela 11 – Valores de Erro entre Valor Estimado e Medido para a Densidade

O valor RMS (*Root Mean Square*) representa outra figura de mérito. Para um conjunto de valores x_1, x_2, \dots, x_n , o RMS é dado por:

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$$

O valor RMS do erro entre valores estimados e medidos é importante para comparar resultados, como será visto mais a frente nesta seção.

Já para a propriedade TIAC, as tabelas 12 e 13 fornecem os parâmetros e características do melhor sistema obtido.

TIAC	
Característica	Valor/Tipo
Número de Amostras	153
Wavelet	Daubechies 4
Escala Wavelet	5
Épocas ANFIS	1

Tabela 12 – Características utilizadas para geração do sistema Fuzzy para a propriedade TIAC

TIAC	
Parâmetro A.G.	Valor
Início da Região 1	966
Final da Região 1	1228
Início da Região 2	2606
Final da Região 2	3316
Raio de Influência	0.406
Fator de Esmagamento	0.854
Taxa de Aceitação	1.282
Taxa de Rejeição	0.088

Tabela 13 – Parâmetros do melhor indivíduo obtido pelo Algoritmo Genético (A.G.) para a propriedade TIAC

Considerando também 70% das amostras para treino e 30% para teste, o coeficiente de correlação de Pearson ρ entre a saída gerada por esse sistema e o valor real da TIAC foi:

- Amostras de Teste: $\rho = 0.93731$
- Amostras de Treino: $\rho = 0.95092$
- Todas as Amostras: $\rho = 0.94499$

A figura 47 mostra o erro entre o valor estimado e o valor medido com relação ao número da amostra.

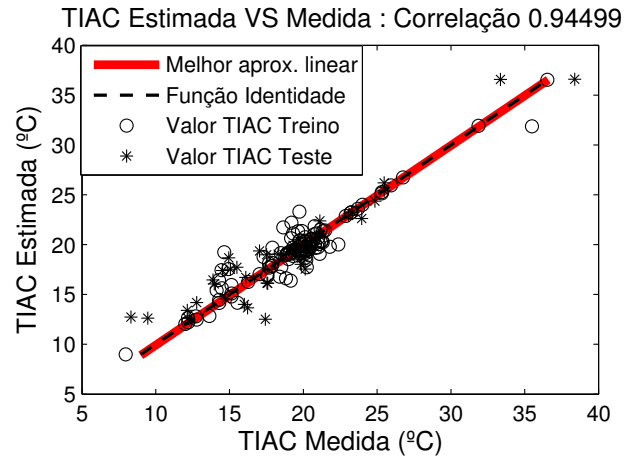


Figura 46 – TIAC Medida Vs. Densidade Estimada pelo Sistema Fuzzy para a TIAC

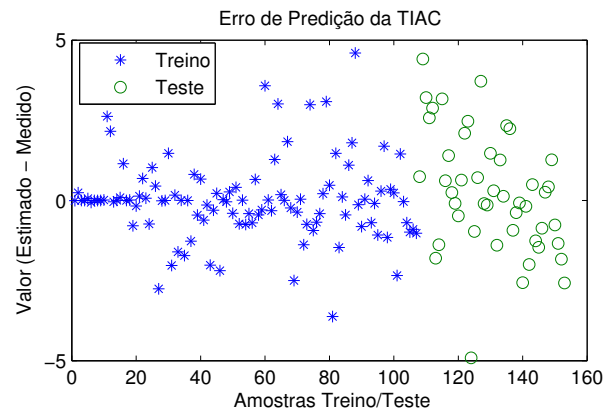


Figura 47 – Erro entre o Valor Estimado e o Medido para a TIAC

TIAC			
Tipo de Erro	Treino	Teste	Total
Máximo Absoluto	4.6006	4.9117	4.9177
Médio Absoluto	0.8189	1.4473	1.0078
RMS	1.2528	1.8663	1.4645

Tabela 14 – Valores de Erro entre Valor Estimado e Medido para a TIAC

Os valores de correlação referentes a cada propriedade mostram que o sistema Fuzzy consegue prevê-las com eficácia. Tal fato é comprovado também pelo erro médio absoluto (Tabelas 14 e 11) entre os valores previstos e os reais. Para exemplificar a qualidade dos valores obtidos, a tabela 15 faz uma comparação com o resultado de 2 artigos da revista *Fuel*: Artigo 1 - (DUARTE et al., 2016) e Artigo 2 - (FILGUEIRAS et al., 2014). Nesses artigos considera-se o grau API, que é uma reparametrização da densidade d :

$$^{\circ}\text{API} = \frac{141.5}{d} - 131.5$$

Já para os valores TIAC, a tabela 16 faz uma comparação do método proposto com o Artigo 1 (DUARTE et al., 2016).

Comparação Valores API			
Figura de Mérito	Método Proposto	Artigo 1	Artigo 2
ρ^2	0.9665	0.945	0.9751
RMS - Teste	1.9364	0.8	0.25

Tabela 15 – Comparação de resultados API

Comparação Valores TIAC		
Figura de Mérito	Método Proposto	Artigo 1
ρ^2	0.8786	0.857
RMS - Teste	1.8663	3.8

Tabela 16 – Comparação de resultados TIAC

Notamos também que a propriedade Densidade possui uma primeira região com mais pontos que a da TIAC, mas com relação à segunda região, a TIAC possui mais pontos. Isso pode significar que a informação no espectro necessária para obtenção da Densidade está mais concentrada na região 1 e mais distribuída na região 2, o que é o oposto do que acontece com TIAC.

Além disso, o sistema Fuzzy da propriedade TIAC demonstrou um menor desempenho do que o da Densidade. Isso ocorre porque a TIAC é mais difícil de se obter do espectro do petróleo. Porém, ambos os resultados são considerados bons em relação à propriedade.

Apesar do número de épocas do método ANFIS ter sido apenas 1 (o que significa que não houve uma grande melhora nos sistemas ao utilizá-lo), em grande parte dos sistemas tal método aprimora o resultado consideravelmente, o que reforça a necessidade de utilizá-lo junto com os demais descritos aqui.

4 Conclusão

O uso das Wavelets nas mais diversas aplicações tem demonstrado a importância da teoria desenvolvida para tais funções. Sua propriedade de escala permite explorar aspectos locais em funções e a variedade de funções existentes permite uma liberdade de escolha para determinar qual das Wavelets melhor se aplica ao problema proposto.

O método de criação de Wavelets demonstrado no segundo capítulo desta dissertação expõe as principais propriedades existentes na família de soluções gerada. Entre elas, uma das mais importantes é o fato das Wavelets serem uma base para $L^2(\mathbb{R})$, o que significa que podemos representar quase todas as funções presentes em problemas práticos por um somatório de Wavelets com seus respectivos coeficientes.

Com relação ao problema prático abordado nesta dissertação, a predição das propriedades Densidade e TIAC através do espectro de amostras de petróleo mostrou-se complexa e não intuitiva, o que levou à necessidade de utilizar os métodos aqui apresentados. Basicamente, o método completo possui os seguintes passos (em ordem lógica):

- Determinar as regiões do espectro a serem utilizadas no método;
- Determinar uma escala Wavelet para as regiões;
- Escolher uma Wavelet;
- Obter coeficientes a partir do produto interno de cada subregião (gerada pela escala escolhida) com a Wavelet selecionada;
- Separar amostras para treino e teste;
- Utilizar o método de clusterização, considerando os coeficientes obtidos referentes às amostras de treino como entrada e como saída os valores da propriedade relativo a cada amostra, e então obter um sistema Fuzzy;
- Aprimorar o desempenho do sistema por meio do método ANFIS;
- Utilizar um Algoritmo Genético para testar várias combinações de regiões e parâmetros de clusterização, com o intuito de obter o melhor sistema possível.

Com relação às regiões, notamos que a variação por alguns pontos apenas já pode comprometer o desempenho do sistema. Isso acontece pois é a combinação dos picos e vales da Wavelet com os do espectro que vão definir o valor dos coeficientes. Assim, tais variações podem fazer com que o produto interno filtre um pequeno intervalo importante

para a determinação do valor da propriedade. Por essa razão, a escolha de uma boa escala é fundamental. Uma escala muito grande significa muitas variáveis no processo, o que aumenta bastante o tempo de simulação. Além disso, não é garantido que o aumento da escala também aumente o desempenho.

Para a escolha da Wavelet, é importante citar que o sistema funciona basicamente para um tipo específico de Wavelet. Caso mude-se a Wavelet, é necessário utilizar o Algoritmo Genético novamente, para encontrar os parâmetros específicos para a dada Wavelet. Porém, é possível encontrar bons resultados com tipos variados de Wavelet, e o melhor resultado dependerá de sorte probabilística e tempo de simulação, já que não é possível afirmar qual é a melhor Wavelet para cada propriedade.

O método de clusterização também é bastante sensível aos valores dos parâmetros fornecidos. Quando a combinação de um Raio de Influência com um Fator de Esmagamento geram muitos núcleos, automaticamente temos muitas regras, o que aumenta o tempo de simulação. Além disso, existem combinações de valores que não conseguem fornecer um sistema Fuzzy. Portanto, é necessário estabelecer o conjunto adequado no qual tais parâmetros podem variar.

Devido ao elevado número de combinações existentes para os parâmetros do método como um todo, o uso do algoritmo genético foi de extrema importância para a determinação dos sistemas Fuzzy de alto desempenho. Seria inviável testar todas as combinações de parâmetro, até mesmo computacionalmente.

É interessante reforçar que o sistema que prediz uma propriedade é um único sistema Fuzzy, com suas regras e conjuntos otimizados para o problema. Sendo assim, todos os outros métodos descritos são apenas para otimizar tais regras e conjuntos.

Por fim, a teoria Wavelet aqui desenvolvida, juntamente com os métodos computacionais descritos, conseguiram ser utilizados em conjunto para o desenvolvimento de um sistema que prevê propriedades de amostras de petróleo apenas pelo espectro por emissão infravermelho. A comparação com outros métodos da literatura feita na seção dos resultados comprova que o método consegue prever as propriedades com desempenho comparável aos da literatura (e até mesmo melhor em alguns casos). Ainda assim, existe a possibilidade de se obter resultados ainda melhores, utilizando outros tipos de Wavelet ou simulando por longos períodos até conseguir um indivíduo melhor no algoritmo genético. Um artigo foi elaborado com a descrição do método e seus resultados e será submetido à revista *Chemometrics and Intelligent Laboratory Systems*.

Referências

- CHAU, F.-T. et al. *Chemometrics: From Basics to Wavelet Transform*. [S.l.]: John Wiley & Sons, Inc., Hoboken, New Jersey, 2004. ISBN 0-471-20242-8. Citado na página 15.
- CHIU, S. L. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, v. 2, p. 267–278, 1994. Citado 3 vezes nas páginas 96, 100 e 101.
- DAUBECHIES, I. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, XLI, p. 909–996, 1988. Citado 4 vezes nas páginas 45, 47, 64 e 66.
- DAUBECHIES, I. *Ten Lectures on Wavelets*. [S.l.]: SIAM: Society for Industrial and Applied Mathematics, 1992. (CBMS-NSF Regional Conference Series in Applied Mathematics). ISBN 9780898712742. Citado na página 15.
- DRIANKOV, D.; HELLENDORRN, H.; REINFRANK, M. *An Introduction to Fuzzy Control*. [S.l.]: Springer-Verlag Berlin Heidelberg New York, 1996. ISBN 978-3-642-08234-4. Citado 3 vezes nas páginas 74, 75 e 95.
- DUARTE, L. M. et al. Determination of some physicochemical properties in brazilian crude oil by h nmr spectroscopy associated to chemometric approach. *Fuel*, n. 181, p. 660—669, 2016. Citado na página 114.
- FILGUEIRAS, P. R. et al. Determination of api gravity, kinematic viscosity and water content in petroleum by atr-ftir spectroscopy and multivariate calibration. *Fuel*, n. 116, p. 123—130, 2014. Citado na página 114.
- HAAR, A. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, v. 69, p. 331–371, 1910. Citado na página 15.
- JANG, J. R. Anfis: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, v. 23, n. 3, 1993. Citado na página 108.
- MORLET, J. et al. Wave propagation and sampling theory-part i: Complex signal and scattering in multilayered media. *Society of Exploration Geophysicists*, v. 47, n. 2, p. 203–221, 1982. Citado na página 15.
- PASSINO, K. M.; YURKOVICH, S. *Fuzzy Control*. [S.l.]: Addison Wesley Longman, Inc., 1998. ISBN 0–201–18074–X. Citado na página 74.
- REED, M.; SIMON, B. *Methods of Modern Mathematical Physics: Vol. 1 - Functional Analysis*. [S.l.]: Academic Press, Inc, 1980. ISBN 0-12-585050-6. Citado 6 vezes nas páginas 16, 24, 26, 27, 41 e 42.
- ROYDEN, H. L.; FITZPATRICK, P. M. *Real Analysis*. [S.l.]: Pearson Education Asia Limited and China Machine Press, 2010. ISBN 978-0-13-143747-0. Citado 2 vezes nas páginas 59 e 60.
- SNEE, R. D. Validation of regression models: Methods and examples. *Technometrics*, v. 19(4), p. 415–428, 1977. Citado na página 69.

STRICHARTZ, R. S. How to make wavelets. *The American Mathematical Monthly*, v. 100, n. 6, p. 539–556, 1993. Citado 2 vezes nas páginas [21](#) e [41](#).