V Escola de Redes Neurais, Promoção: Conselho Nacional de Redes Neurais
pp. c001-c027, 19 de julho, 1999 - ITA, São José dos Campos - SP

Wavelets e Redes Neurais
1^a Parte: Noções sobre Wavelets

Prof^a Tânia Nunes Rabello

ITA, 19 de Julho de 1999

1. Histórico

A análise wavelet tem um desenvolvimento recente, embora desde 1952 as "wavelets "já figurassem implicitamente em trabalhos dos matemáticos A.Calderón e Zygmund.

No início dos anos 80, numerosos cientistas já utilizavam as "wavelets"como uma alternativa à análise de Fourier tradicional. As "wavelets" de J.S.Liénard e de X.Rodet relacionadas ao tratamento numérico de sinais acústicos (fala ou música) e as de J.Morlet, serviam para estocar e interpretar os sinais sísmicos recolhidos em campanhas de prospecção de petróleo. O físico A.Grossmann, juntamente com Morlet desenvolveram o formalismo da transformada wavelet contínua sem utilizar a teoria de representação de grupo e ainda, matemáticos tais como R.Coifman e G,Weiss criaram os "átomos"e as "moléculas" que deviam constituir as componentes de bases de diversos espaços funcionais. Enim, L.Carleson utilizava funções muito semelhantes às "wavelets" a fim de construir uma base incondicional do espaço H^1 de Stein e Weiss.

Nos últimos dez anos, aproximadamente, o interesse pelas wavelets cresceram assustadoramente. Por volta de 1985, foi dado um novo impulso a esta teoria através das contribuições de matemáticos e exprets em processamento de sinal. Um dos pioneiros foi Y.Meyer, um matemático da escola de Calderón-Zygmund, salientamos também as contribuições de I.Daubechies, S.Mallat, entre outros. E xistem várias razões para este sucesso. Por um lado, o conceito de wavelets pode ser visto como uma síntese de idéias originadas durante os últimos vinte ou trinta anos em engenharia(código subbanda), física(estados coerentes, grupo de renormalização) e matemática pura(estudo de operadores de Calderón-Zygmund). Como uma conseqüência dessa interdisciplinaridade, as wavelets apelam a cientistas de várias áreas. Por outro lado, as wavelets são ferramentas matemáticas bem simples com grande variedade de possíveis aplicações.

2. Notação

Daremos aqui algumas notações que serão usadas durante todo o texto.

Denotaremos por $\mathbb{Z}$, $\mathbb{R}$ e $\mathbb{R}^2$, o conjunto dos números inteiros, reais e dos pares ordenados de números reais, respectivamente.

Os espaços funcionais com os quais trabalharemos durante todo o texto são os espaços vetoriais $L^p(\mathbb{R})$, com $p = 1, 2$, cuja definição é a seguinte:

$L^p(\mathbb{R}) = \{u : \mathbb{R} \to \mathbb{R};\ u$ é mensurável e $|u|^p$ é integrável sobre $\mathbb{R}\}$

O espaço vetorial $L^1(\mathbb{R})$ é um espaço de Banach, munido da norma:

$$\|u\|_1 = \int_{\mathbb{R}} |u(x)|\, dx.$$

Enquanto que $L^2(\mathbb{R})$ é um espaço de Hilbert, munido do seguinte produto interno:

$$\langle u, v\rangle = \int_{\mathbb{R}} u(x)v(x)dx,$$

que define a seguinte norma:

$$\|u\| = \left[\int_{\mathbb{R}} |u(x)|^2\, dx\right]^{1/2}.$$

Portanto o espaço $L^2(\mathbb{R})$ é o espaço dos sinais de energia finita.

Para cada $f \in L^1(\mathbb{R})$, definimos a transformada de Fourier de f por:

$$(\mathcal{F}f)(\xi) = \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x)dx,$$

e mostra-se que $\left|\widehat{f}(\xi)\right| \leq (2\pi)^{-1/2} \|f\|_1$. Podemos, a partir de um processo de li mite, estender a definição da transformada de Fourier ao espaço $L^2(\mathbb{R})$, e aí obtemos que a transformada de Fourier $\mathcal{F}$ é um operador linear, isto é

$$\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

é tal que $\mathcal{F}(\alpha u + \beta v) = \alpha\mathcal{F}(u) + \beta\mathcal{F}(v)$. Em $L^2(\mathbb{R})$, obtemos ainda a identidade de Parseval, ou seja $\langle f, g\rangle = \left\langle \widehat{f}, \widehat{g} \right\rangle$.

Ainda, dadas $f, g \in L^2(\mathbb{R})$, definimos a convolução de f e g por:

$$(f * g)(x) = \int_{\mathbb{R}} f(y)g(x-y)dy$$

temos então a seguinte propriedade:

$$\mathcal{F}(f * g)(\xi) = \sqrt{2\pi}(\mathcal{F}f)(\xi)(\mathcal{F}g)(\xi)$$

3. Wavelet contínua

A transformada wavelet é uma ferramenta que decompõe dados, funções ou ope radores em diferentes componentes de freqüência e então estuda cada componente com uma resolução condizente com sua escala. Nos restringiremos às wavelets unidimensionais.

Definição: Seja $\psi \in L^2(\mathbb{R})$, satisfazendo a seguinte condição:

$$\int_{-\infty}^{+\infty} |\xi|^{-1} \left|\widehat{\psi}(\xi)\right|^2 d\xi < +\infty$$

.Definimos a transformada wavelet de uma função $f \in L^2(\mathbb{R})$, com respeito à função de análise ψ, por:

$$(T^{wav} f)(a, b) = \int_{-\infty}^{+\infty} f(x)\psi_{a,b}(x)dx,$$

onde, $\psi_{a,b}(x) = |a|^{-1/2}\, \psi\left(\frac{x-b}{a}\right)$.

Nota: As funções $\psi_{a,b}$ são chamadas wavelets, enquanto que a finção ψ é chamada wavelet mãe ou wavelet básica. Os parâmetros a e b são respectivamente os parâmetros de dilatação ou escala e de translação. Quando a varia, $\psi_{a,0}(s) = |a|^{-1/2}\, \psi(s/a)$ cobre diferentes domínios de freqüência(valores grandes do parâmetro de escala $|a|$ correspondem a baixas freqüências, ou largas escalas $\psi_{a,0}$; pequenos valores de $|a|$ correspondem a altas freqüências ou escalas finas $\psi_{a,0}$). A variação do parâmetro b também nos permite mover o centro de loca lização do tempo: cada $\psi_{a,b}(s)$ é localizada em torno de $s = b$. Portanto a transformada wavelet provê uma descrição tempo-freqüência de f. A normalização foi escolhida de modo que $\|\psi_{a,b}\| = \|\psi\|$ para todo a, b.

Nota: Observe que a transformada wavelet representa o produto escalar de f com cada função da família de wavelets $\{\psi_{a,b}\}$. Para a fixado, ela também pode ser interpretada como a convolução $(\psi_a^- * f)(b)$, onde $\psi_a^-(x) = \psi_{a,0}(-x)$.

Nota: Observe ainda que como $\psi \in L^2(\mathbb{R})$ então $\widehat{\psi} \in L^2(\mathbb{R})$, isto significa que estas funções tem decaimento para zero em $\pm\infty$, intuitivamente esta condição assegura que ψ e $\widehat{\psi}$ tem alguma concentração, ou seja são pequenas ondas ou me lhor ondas localizadas, daí o nome wavelet. A condição, $\int\limits_{-\infty}^{+\infty} |\xi|^{-1} \left|\widehat{\psi}(\xi)\right|^2 d\xi < +\infty$, denominada condição de admissibilidade é na prática(por exemplo se $\psi \in L^1(\mathbb{R})$) equivalente à ψ não ter componente de freqüência zero:$\widehat{\psi}(0) = 0$. Isto implica que ψ deve ter média zero, ou seja $\int\limits_{-\infty}^{+\infty} \psi(x)dx = 0$. Note que $|(T^{wav}f)(a,b)| \leq \|f\|, \forall f \in L^2(\mathbb{R})$.

3.1. A transformada wavelet e a análise tempo-freqüência

Em várias aplicações, dado um sinal $f(t)$, estamos interessados em seu conteúdo de freqüência localmente no tempo. A transformada de Fourier nos dá uma representação do conteúdo da freqüência de f, mas a informação relativa à localização no tempo de, por exemplo, altas freqüências não pode ser obtida facilmente de $\mathcal{F}f$. A localização tempo-freqüência pode ser obtida "janelando" primeiramente o sinal f e em seguida tomando sua transformada de Fourier, é o que chamamos de transformada de Fourier janelada de f, definida por:

$$(T^{win}f)(w,t) = \int_{-\infty}^{+\infty} f(s)g(s-t)e^{-iws}ds$$

Existem várias escolhas possíveis para a função janela g em análise de sinais, sendo as principais as de suporte compacto e regularidade razoável. Uma das escolhas mais populares é uma gaussiana. Em todas as aplicações, g é suposta bem concentrada em tempo e freqüência: se g e $\widehat{g}$ são ambas concentradas em torno do zero, então $(T^{win}f)(w,t)$ pode ser interpretada como o "conteúdo" de f em torno do tempo t e da freqüência w. A transformada de Fourier janelada provê assim uma descrição de f no plano tempo-freqüência. Mas, já que a freqüência de um sinal é diretamente proprcional ao comprimento do seu ciclo, segue que para informações espectrais de alta freqüência, o intervalo de tempo deve ser relativamente pequeno para dar uma melhor precisão, enquanto que para informações espectrais de baixa freqüência, o intervalo de tempo deve ser relativamente mais amplo para dar informações mais completas. Em outras palavras, é importante ter uma janela tempo-freqüência flexível, que se estreita automaticamente em centros de alta freqüência e se alarga em centros de baixa freqüência. Felizmente, a transformada wavelet T^{wav} relativa a uma wavelet mãe tem esta capacidade. Para tornar isto mais claro, daremos algumas definições.

Definição: Uma função não nula $w \in L^2(\mathbb{R})$ é denominada função janela se $xw(x)$ também pertence a $L^2(\mathbb{R})$. O centro t^*e o

raio Δ_w de uma função janela w são definidos por

$$t^* = \frac{1}{\|w\|^2} \int_{-\infty}^{+\infty} x\,|w(x)|^2\,dx$$

e

$$\Delta_w = \frac{1}{\|w\|} \left\{ \int_{-\infty}^{+\infty} (x - t^*)\,|w(x)|^2\,dx \right\}^{1/2}$$

respectivamente; e a largura da função janela w é definida por $2\Delta_w$.

Suponha que ψ é uma wavelet mãe tal que ψ e $\widehat{\psi}$ são funções janelas com centros e raios dados por $t^*, w^*, \Delta_\psi, \Delta_{\widehat{\psi}}$, respectivamente. Então, em primeiro lugar, é claro que a transformada wavelet de um sinal f, localiza o sinal com uma janela no tempo, dada por $[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi]$, sendo o centro da janela igual a $b + at^*$ e a largura dada por $2a\Delta_\psi$. Esta é a chamada "localização-tempo" em análise de sinais. Por outro lado, se considerarmos a fução janela η, dada por

$$\eta(w) = \widehat{\psi}(w + w^*)$$

então η também é uma função janela com centro em 0 e raio dado por $\Delta_{\widehat{\psi}}$ e pela identidade de Parseval a transformada wavelet de f torna-se

$$(T^{wav}f)(a,b) = a\,|a|^{-1/2} \int_{-\infty}^{+\infty} \widehat{f}(w) e^{ibw} \overline{\eta\left(a\left(w - \frac{w^*}{a}\right)\right)} dw.$$

Então com exceção do fator multiplicativo $a\,|a|^{-1/2}$ e da mudança de fase, linear e^{ibw}, determinada pela translação da janela de tempo, segue que a transformada wavelet também nos dá informação localizada do espectro de $\widehat{f}$, com janela de freqüência igual a $\left[\frac{w^*}{a} - \frac{1}{a}\Delta_{\widehat{\psi}}, \frac{w^*}{a} + \frac{1}{a}\Delta_{\widehat{\psi}}\right]$, cujo centro é $\frac{w^*}{a}$ e a largura é dada por $\frac{2}{a}\Delta_{\widehat{\psi}}$. Esta é a chamada "localização-freqüência". Assim do exposto acima, temos uma "janela tempo-freqüência":

$$[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi] \text{ x } \left[\frac{w^*}{a} - \frac{1}{a}\Delta_{\widehat{\psi}}, \frac{w^*}{a} + \frac{1}{a}\Delta_{\widehat{\psi}}\right]$$

para a análise tempo-freqüência, usando a transformada wavelet relativa a uma wavelet mãe ψ, sob as condições descritas acima.

Alguns comentários se fazem necessários. Em primeiro lugar, já que devemos considerar freqüências positivas, a wavelet mãe ψ deve ser escolhida de modo que o centro w^* de $\widehat{\psi}$ seja um número positivo. Na prática, este número positivo, assim como o parâmetro de escala a, é selecionado de modo que w^*/a seja o centro de freqüência da banda de freqüência $\left[\frac{w^*}{a} - \frac{1}{a}\Delta_{\widehat{\psi}}, \frac{w^*}{a} + \frac{1}{a}\Delta_{\widehat{\psi}}\right]$ de interesse. Então a razão entre o centro de freqüência e a largura da banda de freqüência é dada por

$$\frac{w^*/a}{2\Delta_{\widehat{\psi}}/a} = \frac{w^*}{2\Delta_{\widehat{\psi}}},$$

a qual independe da localização do centro de freqüência. Esta é a chamada "constante-Q" na análise de freqüência. A importância da janela tempo-freqüência é que ela se estreita para grandes centros de freqüência w^*/a e se alarga para pequenos centros de freqüência w^*/a, embora a área da janela seja constante, dada por $4\Delta_\psi\Delta_{\widehat{\psi}}$. E isto é o que se deseja em análise tempo-freqüência.

Assim, concluímos que a transformada wavelet é um instrumento melhor que a transformada de Fourier janelada para a análise de fenômenos rápidos de alta freqüência, tais como sinais transientes ou singularidades em funções ou núcleos de integrais.

3.2. Propriedades da wavelet contínua

Daremos a seguir algumas das principais propriedades da transformada wavelet. O resultado principal é a resolução da identidade, que nos mostra que uma função $f \in L^2(\mathbb{R})$ pode ser recoberta de sua transformada wavelet. Não daremos aqui as demonstrações, que podem ser encontradas em Daubechies[2].

Teorema. Seja $\psi \in L^2(\mathbb{R})$ uma wavelet básica. Então para toda $f, g \in L^2(\mathbb{R})$, tem-se que:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (T^{wav}f)(a,b)\overline{(T^{wav}g)(a,b)}\frac{da\,db}{a^2} = C_\psi \langle f, g\rangle ,$$

onde $C_\psi = 2\pi \int_{-\infty}^{+\infty} \left|\widehat{\psi}(\xi)\right|^2 |\xi|^{-1} d\xi$.

Proposition 1. *A resolução da identidade pode ser lida como*

$$f = C_\psi^{-1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (T^{wav}f)(a,b)\psi_{a,b}\frac{da\,db}{a^2},$$

com a convergência da integral no sentido fraco, isto é,

$$C_\psi^{-1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (T^{wav}f)(a,b)\overline{(T^{wav}g)(a,b)}\frac{da\,db}{a^2} = \langle f, g\rangle , \forall g \in L^2(\mathbb{R}).$$

No entanto, prova-se que tal convergência também vale no sentido forte de $L^2(\mathbb{R})$, isto é

$$\lim_{\substack{A_1 \to 0 \\ A_2, B \to +\infty}} \left\| f - C_\psi^{-1} \iint_{\substack{A_1 \leq |a| \leq A_2 \\ |b| \leq B}} (T^{wav}f)(a,b)\psi_{a,b}\frac{da\,db}{a^2} \right\| = 0.$$

Aqui a integral é vista como um elemento de $L^2(\mathbb{R})$, através do teorema de representação de Riesz.

A fórmula acima, mostra que qualquer $f \in L^2(\mathbb{R})$ pode ser arbitrariamente bem aproximada por uma superposição de wavelets. Esta afirmação parece paradoxal, já que todas as wavelets têm média zero, assim como qualquer superposição delas, enquanto que f não tem necessariamente média zero. A explicação deste aparente paradoxo está no fato de que a convergência é no sentido de $L^2(\mathbb{R})$ e não no sentido de $L^1(\mathbb{R})$. Quando $A_1 \to 0, A_2, B \to +\infty$, então

$$f(x) - C_\psi^{-1} \iint_{\substack{A_1 \leq |a| \leq A_2 \\ |b| \leq B}} (T^{wav}f)(a,b)\psi_{a,b}\frac{da\,db}{a^2},$$

é uma função que assume valores pequenos, e que tende lentamente para zero, quando x cresce, tem a mesma integral que f, mas sua norma em $L^2(\mathbb{R})$ tende a zero. Isto é análogo ao que acontece no seguinte exemplo: considere as funções

$$g_n(x) = \begin{cases} (2n)^{-1}, & |x| \leq n \\ 0, & \text{caso contrário} \end{cases}$$

então, $\int_{\mathbb{R}} g_n(x)dx = 1, \forall n$, ainda que $g_n \to 0$, para todo $x \in \mathbb{R}$ e $\|g_n\| = (2n)^{-1/2} \to 0$, quando $n \to +\infty$; mas (g_n) não converge em $L^1(\mathbb{R})$.

Existem variações da resolução da identidade, uma das quais consiste em introduzir diferentes funções para a reconstrução e para a decomposição. Mais explicitamente, se ψ^1, ψ^2 satisfazem

$$\int_{\mathbb{R}} |\xi|^{-1} \left|\widehat{\psi^1}(\xi)\right| \left|\widehat{\psi^2}(\xi)\right| d\xi < +\infty$$

então, prova-se que

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{a^2} \left\langle f, \psi^1_{a,b} \right\rangle \left\langle \psi^2_{a,b}, g \right\rangle da\, db = C_{\psi^1,\psi^2} \left\langle f, g \right\rangle$$

com $C_{\psi^1,\psi^2} = 2\pi \int_{\mathbb{R}} |\xi|^{-1} \overline{\widehat{\psi^1}(\xi)}\widehat{\psi^2}(\xi)d\xi$. Assim, se $C_{\psi^1,\psi^2} \neq 0$, podemos escrever a igualdade acima como

$$f = C^{-1}_{\psi^1,\psi^2} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{a^2} \left\langle f, \psi^1_{a,b} \right\rangle \psi^2_{a,b} da\, db.$$

Note que ψ^1, ψ^2 podem ter propriedades muito diferentes. Por exemplo uma pode não ser regular, enquanto que outra pode ser bem regular; ambas não necessitam nem mesmo serem admissíveis. Em Holschneider e Tchamitchian(1990), a liberdade na escolha de ψ^1, ψ^2 é usada para provar alguns resultados bastante interessantes. Um destes resultados é o estudo de propriedades de regularidade local de funções não diferenciáveis, a saber:

Teorema. Suponha que $\int\limits_{-\infty}^{+\infty} (1 + |x|) |\psi(x)| dx < +\infty$, e $\widehat{\psi}(0) = 0$. Se $f \in L^2(\mathbb{R})$ é uma função limitada e Hölder contínua com expoente α, $0 < \alpha \leq 1$, isto é,

$$|f(x) - f(y)| \leq C |x - y|^\alpha ,$$

então sua transformada wavelet satisfaz

$$|T^{wav}(a,b)| = \left|\left\langle f, \psi_{a,b} \right\rangle\right| \leq C_1 \left|a\right|^{\alpha+1/2}.$$

Reciprocamente, temos o seguinte resultado

Teorema. Suponha que ψ tem suporte compacto e que $f \in L^2(\mathbb{R})$ é limitada e contínua, Se, para algum $\alpha \in]0,1[$, a transformada wavelet de f satisfaz

$$\left|\left\langle f, \psi_{a,b} \right\rangle\right| \leq C \left|a\right|^{\alpha+1/2},$$

então f é Hölder contínua com expoente α.

Os teoremas acima mostram que a Hölder continuidade de uma função f pode ser caracterizada pelo decaimento em a do valor absoluto de sua transformada wavelet(exceto para $a = 1$, não temos a equivalência). Note que não assumimos nenhuma regularidade para ψ, a não ser ter média zero. A diferenciabilidade de ordem superior de f e a Hölder continuidade de sua derivada de maior ordem podem ser caracterizadas analogamente por meio do decaimento dos coeficientes wavelets, se ψ tem mais momentos nulos, isto é, supondo que $\int_{\mathbb{R}} x^m \psi(x)dx = 0$, para $m = 0, 1, \ldots, n$, temos o seguinte resultado:

$f \in C^n$, com $f^{(m)}$, $m = 0, 1, \ldots, n$ limitadas e quadrado integráveis, e $f^{(n)}$ Hölder contínua com expoente $\alpha \in]0,1[\Leftrightarrow \left|\left\langle f, \psi_{a,b} \right\rangle\right| \leq C \left|a\right|^{n+\alpha+1/2}$, uniformemente em a.

Novamente não exigimos nenhuma regularidade para ψ.

O mais importante em todas estas caracterizações é que elas envolvem somente o valor absoluto da transformada wavelet. Note que também podemos concluir sobre a regularidade de f a partir do decaimento em w do valor absoluto de sua transformada de Fourier janelada $T^{win}(w,t)$, se a janela g escolhida for suficientemente regular. No entanto o valor computado do expoente α, a partir de $|T^{win}(w,t)|$ não será ótimo. Para obter uma verdadeira caracterização, a fase de $|T^{win}(w,t)|$ deve ser levada em conta, por exemplo via estimativas tipo Littlewood-Paley.

A transformada wavelet pode ser usada também para caracterizar regularidade local, o que não acontece com a transformada de Fourier janelada, mesmo levando em conta a informação de fase. Os dois próximos teoremas são devidos à Holschneider e Tchamitchian(1990).

Teorema. Suponha que $\int\limits_{-\infty}^{+\infty} (1+|x|)\,|\psi(x)|\,dx < +\infty$ e $\int_{\mathbb{R}} \psi(x)dx = 0$. Se uma função f, limitada é Hölder contínua em x_0, com expoente $\alpha \in]0,1[$, isto é,

$$|f(x_0+h) - f(x_0)| \leq C\,|h|^{\alpha},$$

então

$$\left|\left\langle f, \psi_{a,x_0+b}\right\rangle\right| \leq C\,|a|^{1/2}\,(|a|^{\alpha} + |b|^{\alpha}).$$

Teorema. Suponha que ψ tem suporte compacto e $f \in L^2(\mathbb{R})$ é limitada e contínua. Se, para algum $\gamma > 0$ e $\alpha \in]0,1[$,

$$\left|\left\langle f, \psi_{a,b}\right\rangle\right| \leq C\,|a|^{\gamma+1/2} \text{ uniformemente em } b,$$

e

$$\left|\left\langle f, \psi_{a,x_0+b}\right\rangle\right| \leq C\,|a|^{1/2}\left(|a|^{\alpha} + \frac{|b|^{\alpha}}{|\log|b||}\right),$$

então f é Hölder contínua em x_0 com expoente α

Estes teoremas justificam o nome de "microscópio matemático", que é dado algumas vezes à transformada wavelet.

4. Transformada wavelet discreta

A transformada wavelet contínua é altamente redundante, e na prática é avaliada somente num grid no plano tempo-escala. Neste caso os parâmetros de escala a e translação b ambos assumem valores discretos. Para a, escolhemos as potências inteiras de um parâmetro de escala fixado $a_0 > 1$, isto é $a = a_0^m$. Como já visto, diferentes valores de m, correspondem a wavelets de diferentes larguras. Segue portanto que os valores do parâmetro de translação b dependerão de m : wavelets estreitas(altas freqüências) são transladadas por pequenos passos, para cobrir todo o domínio do tempo, enquanto que wavelets largas(baixas freqüências) são transladadas por passos largos. Já que a largura de $\psi(a_0^m x)$ é proporcional a a_0^m, nós escolhemos portanto para discretizar b por $b = nb_0a_0^m$, onde $b_0 > 0$ fixado, As wavelets correspondentes aos parâmetros discretos são, portanto

$$\psi_{m,n}(x) = a_0^{-m/2}\psi(a_0^{-m}(x - nb_0a_0^m)) = a_0^{-m/2}\psi(a^{-m}x - nb_0).$$

No caso discreto, não existe em geral uma fórmula análoga à resolução da identidade do caso contínuo. Surgem, então as seguintes questões:

1)É possível caracterizar f completamente, conhecendo-se $T^{wav}f$?

2)É possível reconstruir f de modo numericamente estável a partir de $T^{wav}f$?

Estas questões referem-se à reconstrução de f a partir da transformada wavelet discreta. Podemos também considerar o problema dual, a possibilidade de expandir f em termos de uma superposição de wavelets, surgem, então as seguintes questões:

A)Qualquer função pode ser escrita como uma superposição de $\psi_{m,n}$?

B)Existe um algoritmo numericamente estável para calcular os coeficientes para uma tal expansão?

Para responder a estas questões, teremos os dois próximos parágrafos.

4.1. Frames

Os frames foram introduzidos por Duffin e Schaeffer(1952), no contexto de séries de Fourier não harmônicas; eles são também revisadas em Young(1980). Vejamos sua definição. Neste parágrafo estaremos trabalhando num espaço de Hilbert H, munido de um produto interno $\langle .,. \rangle$.

Definição: Uma família de funções $(\varphi_j)_{j\in J}$ num espaço de Hilbert H é denominada "frame", se existem $A > 0, B < +\infty$ tais que, $\forall f \in H$, tem-se

$$A \|f\|^2 \leq \sum_{j\in J} \left|\langle f, \varphi_j \rangle\right|^2 \leq B \|f\|^2 .$$

Chamamos A e B de limitantes do frame Se os dois limitantes do frame são iguais, então a família $(\varphi_j)_{j\in J}$ será denominada de "tight frame". Portanto num "tight frame" temos

$$\sum_{j\in J} \left|\langle f, \varphi_j \rangle\right|^2 = A \|f\|^2 , \forall f \in H,$$

o que implica, pela identidade de polarização que

$$A \langle f, g \rangle = \sum_{j\in J} \langle f, \varphi_j \rangle \langle \varphi_j, g \rangle$$

isto é,

$$\langle f, g \rangle = A^{-1} \sum_{j\in J} \langle f, \varphi_j \rangle \langle \varphi_j, g \rangle , \forall g \in H,$$

o que implica,

$$f = A^{-1} \sum_{j\in J} \langle f, \varphi_j \rangle \varphi_j,$$

no sentido fraco.

A fórmula acima nos lembra a expansão de f em bases ortonormais, mas é importante notar que os "frames" e até mesmo os "tight frames" não são em geral bases ortonormais, como ilustraremos no exemplo a seguir.

Exemplo: Seja $H = \mathbb{C}^2$ e considere os vetores $e_1 = (0,1)$, $e_2 = (-\frac{\sqrt{3}}{2}, -\frac{1}{2})$, $e_3 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$. Assim, $\forall v = (v_1, v_2) \in H$, temos:
$\sum\limits_{j=1}^{3} |\langle v, e_j \rangle|^2 = |v_2|^2 + \left| -\frac{\sqrt{3}}{2} v_1 - \frac{1}{2} v_2 \right|^2 + \left| \frac{\sqrt{3}}{2} v_1 - \frac{1}{2} v_2 \right|^2 = \frac{3}{2}[|v_1|^2 + |v_2|^2] = \frac{3}{2} \|v\|^2$.

Logo $\{e_1, e_2, e_3\}$ é um "tight frame", mas definitivamente não é uma base ortonormal, já que não é nem mesmo linearmente independente. Observe também que neste exemplo o limitante do frame é $A = \frac{3}{2}$, que nos dá a razão de redundância do frame(três vetores num espaço bidimensional). Se a razão de redundância do tight frame for igual a 1, então o "tight frame" é uma base ortonormal, como veremos a seguir.

Teorema. Seja $(\varphi_j)_{j \in J}$ um tight frame num espaço de Hilbert H, com limitante $A = 1$ e $\|\varphi_j\| = 1, \forall j \in J$. Então $\{\varphi_j, j \in J\}$ constitui uma base ortonormal de H.

Prova:

$$1 = \|\varphi_k\|^2 = \sum_{j \in J} |\langle \varphi_k, \varphi_j \rangle|^2 = \|\varphi_k\|^2 + \sum_{\substack{j \in J \\ j \neq k}} |\langle \varphi_k, \varphi_j \rangle|^2 = 1 + \sum_{\substack{j \in J \\ j \neq k}} |\langle \varphi_k, \varphi_j \rangle|^2,$$

então $\sum\limits_{\substack{j \in J \\ j \neq k}} |\langle \varphi_k, \varphi_j \rangle|^2 = 0$, portanto $|\langle \varphi_k, \varphi_j \rangle|^2 = 0$, $\forall j \in J, j \neq k$. Ou seja, $\{\varphi_j, j \in J\}$ é um conjunto ortonormal de H. Resta provar que tal conjunto é completo. De fato, seja $f \in H$, tal que $\langle f, \varphi_j \rangle = 0$, $\forall j \in J$. Como

$$\|f\| = \sum_{j \in J} |\langle f, \varphi_j \rangle|^2 = 0 \Rightarrow f = 0.$$

Assim $\{\varphi_j, j \in J\}$ é uma base de H, logo

$$f = \sum_{j \in J} \langle f, \varphi_j \rangle \varphi_j, \forall f \in H.$$

A fórmula acima nos dá uma maneira trivial de recobrir f a partir de $\left\langle f, \varphi_j \right\rangle$, se o "frame" é um tight frame com limitante igual a 1. Vejamos o que acontece no caso geral.Primeiramente introduziremos o operador "frame".

Seja $(\varphi_j)_{j\in J}$ um "frame" com limitantes A e B num espaço de Hilbert H. Definamos então o operador "frame" $F : H \to l_2(J)$, por:

$$(Ff)_j = \left\langle f, \varphi_j \right\rangle, \forall f \in H, \forall j \in J.$$

É fácil mostrar que F é uma transformação linear e mais

$$\|Ff\|^2 = \sum_{j\in J} \left|\left\langle f, \varphi_j \right\rangle\right|^2 \leq B\|f\|^2, \forall f \in H,$$

o que implica que F é limitado e $\|F\| \leq B^{1/2}$.

A partir de F, definimos o operador adjunto F^*, que deve satisfazer a seguinte igualdade

$$\langle F^*c, f\rangle = \langle c, Ff\rangle, \forall c \in l_2(J), \forall f \in H,$$

o que nos permite definir $F^* : l_2(J) \to H$, por

$$F^*c = \sum_{j\in J} c_j \varphi_j, \forall c = (c_j)_{j\in J} \in l_2(J).$$

Como $\|F\| = \|F^*\|$, temos que F^* também é limitado e podemos provar que o operador F^*F satisfaz a seguinte desigualdade

$$\langle Af, f\rangle \leq \langle F^*Ff, f\rangle \leq \langle Bf, f\rangle, \forall f \in H,$$

o que implica que F^*F é inversível. Assim aplicando $(F^*F)^{-1}$ aos vetores φ_j obtemos uma nova família de vetores, a qual denotamos por $(\widetilde{\varphi}_j)_{j\in J}$ e definida por $\widetilde{\varphi}_j = (F^*F)^{-1}\varphi_j$. Prova-se que $(\widetilde{\varphi}_j)_{j\in J}$ é também um "frame", com limitantes A^{-1} e B^{-1}. A este "frame", denominaremos de "frame dual" de $(\varphi_j)_{j\in J}$. Ainda podemos provar que

$$\sum_{j\in J} \left\langle f, \varphi_j \right\rangle \widetilde{\varphi}_j = f = \sum_{j\in J} \left\langle f, \widetilde{\varphi}_j \right\rangle \varphi_j.$$

Isto significa que temos uma fórmula de reconstrução para f, a partir de $\left\langle f, \varphi_j \right\rangle$ e ao mesmo tempo uma receita para escrever f como uma superposição de φ_j, o que demonstra que os conjuntos são realmente duais. Assim, dado um "frame", o que é necessário fazer para aplicar a fórmula acima, é computar $\widetilde{\varphi}_j$. Voltaremos a isto em breve, mas primeiramente endereçamos uma questão que frequentemente aparece neste ponto. Enfatizamos que "frame" e mesmo "tight frame" não são bases ortonormais porque o conjunto $\{\varphi_j ; j \in J\}$ é tipicamente linearmente dependente. Isto significa que para uma dada $f \in H$, existem várias diferentes superposições de φ_j. O que então aponta a segunda parte da fórmula acima como especialmente interessante? Daremos uma idéia da resposta com um exemplo bem simples.

Exemplo: Consideramos, como em exemplo anterior $H = \mathbb{C}^2$ e o conjunto $\{e_1, e_2, e_3\}$, onde $e_1 = (0,1)$; $e_2 = (-\frac{\sqrt{3}}{2}, -\frac{1}{2})$ e $e_3 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$. Já provamos que o conjunto acima é um "tight frame" com limitante $A = \frac{3}{2}$. Pode-se provar ainda que

$$v = \frac{2}{3} \sum_{j=1}^{3} \langle v, e_j \rangle \, e_j, \forall v \in H.$$

É possível provar ainda que, se $v = \sum\limits_{j=1}^{3} \beta_j e_j$ então $\beta_j = \frac{2}{3} \left[\langle v, e_j \rangle + \alpha \right]$, para algum $\alpha \in \mathbb{C}$. De algum modo parece mais "econômico" a primeira fórmula que a segunda. Vejamos como podemos precisar esta afimação intuitiva.

$$\sum_{j=1}^{3} \left| \langle v, e_j \rangle \right|^2 = \frac{3}{2} \left\| v \right\|^2,$$

enquanto que

$$\sum_{j-1}^{3} \left| \langle v, e_j \rangle + \alpha \right|^2 = \frac{3}{2} \left\| v \right\|^2 + 3 \left| \alpha \right|^2.$$

Logo, se $\alpha \neq 0$, $\sum_{j-1}^{3} \left| \langle v, e_j \rangle + \alpha \right|^2 > \sum_{j=1}^{3} \left| \langle v, e_j \rangle \right|^2$, o que significa que a norma em $l_2(J)$ dos primeiros coeficientes é menor que a mesma norma dos demais coeficientes. É neste sentido que dizemos que os coeficientes $\left\langle f, \widetilde{\varphi}_j \right\rangle$ são mais econômicos, como enunciaremos no seguinte resultado.

Teorema. Se $f = \sum\limits_{j \in J} c_j \varphi_j$, para algum $c = (c_j) \in l_2(J)$ e se nem todo c_j é igual a $\left\langle f, \widetilde{\varphi}_j \right\rangle$, então $\sum\limits_{j \in J} \left| c_j \right|^2 > \sum\limits_{j \in J} \left| \left\langle f, \widetilde{\varphi}_j \right\rangle \right|^2$.

Analogamente, também não temos unicidade na reconstrução de f a partir de $\left\langle f, \varphi_j \right\rangle$, isto é existem outras famílias (u_j) de vetores de H tais que

$$f = \sum_{j \in J} \left\langle f, \varphi_j \right\rangle u_j$$

. No entanto os vetores $\widetilde{\varphi}_j$, mostram-se os mais econômicos, no seguinte sentido:

$$\begin{aligned}\text{Se } (u_j)_{j\in J} \text{ é tal que } f &= \sum_{j\in J} \langle f, \varphi_j\rangle u_j, \forall f \in H,\\ \text{então } \sum_{j\in J} |\langle u_j, g\rangle|^2 &\geq \sum_{j\in J} \left|\langle \widetilde{\varphi}_j, g\rangle\right|^2, \forall g \in H.\end{aligned}$$

Assim, para reconstruir f a partir de $\langle f, \varphi_j\rangle$ necessitamos computar $\widetilde{\varphi}_j$, o que envolve a inversa de F^*F e que não é simples. No entanto se A e B são "próximos" um do outro, isto é, $r = \dfrac{B}{A} - 1 << 1$, então F^*F está "próximo" de $\dfrac{A+B}{2}Id$ e portanto $(F^*F)^{-1}$ está próximo de $\dfrac{2}{A+B}Id$, logo $\widetilde{\varphi}_j$ está próximo de $\dfrac{2}{A+B}\varphi_j$, no seguinte sentido

$$\left\langle \frac{A-B}{B(A+B)}f, f\right\rangle \leq \left\langle (F^*F)^{-1}f - \frac{2}{A+B}f, f\right\rangle \leq \left\langle \frac{B-A}{A(A+B)}f, f\right\rangle, \forall f \in H,$$

o que nos permite concluir que $\widetilde{\varphi}_j \cong \dfrac{2}{A+B}\varphi_j$. Mais precisamente,

$$f = \frac{2}{A+B}\sum_{j\in J} \langle f, \varphi_j\rangle \varphi_j + Rf,$$

onde R é um operador auto-adjunto tal que

$$\|R\| = \sup\{|\langle Rf, f\rangle| ; f \in H e \|f\| = 1\} \leq \frac{B-A}{A+B} = \frac{r}{r+2}.$$

Assim se r é pequeno, obtemos uma fórmula de reconstrução para f, abandonando o resto Rf na fórmula acima e cujo erro na norma de $L^2(\mathbb{R})$ é de até $\dfrac{r}{r+2}\|f\|$.

Mesmo se r não for pequeno, podemos obter um algoritmo para a reconstrução de f, com convergência exponencial a partir de um método iterativo, que pode ser encontrado em [2].

4.2. Frames de wavelet

No parágrafo anterior, vimos que se tivermos um "frame" num espaço de Hilbert H, teremos um algoritmo de reconstrução numericamente estável para toda $f \in$

H. Neste parágrafo daremos condições necessárias e condições suficientes, para podermos determinar quando uma função ψ gera uma família de wavelets $\{\psi_{m,n}; m, n \in \mathbb{Z}\}$ que constitui um frame, e assim teremos um algoritmo numericamente estável de reconstrução em $L^2(\mathbb{R})$. Primeiramente, daremos uma condição necessária.

Teorema. Seja $\psi \in L^2(\mathbb{R})$ e $\psi_{m,n}(x) = a_0^{-m/2}\psi(a_0^{-m}x - nb_0), m, n \in \mathbb{Z}$. Se $\{\psi_{m,n}; m, n \in \mathbb{Z}\}$ constitui um "frame" para $L^2(\mathbb{R})$ com limitantes A e B, então

$$\frac{b_0 \ln a_0}{2\pi} A \leq \int_0^{+\infty} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi \leq \frac{b_0 \ln a_0}{2\pi} B$$

e

$$\frac{b_0 \ln a_0}{2\pi} A \leq \int_{-\infty}^{0} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi \leq \frac{b_0 \ln a_0}{2\pi} B.$$

Nota: As fórmulas acima impõem a condição de admissibilidade para que ψ dê origem a um frame, que é a mesma restrição que no caso contínuo.

Se $\{\psi_{m,n}; m, n \in \mathbb{Z}\}$ constitui um tight frame($A = B$), então as fórmulas acima implicam que

$$A = \frac{2\pi}{b_0 \ln a_0} \int_0^{+\infty} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi = \frac{2\pi}{b_0 \ln a_0} \int_{-\infty}^{0} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi.$$

Em particular, se $\{\psi_{m,n}; m, n \in \mathbb{Z}\}$ constitui uma base ortonormal($A = B = 1$) de $L^2(\mathbb{R})$, então

$$\int_0^{+\infty} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi = \int_{-\infty}^{0} \frac{\left|\widehat{\psi}(\xi)\right|^2}{\xi} d\xi = \frac{b_0 \ln a_0}{2\pi}.$$

Nem todas as escolhas de ψ, a_0, b_0 levam a "frame" de wavelets, mesmo se ψ é admissível. Por isso, daremos a seguir condições suficientes sobre ψ, a_0, b_0, sob as quais nós de fato obtemos um "frame". Obteremos também uma estimativa para os limitantes do "frame".

Proposition 1. *Se ψ, a_0 são tais que*

$$\inf_{1 \leq |\xi| \leq a_0} \sum_{m=-\infty}^{+\infty} \left|\widehat{\psi}(a_0^m \xi)\right|^2 > 0$$

e

$$\sup_{1\le|\xi|\le a_0}\sum_{m=-\infty}^{+\infty}\left|\widehat{\psi}(a_0^m\xi)\right|^2<+\infty$$

e se $\beta(s)=\sup_{\xi}\sum_{m=-\infty}^{+\infty}\left|\widehat{\psi}(a_0^m\xi)\right|\left|\widehat{\psi}(a_0^m\xi+s)\right|$ *decai pelo menos tão rápido quanto* $(1+|s|)^{-(1+\varepsilon)}$, *com* $\varepsilon>0$, *então existe* $(b_0)^*>0$ *tal que* $\{\psi_{m,n};m,n\in\mathbb{Z}\}$ *constitui um "frame", para todas as escolhas de* $b_0<(b_0)^*$ *e as expressões para os limitantes do "frame"* $\{\psi_{m,n};m,n\in\mathbb{Z}\}$ *são:*

$$A=\frac{2\pi}{b_0}\left\{\inf_{1\le|\xi|\le a_0}\sum_{m=-\infty}^{+\infty}\left|\widehat{\psi}(a_0^m\xi)\right|^2-\sum_{\substack{k=-\infty\\k\neq0}}^{+\infty}\left[\beta\left(\frac{2\pi}{b_0}k\right)\beta\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}$$

$$B=\frac{2\pi}{b_0}\left\{\sup_{1\le|\xi|\le a_0}\sum_{m=-\infty}^{+\infty}\left|\widehat{\psi}(a_0^m\xi)\right|^2+\sum_{\substack{k=-\infty\\k\neq0}}^{+\infty}\left[\beta\left(\frac{2\pi}{b_0}k\right)\beta\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}$$

As condições acima são satisfeitas se, por exemplo $\left|\widehat{\psi}(\xi)\right|\le C\,|\xi|^{\alpha}\,(1+|\xi|)^{-\gamma}$, *com* $\alpha>0$ *e* $\gamma>\alpha+1$.

Vimos então que é possível reconstruir f a partir de $\left(\langle f,\psi_{m,n}\rangle\right)_{m,n\in\mathbb{Z}}$, onde $\psi_{m,n}(x)=a_0^{-m/2}\psi(a_0^{-m}x-nb_0)$, de forma numericamente estável, desde que $\{\psi_{m,n};m,n\in\mathbb{Z}\}$ constitua um "frame".

No próximo parágrafo veremos como podemos construir bases ortonormais do espaço $L^2(\mathbb{R})$.

5. Bases ortonormais de wavelets e Análise de multiresolução

A primeira construção de bases ortonormais de wavelets pareceu um tanto miraculosa. Esta situação mudou com o advento da análise de multiresolução,-formulada por Mallat e Meyer, em 1986. A análise de multiresolução provê uma ferramenta para a compreensão das bases de wavelets e para a construção de novos exemplos. A história da formulação da análise de multiresolução é um belo exemplo de aplicação estimulante do desenvolvimento teórico. Quando Mallat aprendeu sobre a base de Meyer, ele foi trabalhando em análise de imagens, onde a idéia de estudar imagens simultaneamente em diferentes escalas é muito popular há vários anos. Isto o estimulou a ver as bases ortonormais de wavelets como um instrumento para descrever matematicamente o incremento na informação necessário para ir de uma aproximação grosseira para uma aproximação de alta resolução. Esta compreensão está cristalizada na análise de multiresolução.

5.1. Definção e propriedades

Os pesquisadores de várias especialidades queriam encontrar algoritmos manipuláveis, permitindo decompor funções arbitrárias em termos de funções especiais, combinando as vantagens do sistema trigonométrico e do sistema de Haar. Estes dois sistemas ocupam duas posições extremas, no seguinte sentido. As funções do sistema trigonométrico são perfeitamente localizadas em freqüência ou em variável de Fourier, mas não têm localização precisa na variável de espaço(ou tempo). Por outro lado as funções do sistema de Haar são perfeitamente localizadas na variável de espaço, mas têm uma má localização na variável de Fourier. Isto se deve a dois defeitos das funções do sistema de Haar: a falta de regularidade e a falta de oscilação.

A pesquisa de base hilbertianas que fossem ao mesmo tempo bem localizadas na variável de espaço e na variável de Fourier foi motivada por R. Balian, da seguinte maneira. "Podemos ter interesse em teoria de comunicações, em re presentar um sinal oscilante como uma superposição de "ondinhas" elementares tal que cada uma possua ao mesmo tempo uma freqüência e uma localização no tempo bastante bem definidas. A informação útil é, de fato, freqüentemente veiculada ao mesmo tempo pelas freqüências emitidas e pela estrutura temporal do sinal(o exemplo da música é característico). A representação de uma sinal como função do tempo exibe mal o espectro das freqüências em jogo, enquanto que ao contrário sua análise de Fourier mascara o instante de emissão e a duração de cada um dos elementos do sinal. Uma representação adequada deveria combinar as vantagens destas duas descrições complementares, apresentando além disso um caráter discreto melhor adaptado à teoria de comunicações".

A noção de análisa de multiresolução permite conciliar a análisa em variável de espaço com a análise em variável de Fourier e isto de maneira compatível com o princípio de incerteza de Heisenberg. Antes de passarmos à definição propriamente dita vejamos qual a idéia.

Seja $\psi \in L^2(\mathbb{R})$ uma wavelet tal que o conjunto $\{\psi_{j,k};\ j,k \in \mathbb{Z}\}$ seja uma base de $L^2(\mathbb{R})$, onde $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$. Assim, para cada $j \in \mathbb{Z}$, considere o subespaço $W_j = \overline{[\{\psi_{j,k}; k \in \mathbb{Z}\}]}$. É claro que $L^2(\mathbb{R})$ pode ser decomposto como uma soma direta dos subespaços W_j :

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

no sentido que toda função $f \in L^2(\mathbb{R})$ tem uma única decomposição:

$$f(x) = \cdots + g_1(x) + g_0(x) + g_{-1}(x) + \cdots,$$

onde $g_j \in W_j$ para todo $j \in \mathbb{Z}$.

Se $\{\psi_{j,k};\ j,k \in \mathbb{Z}\}$ é uma base ortonormal de $L^2(\mathbb{R})$, então os subespaços W_j são mutuamente ortogonais, isto é,

$$\langle g_j, g_l\rangle = 0, \quad j \neq l, \quad \text{onde } g_j \in W_j \text{ e } g_l \in W_l.$$

Neste caso usa-se a notação:

$$W_j \perp W_l, \quad j \neq l.$$

Consequentemente, a soma direta torna-se uma soma ortogonal, ou seja além da decomposição ser única, seus componentes são mutuamente ortogonais.

Sendo a base ortonormal ou não, tem-se uma decomposição do $L^2(\mathbb{R})$ em soma direta dos subespaços W_j. Para cada $j \in \mathbb{Z}$, considere os subespaços fechados de $L^2(\mathbb{R})$, definidos por:

$$V_j = W_{j+1} \oplus W_{j+2} \oplus \cdots \quad , j \in \mathbb{Z}.$$

Esses subespaços têm as seguintes propriedades:

1)$\cdots \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots$;

2)$\overline{\left(\bigcup_{j\in\mathbb{Z}} V_j\right)} = L^2(\mathbb{R})$;

3)$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$:

4)$V_{j-1} = V_j \oplus W_j$, $j \in \mathbb{Z}$;

5)$f(.) \in V_j \Leftrightarrow f(2.) \in V_{j-1}$.

Então, em contraste com os subespaços W_j, que satisfazem $W_j \cap W_l = \{0\}$, $j \neq l$, a seqüência dos subespaços V_j é uma seqüência encaixante, como descrito por (1), e tem a propriedade de que toda função $f \in L^2(\mathbb{R})$ pode ser aproximada tão bem quanto se queira por suas projeções $P_jf \in V_j$, como descrito em (2). Mas por outro lado, para j cada vez maior, as projeções P_jf devem ter energia arbitrariamente pequenas, como garantido por (3). A principal propriedade desses subespaços que não é descrita pelas propriedades de (1) a (3) é que mais e mais "variações" de P_jf são removidas quando $j \to +\infty$. De fato, essas variações são tiradas fora nível por nível em ordem crescente da "razão de variação" e estocadas nos subespaços complementares W_j como em (4). Este processo pode ser tornado muito eficiente por uma aplicação da propriedade(5).

De fato, se o subespaço de referência V_0 for gerado por uma única função $\phi \in L^2(\mathbb{R})$ no sentido que

$$V_0 = \overline{[\phi_{0,k}; k \in \mathbb{Z}]}$$

onde

$$\phi_{j,k} = 2^{-j/2}\phi(2^{-j}x - k),$$

então todos os subespaços V_j são também gerados pela mesma função ϕ (assim como os subespaços W_j são gerados por ψ) isto é:

$$V_j = \overline{[\phi_{j,k}; k \in \mathbb{Z}]}, j \in \mathbb{Z}.$$

Então, este processo pode ser realizado eficientemente, como veremos mais tarde. Vejamos então a definição de análise de multiresolução.

Definição: Uma análise de multiresolução de $L^2(\mathbb{R})$ é por definição, uma seqüência decrescente $V_j, j \in \mathbb{Z}$, de subespaços vetoriais fechados de $L^2(\mathbb{R})$, isto é $\ldots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots$, tendo as seguintes propriedades:

1. $\bigcap\limits_{j\in\mathbb{Z}} V_j = \{0\}$.
2. $\bigcup\limits_{j\in\mathbb{Z}} V_j$ é denso em $L^2(\mathbb{R})$.
3. $\forall f \in L^2(\mathbb{R}), \forall j \in \mathbb{Z}$, tem-se $f(.) \in V_j \Leftrightarrow f(2.) \in V_{j-1}$.
4. $f \in V_0 \Rightarrow f(. - n) \in V_0, \forall n \in \mathbb{Z}$.
5. $\exists \phi \in V_0$ tal que se $\phi_{0,n}(x) = \phi(x-n)$ então $\{\phi_{0,n}; n \in \mathbb{Z}\}$ é uma base de Riesz de V_0.

Neste caso ϕ é denominada função escala da análise de multiresolução.

Lembremos que uma base de Riesz de um espaço de Hilbert H é um subconjunto $\{e_1, e_2, \ldots, e_n, \ldots\}$ de H tal que $[e_1, e_2, \ldots, e_n, \ldots]$ é denso em H e existem constantes $c_2 > c_1 > 0$ tais que para toda seqüência de escalares $\alpha_0, \alpha_1, \alpha_2, \ldots$ tem-se:

$$c_1 \left(\sum_{k=0}^{+\infty} |\alpha_k|^2 \right)^{1/2} \leq \left\| \sum_{k=0}^{+\infty} \alpha_k e_k \right\| \leq c_2 \left(\sum_{k=0}^{+\infty} |\alpha_k|^2 \right)^{1/2}.$$

De (5) temos o seguinte resultado: Definindo $\phi_{j,n}(x) = 2^{-j/2}\phi(2^{-j}x - n)$, segue que $\{\phi_{j,n}; n \in \mathbb{Z}\}$ é uma base de Riesz de V_j, para cada $j \in \mathbb{Z}$.

O princípio básico da análise de multiresolução é que quando uma coleção decrescente de subespaços fechados de $L^2(\mathbb{R})$ satisfaz as condições de (1) a (5), e mais algumas outras hipóteses de regularidade sobre ϕ, então existe uma base de Riesz de wavelets $\{\psi_{j,k}; j, k \in \mathbb{Z}\}$ de $L^2(\mathbb{R})$, onde $\psi \in L^2(\mathbb{R})$ e $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$.Ainda a wavelet ψ pode ser construída explicitamente. A construção da wavelet no caso geral pode ser encontrada em [1]. Mostraremos aqui somente o caso em que $\{\phi_{0,n}; n \in \mathbb{Z}\}$ é uma base ortonormal de V_0.

Para cada $j \in \mathbb{Z}$ definamos W_j, como o complemento ortogonal de V_j em V_{j-1}, isto é $w \in W_j \Leftrightarrow w \in V_{j-1}$ e $\langle w, u\rangle = 0$, $\forall u \in V_j$. Temos então que

$$V_{j-1} = V_j \oplus W_j.$$

Ainda, se $j \neq k$ então $W_j \perp W_k$, ou seja $\langle w, u\rangle = 0, \forall w \in W_j$ e $\forall u \in W_k$. Portanto, para $j < J$, segue que

$$V_j = V_J \oplus \bigoplus_{k=0}^{J-j-1} W_{J-k},$$

onde todos estes espaços são ortogonais. Assim, em virtude das propriedades (1) e (2) da definição, pode-se mostrar que

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

Além disso, os subespaços W_j herdam a propriedade (3) de V_j, isto é,

$$f \in W_j \Leftrightarrow f(2^j.) \in W_0.$$

Logo, se $\{\psi_{0,k}; k \in \mathbb{Z}\}$ for uma base ortonormal de W_0, segue da equivalência acima que, para cada $j \in \mathbb{Z}$, $\{\psi_{j,k}; k \in \mathbb{Z}\}$ é uma base ortonormal de W_j e portanto $\{\psi_{j,k}; j, k \in \mathbb{Z}\}$ é uma base ortonormal de $L^2(\mathbb{R})$. Assim, nossa tarefa se reduz a encontrar $\psi \in W_0$ tal que $\{\psi(. - k); k \in \mathbb{Z}\}$ constitua uma base ortonormal de W_0.

Para construir esta ψ, a partir da função escala ϕ, utilizamos propriedades decorrentes da definição da análise de multiresolução e concluímos que

$$f \in W_0 \Leftrightarrow \widehat{f}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2 + \pi)}\nu(\xi)\widehat{\phi}(\xi/2),$$

onde ν é $2\pi - periódica$(ou constante) e $m_0(\xi) = \dfrac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} \left\langle \phi, \phi_{-1,n} \right\rangle e^{-in\xi}$. A fórmula geral acima para a transformada de Fourier de $f \in W_0$ sugere que tomemos

$$\widehat{\psi}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2 + \pi)}\widehat{\phi}(\xi/2),$$

como uma candidata para nossa wavelet. Assim, temos o seguinte resultado

Teorema. Se uma seqüência de subespaços fechados $(V_j)_{j\in\mathbb{Z}}$ em $L^2(\mathbb{R})$ satisfaz as condições de uma análise de multiresolução, sendo $\{\phi_{0,n}; n \in \mathbb{Z}\}$ uma base ortonormal de V_0, então existe uma base ortonormal associada de wavelets $\{\psi_{j,k}; j,k \in \mathbb{Z}\}$ para $L^2(\mathbb{R})$ tal que

$$P_{j-1}f = P_j f + \sum_{k\in\mathbb{Z}} \left\langle f, \psi_{j,k} \right\rangle \psi_{j,k}, \forall f \in L^2(\mathbb{R}).$$

Uma possibilidade para a construçãoda wavelet ψ é

$$\widehat{\psi}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2+\pi)}\widehat{\phi}(\xi/2),$$

(com m_0 definida como acima), ou equivalentemente

$$\psi = \sum_{n\in\mathbb{Z}} (-1)^{n-1} h_{-n-1}\phi_{-1,n},$$

onde $h_n = \left\langle \phi, \phi_{-1,n} \right\rangle$ e a convergência da série acima é no sentido de $L^2(\mathbb{R})$.

Note que ψ não é unicamente determinada pela análise de multiresolução $(V_j)_{j\in\mathbb{Z}}$ pois se ψ a satisfaz, então $\psi^{\#}$ definida por

$$\widehat{\psi^{\#}}(\xi) = \rho(\xi)\widehat{\psi}(\xi),$$

com ρ $2\pi-periódica$ e $|\rho(\xi)| = 1$, $q.s.$, também a satisfaz Em particular, podemos escolher $\rho(\xi) = \rho_0 e^{im\xi}$ com $m \in \mathbb{Z}$, $|\rho_0| = 1$, o que corresponde a uma mudança de fase e uma translação por m para ψ. Usaremos esta liberdadade para definir, ψ por

$$\psi = \sum_{n\in\mathbb{Z}} g_n \phi_{-1,n}, \text{ com } g_n = (-1)^n h_{-n+1},$$

ou ocasionalmente $g_n = (-1)^n h_{-n+1+2N}$, com $N \in \mathbb{Z}$ escolhido apropriadamente.

Embora as bases ortonormais de wavelets de interesse prático estejam associadas à análise de multiresolução, é possível construir uma base ortonormal $\{\psi_{j,k}; j,k \in \mathbb{Z}\}$ de $L^2(\mathbb{R})$, a partir de uma wavelet ψ, mas que não esteja associada a uma análise de multiresolução. Por exemplo, se definirmos ψ da seguinte maneira

$$\widehat{\psi}(\xi) = \begin{cases} (2\pi)^{-1/2}, & \text{se } \frac{4\pi}{7} \leq |\xi| \leq \pi \text{ ou } 4\pi \leq |\xi| \leq \frac{32\pi}{7}. \\ 0, & \text{caso contrário.} \end{cases}$$

Vejamos dois exemplos simples de como aplicar a receita acima para construção de bases ortonormais associadas a uma análise de multiresolução.

Exemplo: Considere ϕ a função característica do intervalo $[0,1]$, então obtemos

$$h_n = \sqrt{2}\int_{-\infty}^{+\infty} \phi(x)\overline{\phi(2x-n)}dx = \begin{cases} 1/\sqrt{2}, & \text{se } n = 0,1, \\ 0 & \text{caso contrário.} \end{cases}$$

Consequentemente, $\psi = \frac{1}{\sqrt{2}}\phi_{-1,0} - \frac{1}{\sqrt{2}}\phi_{-1,1}$ ou seja,

$$\psi(x) = \begin{cases} 1, & \text{se } 0 \le x < \frac{1}{2} \\ -1, & \text{se } \frac{1}{2} \le x < 1 \\ 0, & \text{caso contrário.} \end{cases}$$

Esta é a base de Haar.

Exemplo: Considere ϕ, tal que sua transformada de Fourier é definida por

$$\widehat{\phi}(\xi) = \begin{cases} (2\pi)^{-1/2}, & |\xi| \le 2\pi/3 \\ (2\pi)^{-1/2}\cos[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\xi|-1)], & 2\pi/3 < |\xi| \le 4\pi/3 \\ 0, & \text{caso contrário,} \end{cases}$$

onde ν é uma função de classe C^k ou C^∞, satisfazendo

$$\nu(x) = \begin{cases} 0, & x \le 0 \\ 1, & x \ge 1 \end{cases}$$

e ainda

$$\nu(x) + \nu(1-x) = 1.$$

A partir de $\widehat{\phi}$, obtemos

$$m_0(\xi) = \sqrt{2\pi}\sum_{l\in\mathbb{Z}} \widehat{\phi}(2(\xi + 2\pi l)) \tag{5.1}$$

e como

$$\widehat{\psi}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2+\pi)}\widehat{\phi}(\xi/2) \tag{5.2}$$

então

$$\widehat{\psi}(\xi) = \sqrt{2\pi}e^{i\xi/2}[\widehat{\phi}(\xi+2\pi) + \widehat{\phi}(\xi-2\pi)]\widehat{\phi}(\xi/2).$$

isto é

$$\widehat{\psi}(\xi) = \begin{cases} \frac{1}{\sqrt{2\pi}}e^{i\xi/2}\,\text{sen}[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\xi|-1)], & \frac{2\pi}{3} \le |\xi| \le \frac{4\pi}{3} \\ \frac{1}{\sqrt{2\pi}}e^{i\xi/2}\cos[\frac{\pi}{2}\nu(\frac{3}{4\pi}|\xi|-1)], & \frac{4\pi}{3} < |\xi| \le \frac{8\pi}{3} \\ 0, & \text{caso contrário.} \end{cases}$$

Esta é a base de Meyer.

Nota: Para construir novas bases ortonormais, podemos, a partir de funções escalas, construir uma análise de multiresolução, definindo $V_j = \overline{[\{\phi_{j,k}; k \in \mathbb{Z}\}]}$. A pergunta que surge naturalmente é: Que condições a função ϕ deve satisfazer para que possamos obter uma análise de multiresolução? A resposta é a seguinte, escolha ϕ tal que:

1. 1. 1. ϕ e $\widehat{\phi}$ tenham um decaimento razoável,
   2. $\phi(x) = \sum\limits_{n\in\mathbb{Z}} c_n\phi(2x-n)$, onde $\sum\limits_{n\in\mathbb{Z}} |c_n|^2 = 1$ e $0 < \alpha \leq \sum\limits_{l\in\mathbb{Z}} \left|\widehat{\phi}(\xi + 2\pi l)\right|^2 \leq \beta < +\infty$,
   3. $\int\limits_{-\infty}^{+\infty} \phi(x)dx \neq 0$.

A definição de análise de multiresolução pode ser estendida à $L^2(\mathbb{R}^n)$. S.Mallat religa este conceito a algoritmos de tratamento numérico de imagens. Se, no plano, $f(x, y)$ designa a imgem ideal(limite, cuja precisão é absoluta), $f_j(x, y) \in V_j$ representa uma aproximação cuja resolução é (de ordem de grandeza de) 2^{-j}.

O conceito de análise de multiresolução generaliza um algoritmo introduzido por G.Deslauries e S.Dubuc sob o nome de interpolação diádica.

5.2. Decomposições e reconstruções wavelet

A análise de multiresolução leva naturalmente a um esquema hierárquico e rápido para o cálculo dos coeficientes wavelet de uma dada função.

Seja $\phi \in L^2(\mathbb{R})$ uma função escala e $\{V_j\}_{j\in\mathbb{Z}}$ os subespaços gerados por ϕ. Considere $\psi \in L^2(\mathbb{R})$ e $\{W_j\}_{j\in\mathbb{Z}}$ os subespaços gerados por ψ. Assim, da segunda propriedade da definição de multiresolução, segue que qualquer função $f \in L^2(\mathbb{R})$ pode ser aproximada tanto quanto se queira por uma função $f_N \in V_N$, para algum $N \in \mathbb{Z}$. Ainda, como $V_{j-1} = V_j \oplus W_j, \forall j \in \mathbb{Z}$, f_N tem uma única decomposição:

$$f_N = f_{N+1} + g_{N+1},$$

onde $f_{N+1} \in V_{N+1}$ e $g_{N+1} \in W_{N+1}$. Repetindo este processo, temos

$$f_N = g_{N+1} + g_{N+2} + \cdots + g_{N+M} + f_{N+M},$$

onde $f_j \in V_j$ e $g_j \in W_j, \forall j \in \mathbb{Z}$, e M escolhido tal que f_{N+M} seja suficientemente "pequena". A decomposição acima, que é única, é chamada "decomposição wavelet"; e o "resto" é medido em termos da "variação" de f_{N+M} (ou mais precisamente, freqüência ou número de ciclos por unidade de comprimento). Um critério menos eficiente de parada é exigir que $\|f_{N+M}\|$ seja menor que um certo limiar.No que segue, discutiremos um algoritmo para o cálculo dos coeficientes de f_j e g_j.

Como $\phi \in V_0$ e $\psi \in W_0$, ambos contidos em V_{-1}, que é gerado por $\{\phi_{-1,k} = 2^{1/2}\phi(2x-k); k \in \mathbb{Z}\}$ segue que existem duas seqüências $(p_k), (q_k) \in l_2(\mathbb{Z})$ tais que

$$\begin{aligned} \phi(x) &= \sum_{k\in\mathbb{Z}} p_k \phi_{-1,k}(x) \\ \psi(x) &= \sum_{k\in\mathbb{Z}} q_k \phi_{-1,k}(x) \end{aligned}$$

$\forall x \in \mathbb{R}$. Por outro lado, já que $\phi(2.), \phi(2.-1) \in V_{-1}$ e $V_{-1} = V_0 \oplus W_0$, existem quatro seqüências $(a_{-2k}), (b_{-2k}), (a_{1-2k}), (b_{1-2k}) \in l_2(\mathbb{Z})$ $k \in \mathbb{Z}$ tais que:

$$\begin{aligned} \phi(2x) &= \sum_{k\in\mathbb{Z}} [a_{-2k}\phi(x-k) + b_{-2k}\psi(x-k)] \qquad (5.3) \\ \phi(2x-1) &= \sum_{k\in\mathbb{Z}} [a_{1-2k}\phi(x-k) + b_{1-2k}\psi(x-k)], \end{aligned}$$

$\forall x \in \mathbb{R}$. As duas fórmulas acima podem ser combinadas na única fórmula a seguir:

$$\phi(2x-l) = \sum_{k\in\mathbb{Z}} [a_{l-2k}\phi(x-k) + b_{l-2k}\psi(x-k)], \quad k \in \mathbb{Z}, \quad \forall x \in \mathbb{R}. \qquad (5.4)$$

que é chamada relação de decomposição de ϕ e ψ. Agora, temos dois pares de seqüência $\{(p_k), (q_k)\}$ e $\{(a_k), (b_k)\}$, as quais são todas únicas devido a relação de soma direta $V_{-1} = V_0 \oplus W_0$. Estas seqüências são usadas para formular os algoritmos de reconstrução e decomposição. O primeiro par de seqüências é utilizado no algoritmo de reconstrução, enquanto que o segundo par de seqüências e utilizado no algoritmo de decomposição.

Para descrever estes algoritmos, vamos promeiro lembrar que ambas $f_j \in V_j$ e $g_j \in W_j$ têm uma única representação em série:

$$\begin{cases} f_j(x) = \sum_{k\in\mathbb{Z}} c_k^j \phi(2^{-j}x - k), \\ \text{com } \mathrm{c}^j = (c_k^j) \in l_2(\mathbb{Z}); \end{cases} \qquad (5.5)$$

e

$$\left\{ \begin{array}{l} g_j(x) = \sum_{k\in\mathbb{Z}} d_k^j \psi(2^{-j}x - k), \\ \quad \text{com } \mathrm{d}^j = (d_k^j) \in l_2(\mathbb{Z}); \end{array} \right. \tag{5.6}$$

onde nós intencionalmente suprimimos o coeficiente de normalização $2^{-j/2}$, escrevendo $\phi(2^{-j}x - k)$ e $\psi(2^{-j}x - k)$ ao invés de usar $\phi_{j,k}$ e $\psi_{j,k}$. A partir de 5.4, segue que:

$$\phi(2^{-j+1}x - k) = \phi(2(2^{-j}x) - k) = \sum_{l\in\mathbb{Z}} [a_{k-2l}\phi(2^{-j}x - l) + b_{k-2l}\psi(2^{-j}x - l)]. \tag{5.7}$$

Como $f_j + g_j \in V_j \oplus W_j = V_{j-1}$ então tem-se de 5.5,5.6 e 5.7, que:

$$\begin{aligned} \sum_{k\in\mathbb{Z}} c_k^{j-1}\phi(2^{-j+1}x - k) &= \sum_{k\in\mathbb{Z}} c_k^{j-1} \left\{ \sum_{l\in\mathbb{Z}} [a_{k-2l}\phi(2^{-j}x - l) + b_{k-2l}\psi(2^{-j}x - l)] \right\} = \\ = \sum_{l\in\mathbb{Z}} &\left\{ \left[\sum_{k\in\mathbb{Z}} c_k^{j-1} a_{k-2l} \right] \phi(2^{-j}x - l) + \left[\sum_{k\in\mathbb{Z}} c_k^{j-1} b_{k-2l} \right] \phi(2^{-j}x - l) \right\} = \\ = \sum_{l\in\mathbb{Z}} &[c_l^j \phi(2^{-j}x - l) + d_l^j \psi(2^{-j}x - k)]. \end{aligned}$$

Assim, obtemos que:

$$\left\{ \begin{array}{l} c_l^j = \sum_{k\in\mathbb{Z}} c_k^{j-1} a_{k-2l} \\ d_l^j = \sum_{k\in\mathbb{Z}} c_k^{j-1} b_{k-2l} \end{array} \right.$$

Observe que ambas as seqüências c^j e d^j são obtidas de c^{j-1}, usando as seqüências de decomposição como pesos. Este é o chamado **algoritmo de decomposição**.

Reciprocamente, como $f_j \in V_j = V_{j+1} \oplus W_{j+1}$, segue que:

$$\sum_{k\in\mathbb{Z}} c_k^j \phi(2^{-j}x - k) = \sum_{l\in\mathbb{Z}} [c_l^{j+1}\phi(2^{-j-1}x - l) + d_l^{j+1}\psi(2^{-j-1}x - l)] \tag{5.8}$$

No entanto tem-se a partir de 5.3:

$$\begin{aligned} \phi(2^{-j-1}x - l) &= \sum_{i\in\mathbb{Z}} p_i \phi(2(2^{-j-1}x - l) - i) = \sum_{i\in\mathbb{Z}} p_i \phi(2^{-j}x - 2l - i) \\ &= \sum_{k\in\mathbb{Z}} p_{k-2l}\phi(2^{-j}x - k) \end{aligned} \tag{5.9}$$

$$\begin{aligned}\psi(2^{-j-1}x-l) &= \sum_{i\in\mathbb{Z}} q_i\phi(2(2^{-j-1}x-l)-i) = \sum_{i\in\mathbb{Z}} q_i\phi(2^{-j}x-2l-i)\\ &= \sum_{k\in\mathbb{Z}} q_{k-2l}\phi(2^{-j}x-k) \qquad (5.10)\end{aligned}$$

Assim, substituindo 5.9 e 5.10 em 5.8, obtemos:

$$\begin{aligned}\sum_{k\in\mathbb{Z}} c_k^j\phi(2^{-j}x-k) &= \sum_{l\in\mathbb{Z}}\left\{c_l^{j+1}\left[\sum_{k\in\mathbb{Z}} p_{k-2l}\phi(2^{-j}x-k)\right]+\right.\\ &\quad \left.+\, d_l^{j+1}\left[\sum_{k\in\mathbb{Z}} q_{k-2l}\phi(2^{-j}x-k)\right]\right\} \qquad (5.11)\\ &= \sum_{k\in\mathbb{Z}}\left\{\sum_{l\in\mathbb{Z}}[c_l^{j+1}p_{k-2l}+d_l^{j+1}q_{k-2l}]\right\}\phi(2^{-j}x-k).\end{aligned}$$

Logo de 5.11, concluímos que:

$$c_k^j = \sum_{l\in\mathbb{Z}}[c_l^{j+1}p_{k-2l}+d_l^{j+1}q_{k-2l}],$$

ou seja, podemos reconstruir a seqüência c^j a partir das seqüências c^{j+1} e d^{j+1}, utilizando as seqüências de reconstrução como pesos. Este é o chamado **algoritmo de reconstrução**.

Em engenharia elétrica estes algoritmos são chamados passos de análise e síntese de um esquema de filtros subbandas, com reconstrução exata.

6. Referências

[1] Chui, C.K. An Introduction to wavelets. Academic Press, London, 1992.
[2] Daubechies, I. Ten Lectures on Wavelets. CBMS-NSF Series in Appl. Math. SIAM publi., Philadelphia, 1992.
[3] Meyer, Y. Ondelettes et Opérateurs. Hermann, Paris, 1990.
[4] Gomes, S.M. ; Cortina, E. Wavelet Transform: A Local Time-Frequency Analysis. 40^0 Seminário Brasileiro de Análise, 1994.