

Identificação de Sistemas Utilizando Wavelets

LEONARDO C. ARAUJO¹

¹CEFALA — Center for Research on Speech, Acoustics, Language and Music

leoca@cefala.org

Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

Abstract. A identificação paramétrica de sistemas variantes no tempo (TV) pode ser efetuada se cada coeficiente TV puder ser expandido como um conjunto finito de uma sequência base. O problema se torna invariante no tempo com relação aos parâmetros da expansão. O uso de uma base de funções wavelets é adequado devido a sua flexibilidade em capturar as características do sinal em diferentes escalas e instantes de tempo.

1 Introdução

Um modelo matemático para sistemas dinâmicos não pode sempre ser inferido com base nos primeiros princípios. Para sistemas complexos, em particular, a tentativa axiomática de modelagem utilizando leis elementares para discrever as equações diferenciais que regem a dinâmica do sistema é uma aproximação que se torna, em geral, impossível de ser seguida. Uma alternativa para esse problema é determinar o modelo para o sistema baseando-se apenas na relação entrada/saída medida, o que chamamos de identificação de sistemas.

A identificação de sistemas variantes no tempo tem se tornado importante no modelamento de sistemas caóticos. Nas últimas décadas, a descoberta do caos tem impulsionado os estudos de sistemas dissipativos. O modelamento tem uma grande influência na análise, entendimento e controle dos sistemas caóticos. Um modelo para um sistema caótico, pode, por exemplo, fornecer uma quantidade ilimitada de dados, sendo assim possível a computação mais precisa de invariantes dinâmicas tais como dimensões fractais. Um modelo pode também ser utilizado para prever mudanças estruturais na estabilidade do sistema, conhecidas como bifurcações.

Grande parte dos sinais encontrados em aplicações não satisfazem o pressuposto de estacionariedade, o que explica o crescente interesse em processamento de sinais não estacionários, incluindo a representação tempo-frequência e a análise espectral variante no tempo (TV).

Modelos variantes no tempo surgem naturalmente na análise de diversos sinais da natureza, tais como, sinais acústicos da fala (devido à constante mudança na configuração do trato vocal), canais de comunicação (devido às mudanças do meio de propagação, dos diversos obstáculos, etc), sinais sísmicos da Terra, etc. Os sistemas variantes no tempo aparecem em diversas aplicações de identificação de sistemas, tais como cancelamento de eco, estimação de tempo de atraso, controle e processamento de arrays.

A abordagem mais popular ao problema de identificação de sistemas variantes no tempo é empregar um algoritmo adaptativo e, assumindo que a variação no tempo é lenta, ser capaz de rastrear a trajetória do sistema. A popularidade dos algoritmos adaptativos é um indicativo da importância da identificação de sistemas variantes no tempo. Apesar da sua ampla utilização, esses algoritmos não são capazes de lidar com sistemas que variam muito rapidamente. Se os coeficientes mudarem muito rapidamente em relação ao tempo de convergência do algoritmo, o algoritmo adaptativo não será capaz de rastrear a evolução do sistema.

Para contornar esse problema, um modelamento mais explícito dos coeficientes é necessário. Uma tentativa é impor uma estrutura probabilística nas trajetórias dos coeficientes e considerá-los como processos estocásticos. Então, os coeficientes são estimados com uma filtragem de Kalman. No entanto, não fica claro nesta abordagem como determinar o modelo apropriado para as trajetórias dos coeficientes e/ou como determinar os parâmetros.

Uma abordagem diferente, a qual nos focaremos, é expandir cada coeficiente variante no tempo em um conjunto de seqüências de base. Se a evolução no tempo de cada coeficiente puder ser bem aproximada pela combinação de pequenos números de seqüências de bases, então, a tarefa de identificação é equivalente à estimação de parâmetros nessa expansão. Esse problema de estimação de parâmetros pode ser resolvido pelo método dos mínimos quadrados, por exemplo.

Este método possui dois problemas quando aplicados a sistemas variantes no tempo genéricos: uma classe de funções de base precisa ser escolhida e então devemos selecionar as mais significantes dentre essas funções. Diversas classes foram propostas, mas não existe ainda nenhuma indicativo ou regra para selecionar qual classe deve ser selecionada. Para o problema de escolha das funções significantes dentro de uma classe, tem sido adotado simplesmente truncar a expansão numa certa ordem.

Alguns trabalhos foram desenvolvidos buscando responde a essas duas questões. São propostas diferentes bases de wavelets para melhor modelar as características locais e globais das trajetórias do sistema, são também dados alguns guias para a escolha da base ótima para um dado problema. A escolha das funções mais significantes da base é feita através de um teste estatístico aplicando para tanto o F-teste de Aström.

2 Wavelets em Identificação de Sistemas

As wavelets são funções que, dada uma família destas, elas formam uma base ortonormal para o subespaço $L^2(\mathbb{R})$. Uma propriedade interessante destas funções é que todas elas possuem suporte compacto^{1 2} (são diferentes de zero apenas em um intervalo finito) e satisfazem uma relação de recorrência em que é possível, através de dilatações e compressões de uma função chamada wavelet mãe, obter versões dessa wavelet mãe em escalas diferentes e assim obter uma família de wavelets que formará uma base para o $L^2(\mathbb{R})$. É devido a essas propriedades que as wavelets são capazes de promover uma representação de um sinal em diferentes escalas e instantes de tempo, podendo assim estabelecer uma melhor análise das características do sinal.

A representação utilizando wavelets se mostrou ser assintoticamente próxima do ótimo no que diz respeito a taxa de convergência que é igual a melhor taxa obtida usando esquemas de aproximação não-lineares. Além disso, a boa localização tempo-frequência das wavelets as tornam candidatas ideais para implementar modelos com estruturas adaptativas.

Dentre as diversas famílias de wavelets as B-Splines se mostraram particularmente adequadas para o caso de identificação de sistemas. As Splines se sobrepõem as demais wavelets no que tange a taxa de aproximação. Isto significa que menos níveis de resolução são necessários para aproximar uma função com uma certa precisão. Como cada nível a mais de resolução acrescentado no processo de decomposição dobra o esforço computacional necessário, é evidente então a razão da importância da escolha da família de wavelets.

As B-Splines são funções que são polinômios por partes (concatenação de polinômios). Elas podem ser facilmente computadas para dados que não estejam necessariamente espaçados uniformemente. As wavelets B-Splines possuem suporte local e fornecem uma localização

¹Uma função tem suporte compacto se ela for zero fora de um conjunto compacto. Uma função com suporte compacto só é interessante em um domínio limitado. Podemos também dizer que uma função é de suporte compacto se o seu suporte é um conjunto compacto.

²Um conjunto S é compacto se, para qualquer seqüência de elementos X_1, X_2, \dots em S , uma subseqüência pode sempre ser extraída tendendo no limite a um elemento X de S . Conjuntos compactos são fechados e limitados, e essas condições os caracterizam em um espaço de dimensão finita.

tempo frequência quase ótima. O modelo da estrutura resultante é espacialmente adaptativo. Irregularidades locais de uma dada função podem ser aproximados através, mas economicamente, utilizando apenas um pequeno número significativo de funções de base.

3 Wavelets

Wavelets é o nome dado a uma família de funções que são geradas a partir de compressões e dilatações de uma função chamada wavelet mãe. A primeira menção às wavelets apareceu em um trabalho de A. Haar em 1909 [1]. As wavelets de Haar caíram no anonimato por vários anos, até que em meados da década de 30, foi resgatada por Paul Levy que investigava o movimento Browniano e mostrou que as funções da base de Haar são melhores do que as da base de Fourier para estudar os pequenos e complexos detalhes do movimento Browniano.

Por mais alguns anos as wavelets de Haar permaneceram como sendo as únicas wavelets conhecidas. Até que em 1985, estudando processamento digital de imagens, Stephane Mallat deu um grande impulso às wavelets sendo seguido por Yves Meyer que propôs a primeira wavelet não-trivial (suave). As wavelets de Meyer, ao contrário das wavelets de Haar, são continuamente diferenciáveis, mas no entanto não possuem suporte compacto. Poucos anos mais tarde, Ingrid Daubechies inspirou-se nos trabalhos de Mallat para para construir um conjunto de bases ortonormais de wavelets suaves, com suportes compactos.

3.1 Definição

Wavelet é uma função $\psi(x) \in L^2(\mathbb{R})$ que, a partir de uma relação de translação e compressão, é capaz de gerar uma família ortonormal de funções wavelets que é uma base para o espaço $L^2(\mathbb{R})$. A relação de translação e dilatação é definida como:

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k) \quad (1)$$

no qual $j, k \in \mathbb{Z}$. Deve-se ressaltar que a constante de dilatação pode ser qualquer número real, mas usualmete é adotado o 2 por possibilitar maior adequação à implementação computacional dos algoritmos.

A cada função wavelet temos a elas associadas a sua respectiva função escala. Como veremos posteriormente na sessão sobre Análise de Múltiplas Resoluções (ARM), um sinal originalmente em uma dada resolução pode ser progressivamente decomposto em versões de menor resolução, sendo então projetado num subespaço gerado pelas funções escala de menor resolução. Os detalhes, a informações perdidas nesse processo de projeção (nessa diminuição de resolução do sinal) podem ser obtidas pela projeção do mesmo sinal original no subespaço gerado pelas funções wavelets associadas às funções escala de menor resolução.

Teremos então a decomposição do sinal em duas parcelas: uma parcela de baixa resolução e outra constituída pelos detalhes perdidos nessa mudança de resolução.

Cada função wavelet assim como cada função escala associada pode ser totalmente caracterizada pelos seus coeficientes de filtros a elas associadas. Às funções escala temos os coeficientes de filtro g e às wavelets os coeficientes h . Assim como podemos determinar uma função a partir da outra, podemos também determinar os coeficientes de filtros de uma a partir da outra. Os coeficientes de filtros possuem uma grande importancia, pois eles determinam completamente as características das funções, tanto que o valor da função em cada ponto do domínio pode ser determinado conhecendo-se apenas os coeficientes de filtro.

3.2 Wavelet de Haar

A mais simples das funções wavelets é a wavelet de Haar. A função de Haar é definida como:

$$\psi(x) = \begin{cases} 1, & \text{se } x \in [0, \frac{1}{2}) \\ -1, & \text{se } x \in [\frac{1}{2}, 1) \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

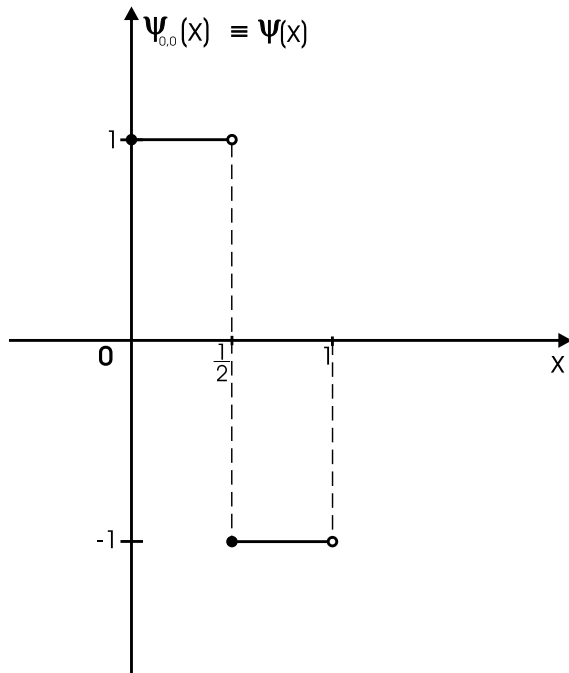


Figura 1: Wavelet de Haar.

Dada a função de Haar, é possível mostrar que a família gerada a partir de sua translação e dilatação é uma base ortonormal de $L^2(\mathbb{R})$.

A função de Haar possui seu suporte restrito ao intervalo $[0, 1)$, ou seja, este é o intervalo no qual a função é

não nula ($supp(\psi) = [0, 1)$). Além disso, a função de Haar possui norma um.

$$\begin{aligned} \|\psi(x)\|^2 &= \langle \psi(x), \psi(x) \rangle \\ &= \int_{-\infty}^{\infty} \psi^2(x) dx \\ &= 1 \end{aligned} \quad (3)$$

Dada uma função obtida por translação e dilatação da função de Haar: $\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k)$, o suporte desta é dado por:

$$supp(\psi_{j,k}(x)) = \{x \in \mathbb{R} / 0 \leq 2^{-j}x - k < 1\} \quad (4)$$

então,

$$supp(\psi_{j,k}) = [2^j k, (k+1)2^j] \quad (5)$$

e o comprimento do suporte é 2^j .

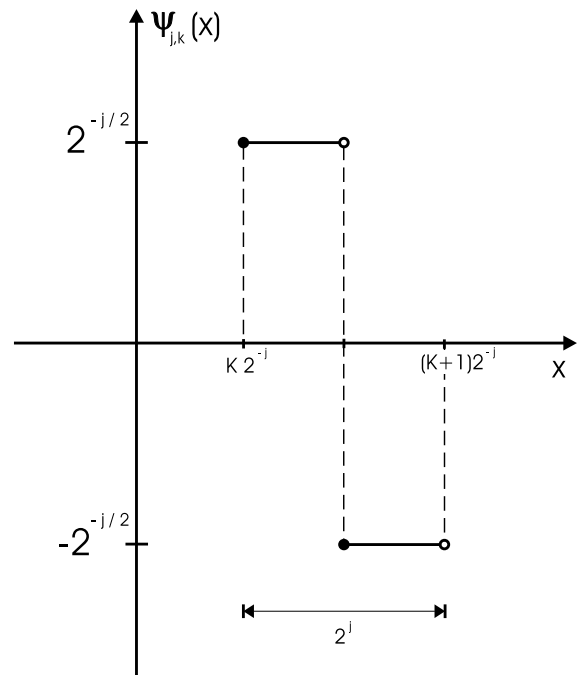


Figura 2: Wavelets de Haar obtidas por dilatações e translações.

Para demonstrar a ortogonalidade do conjunto das funções $\psi_{j,k}$, deve-se mostrar que dadas duas funções $\psi_{j,k}$ quaisquer estas serão ortogonais. Para tanto, fixa-se um j qualquer e prova-se a ortogonalidade nestas condições. Em uma segunda etapa, demonstra-se a ortogonalidade para diferentes valores de j . Feitas ambas etapas, estará provado que para quaisquer j e k a família de funções gerada é ortogonal. Para mostrar que essas funções são ortonormais, basta agora mostrar que quaisquer uma das funções possui norma unitária.

3.3 B-Slines

As B-Splines representam um exemplo em particular de uma base formada a partir de convoluções. As funções B-Splines são definidas a partir de n convoluções do pulso retângular $\beta^0(\cdot)$ que é a B-Spline de ordem zero.

$$\beta^0(x) = \begin{cases} 1, & \text{se } |x| < 0.5 \\ 0.5, & \text{se } |x| = 0.5 \\ 0, & \text{caso contrário} \end{cases} \quad (6)$$

As B-Slines de ordem n , $\beta^n(\cdot)$, são definidas:

$$\beta^n(x) = \underbrace{\beta^0(x) * \beta^0(x) * \dots * \beta^0(x)}_{n \text{ vezes}} \quad (7)$$

Existe também uma expressão explícita que pode ser demonstrada por indução:

$$\beta^n(x) = \frac{1}{n!} \sum_{k=0}^{n+1} C_k^{n+1} (-1)^k (x + \frac{n+1}{2} - k)_+^n \quad (8)$$

onde C_k^{n+1} é o coeficiente binomial, e a função x_+^n é definida por:

$$x_+^n = \begin{cases} x^n, & \text{se } x \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (9)$$

A B-Spline de ordem 3, por exemplo, possui a seguinte expressão:

$$\beta^3(x) = \begin{cases} (4 - 6x^2 + 3x^3)/6, & \text{se } 1 > x \leq 0 \\ (2 - |x|)^3/6, & \text{se } 2 > x \leq 1 \\ 0, & \text{caso contrário} \end{cases} \quad (10)$$

Os filtros correspondentes para esta B-Spline são: $1/6$, $2/3$ e $1/6$.

As B-Splines possuem algumas propriedades interessantes que são listadas abaixo:

1. A derivada da B-Spline de ordem n em relação a x pode ser expressa como uma soma de B-Slines de ordem $n - 1$ avaliadas em deslocamentos de x .

$$\frac{d}{dx} \beta^n(x) = \beta^{n-1}(x + 0.5) - \beta^{n-1}(x - 0.5)$$

2. A integral de um B-Spline de ordem n pode é igual a soma infinita de B-Splines de ordem $n + 1$.

$$\int_{-\infty}^x \beta^n(t) dt = \sum_{k=0}^{\infty} (\beta^{n+1}(x - 0.5 - k))$$

3. A transformada de Fourier está relacionada a convolução de ordem $n + 1$ das B-Splines.

$$\hat{\beta}^n(\omega) = \left(\frac{\text{sen}(\frac{\omega}{2})}{\frac{\omega}{2}} \right)^{n+1}$$

4. As funções B-Splines possuem suporte compacto.

5. As B-Splines são as menores splines polinomiais.

6. $\beta^n(x)$ é um polinômio por partes de ordem n

$$\beta^n(x) = \frac{1}{n!} \sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k (x - k + \frac{n+1}{2})_+^n$$

onde x_+^n é conforme já definido acima.

3.4 Análise de MultiResolução

Dadas funções wavelets bases para o espaço $L^2(\mathbb{R})$ temos a decomposição desse espaço em subespaços encaixantes associados as diversas resoluções.

$$\dots V_0 \subset V_1 \subset \dots \subset V_j \subset \dots \quad (11)$$

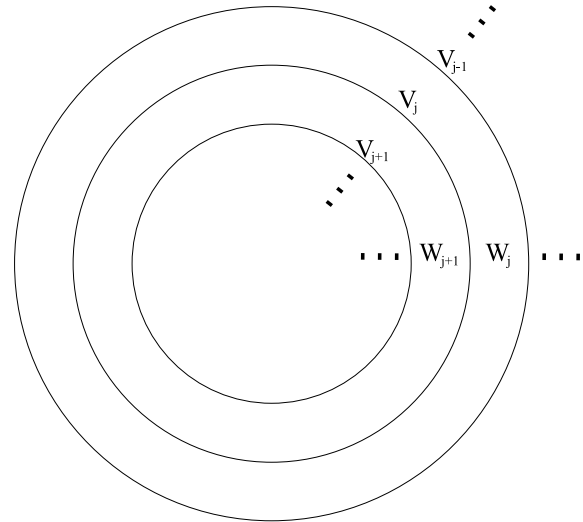


Figura 3: Seqüência de Subespaços encaixantes (ARM).

Um subespaço de maior resolução V_{j-1} pode ser expresso como uma soma de um subespaço de menor resolução V_j com um subespaço W_j com os detalhes, as informações, que as quais se faz um salto de uma resolução para a outra.

$$V_{j-1} = V_j \oplus W_j \quad (12)$$

Temos então um particionamento de V_j em uma parcela de baixas e outra de altas frequências, V_{j-1} e W_j , respectivamente. O espaço $L^2(\mathbb{R})$ pode ser decomposto nos subespaços W_j 's da seguinte forma:

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j \quad (13)$$

Dada um função f em subespaço V , esta pode ser aproximada com a precisão desejada pela sua projeção $f_j = P_j f$ no subespaço V_j . No caso limite temos: $\lim_{j \rightarrow \infty} f_j = f$. Como vimos acima, um subespaço V_j pode ser dividido de forma estável em uma parcela de altas e outra de baixas frequências, W_j e V_{j-1} , respectivamente. Desta forma, a projeção de f em V_j pode ser escrita como a soma das projeções nesses dois subespaços que são partes e formam V_j , $P_j f = P_{j-1} f + Q_j f$, onde $Q_j f$ representa a projeção de f no subespaço W_j .

Cada um dos subespaços V_j 's é gerado pelas funções de escala de diferentes níveis de resolução. É possível então obtermos uma análise nas múltiplas resoluções, fazendo para tanto a projeção nos diversos subespaços V_j . Para existir a análise em múltiplas resoluções (ARM) é necessário termos uma seqüência de subespaços fechados de $L^2(\mathbb{R})$, V_j onde j é um número inteiro satisfazendo às seguintes propriedades:

1. $\forall j \in \mathbb{Z}, V_j \subset V_{j-1}$
2. $\bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$
3. $\overline{\bigcup_{j=-\infty}^{+\infty} V_j} = L^2(\mathbb{R})$
4. $\forall f \in L^2(\mathbb{R}), \forall j \in \mathbb{Z}, f \in V_j \Leftrightarrow f(2^j \cdot) \in V_0$
5. $\forall f \in L^2(\mathbb{R}), \forall j \in \mathbb{Z}, f \in V_j \Leftrightarrow f(\cdot - n) \in V_j$
6. Existe uma função escala $\phi \in V_0$ tal que $\phi_{j,k}(x) = 2^{-j/2} \phi(2^{-j}x - k)$

Uma vez dada uma análise de resolução múltipla, podemos associar a ela uma wavelet:

$$\psi = \sum_n (-1)^{n-1} \overline{h_{-n-1}} \phi_{-1,n} \quad (14)$$

As funções wavelets e as funções escalas formam bases para os subespaços W_j e V_j . O conjunto $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ é uma base para o subespaço W_j . O conjunto $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$ é uma base para V_j . Desta forma, uma dada função f em V pode ser expressa como uma expansão em wavelets:

$$f(x) = \sum_k c_{l,k} \phi_{l,k}(x) + \sum_{j=l+1}^{\infty} \sum_k d_{j,k} \psi_{j,k}(x) \quad (15)$$

É interessante notar que o nível de resolução até qual ponto será feita a decomposição l é um parâmetro que deve ser escolhido.

4 Expansão Wavelet de Sistemas Variantes no Tempo

Vamos tomar inicialmente um modelo de sistema mais simples, um modelo de tempo discreto de um processo autorregressivo com parâmetros variantes no tempo, TV-AR. Tal

processo é descrito pela seguinte equação:

$$y(n) = \sum_{k=1}^p a(n;k) y(n-k) + e(n) \quad (16)$$

no caso acima temos um modelo TV-AR de ordem p onde $a(n;k)$ são os coeficientes variantes no tempo, que são justamente o objeto alvo do processo de identificação. Se o modelo assim descreve um sistema físico, então é natural considerarmos que seja um sistema causal e estável (é importante notar que isso não equivale a condição de termos todos os polos de $a(n;k)$ dentro do círculo de raio unitário para todo n).

Se cada um dos coeficientes $a(n;k)$ for expresso como uma combinação linear de algumas seqüências de base $f_l(n)$, $l = 1, \dots, L$ temos:

$$a(n;k) = \sum_{l=1}^L \xi_{kL} f_l(n) \quad (17)$$

onde ξ_{kL} são coeficientes invariantes no tempo. Substituindo no problema inicial, temos:

$$y(n) = \sum_{k=1}^p \left(\sum_{l=1}^L \xi_{kL} f_l(n) \right) y(n-k) + e(n) \quad (18)$$

Com essa mudança no problema transformamos-o num outro problema onde não há mais nenhum parâmetro variando com o tempo. Tomando a equação acima para vários instantes de tempo, formamos um sistema linear de equações que pode ser resolvido pelo método dos mínimos quadrados.

Se escolhermos as seqüências de base como sendo as funções wavelets teremos que os coeficientes $a(n;k)$ expressos como:

$$a(n;m) = \sum_{k=k_0}^{K_{j_0}} \alpha_{j_0,k}^{(m)} \phi_{j_0,k}(n) + \sum_{j=j_0}^J \sum_{k=k_0}^{K_j} \beta_{j,k}^{(m)} \psi_{j,k}(n) \quad (19)$$

Substituindo na equação para $y(n)$ e tomando os valores para M amostras, podemos re-escrever na forma matricial o sistema de equações, na seguinte forma:

$$Y = H\Theta + \xi \quad (20)$$

A questão agora se resume a determinar o vetor Θ que é o vetor com os coeficientes de wavelets que queremos determinar:

$$\Theta^T = [\theta_0 \alpha^T \beta^T] \quad (21)$$

$$\alpha^T = [\alpha_{j_0, k_0}^{(1)}, \alpha_{j_0, k_0+1}^{(1)}, \dots, \alpha_{j_0, k_{j_0}}^{(1)} \vdots \alpha_{j_0, k_0}^{(2)}, \alpha_{j_0, k_0+1}^{(2)}, \dots, \alpha_{j_0, k_{j_0}}^{(2)} \vdots \dots \vdots \alpha_{j_0, k_0}^{(M)}, \alpha_{j_0, k_0+1}^{(M)}, \dots, \alpha_{j_0, k_{j_0}}^{(M)}] \quad (22)$$

$$\beta^T = [\beta_{j_0}^T, \beta_{j_0+1}^T, \dots, \beta_J^T] \quad (23)$$

$$\beta_j^T = [\beta_{j, k_0}^{(1)}, \beta_{j, k_0+1}^{(1)}, \dots, \beta_{j, k_{j_0}}^{(1)} \vdots \beta_{j, k_0}^{(2)}, \beta_{j, k_0+1}^{(2)}, \dots, \beta_{j, k_{j_0}}^{(2)} \vdots \dots \vdots \beta_{j, k_0}^{(M)}, \beta_{j, k_0+1}^{(M)}, \dots, \beta_{j, k_{j_0}}^{(M)}] \quad (24)$$

A matriz H é a matriz que depende dos valores atrasados da saída e dos valores que assume as wavelets.

$$H = \begin{bmatrix} 1 & A(1) & B(1) \\ 1 & A(2) & B(2) \\ \vdots & \vdots & \vdots \\ 1 & A(N) & B(N) \end{bmatrix} \quad (25)$$

$$A(n) = [P(n) \otimes \Gamma(n)] \quad (26)$$

$$P(n) = [y(n-1), y(n-2), \dots, y(n-M)] \quad (27)$$

$$\Gamma(n) = [\phi_{j_0, k_0}(n), \phi_{j_0, k_0+1}(n), \dots, \phi_{j_0, k_{j_0}}(n),] \quad (28)$$

$$B(n) = [B_{j_0}^T(n), B_{j_0+1}^T(n), \dots, B_J^T(n),] \quad (29)$$

$$B_j(n) = P(n) \otimes \Lambda_j(n) \quad j = j_0, j_0 + 1, \dots, J \quad (30)$$

$$\Lambda_j(n) = [\varphi_{j, k_0}(n), \varphi_{j, k_0+1}(n), \dots, \varphi_{j, k_J}(n)] \quad (31)$$

Os parametros do vetor θ podem ser determinados utilizando para tanto o método dos mínimos quadrados ou uma rotina de predição de erro. O número possível de parâmetros no modelo é, no entanto, muito grande e por isso, a escolha do modelo adequado é de vital importância nos problemas de identificação de sistemas.

4.1 Como Selecionar os Níveis de Multiresolução?

Teoricamente, a expansão em wavelets contem um número infinito de termos, associados a infinitos níveis de resolução. No entanto, na prática, apenas um número finito de funções de base são necessárias para aproximar uma dada função não-linear. A escolha da base apropriada é um passo muito importante na construção de um mapeamento parsimonioso de uma função f do espaço de regressão para o espaço de saída. Iremos então truncar o conjunto que forma a base de forma a incluir apenas as funções escala de uma dada resolução inicial j_0 e as funções wavelets com os níveis de resolução entre j_0 e $J = j_{max}$, sendo J o nível de resolução final. Além do mais, consideraremos apenas as funções cujo suporte se sobrepoem às amostras da função em questão que será aproximada. A resolução máxima será tal que ao menos uma amostra esteja dentro do suporte da wavelet correspondente, ou seja, o suporte da wavelet de resolução máxima deverá ser maior ou igual ao período de amostragem da função/sinal a ser aproximada.

Vamos supor que tenhamos N amostras de um sistema de entrada e saída

$$y_i = f(x_i) + e_i, \quad i = 1, 2, \dots, N \quad (32)$$

onde $x_i = [x_{1,i}, x_{2,i}, \dots, x_{d,i}]^T \in [0, 1]^d$, $f \in L^2(\mathbb{R})$ pode ser definida a partir de x_i e y_i usando para tanto uma expansão em multiplas resoluções usando wavelets sujeito a algumas restrições. Se x_i é uniformemente distribuído no hiper-cubo $[0, 1]^d$, então as leis estatísticas do tipo log-log oferecem um intervalo grosseiro para selecionar a escala deresolução mais alta j_{max} [2]

$$\frac{N}{\ln N} \leq 2^{d \cdot j_{max}} \leq \frac{2N}{\ln N} \quad (33)$$

A tarefa de escolher o nível inicial de resolução pode ser simplificada se os dados de entrada/saída forem normalizados ao intervalo unitário $[0, 1]$. Neste caso a resolução inicial $j = 0$ pode ser escolhida para todos submodelos. Se os dados não estiverem normalizados, $j = \{j_1, \dots, j_{d(i)}\}$ é um multiindex que leva em consideração o universo de percurso de cada entrada e saída. Inicialmente o número de resoluções é escolhido como sendo $N_s = 1$.

Se lidamos com wavelets e funções de escala com suporte compacto, como é o caso das wavelets de Daubechies e das B-Splines, a posição de cada função da base é determinada pelo índice de translação $k = \{k_1, \dots, k_{d(i)}\}$. Então, apenas um número finito de funções de base terá relevância para uma estrutura em particular. Apenas a função que possuir pontos de dados dentro do seu suporte poderá ser considerada como candidata para o modelo. Neste caso, não há necessidade de resolver o problema de posicionamento dos centros das funções de base, que está normalmente associado às wavelets de base radiais.

Referências

- [1] A. Haar, “Zur theorie der orthogonalen funktionensysteme,” *Math. Ann.*, vol. 69, pp. 331–371, 1910.
- [2] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, “Nonlinear black-box modeling in system identification: a unified overview,” *Automatica*, vol. 31, pp. 1691–1724, 1995.
- [3] S. A. Billings and D. Coca, “Discrete wavelet models for identification and qualitative analysis of chaotic systems,” *International Journal of Bifurcation and Chaos*, vol. 9, no. 7, pp. 1263–1284, 1999.
- [4] H.-L. Wei and S. A. Billings, “Identification of time-varying systems using multiresolution wavelet models,” *International Journal of Systems Science*, vol. 33, no. 15, pp. 1217–1228, 2002.
- [5] M. K. Tsatsanis and G. B. Giannakis, “Time-varying system identification and model validation using wavelets,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3512–3523, 1993.