CAPÍTULO 4

Identificación de modelos de Volterra basados en la descomposición PARAFAC

La gran utilización de los modelos de Volterra en numerosos campos hace que el estudio sobre éstos cobre especial importancia. La complejidad paramétrica de la que se caracterizan se convierte en su principal inconveniente, esto es debido a que el número de coeficientes en sus núcleos es elevado.

En este capítulo, se expondrá un método que reduce la complejidad paramétrica con la ayuda de la descomposición de PARAFAC, explicada en el capítulo anterior. Como punto de salida, se considera a los núcleos de Volterra de orden superior a la unidad como tensores simétricos. Los modelos de Volterra-PARAFAC así obtenidos pueden ser vistos como una serie de modelos de Wiener puestos en paralelo.

1. *Introducción*

Los modelos de Volterra en tiempo discreto, de duración finita, llamados también desarrollo en series de Volterra truncados o también modelos polinucleares no recursivos, pueden ser utilizados para representar todo sistema no lineal con memoria que se desvanece. Estos modelos han sido objeto de numerosas aplicaciones en diferentes campos, como por ejemplo la anulación del eco y del ruido, la linealización de altavoces, el modelado, la igualdad y la pre-distorsión de canales de comunicación lineales, el modelado de sistemas fisiológicos, o también el comando de sistemas no lineales.

Los modelos de Volterra que pueden ser vistos como una extensión no lineal del modelo lineal con respuesta de impulso finito (RIF), son interpretados en términos de convoluciones tridimensionales. Poseen dos propiedades interesantes: linealidad vis a vis de sus parámetros, los coeficientes de los núcleos, y la estabilidad garantizada en el sentido de entrada limitada-salida limitada. Su principal inconveniente es su gran complejidad paramétrica en razón del gran número de coeficientes contenidos en sus núcleos. Existen dos maneras principales de reducir esta complejidad: una primera aproximación consiste en desarrollar los núcleos de Volterra sobre las bases de funciones ortonormales como las funciones de Laguerre, o las bases ortonormales generalizadas, por ejemplo, las funciones de bases ortonormales caracterizadas por varios polos. Otra aproximación, consiste en considerar los núcleos de Volterra como tensores y en utilizar las descomposiciones tensoriales como la descomposición PARAFAC. El resultado son los modelos de tipo Volterra-PARAFAC, que pueden ser vistos como la puesta en paralelo de una serie de modelos de Wiener.

2. Filtros adaptativos.

Los filtros adaptativos son un sistema digital compuesto por un filtro lineal programable de entrada $x(n)$ y salida $y(n)$. Los coeficientes se reprograman de una muestra a la siguiente a través de un algoritmo de adaptación, cambiando su valor dinámicamente en el tiempo. El sistema puede ser inestable por lo que se utiliza con mucha frecuencia los sistemas FIR ya que son incondicionalmente estables.

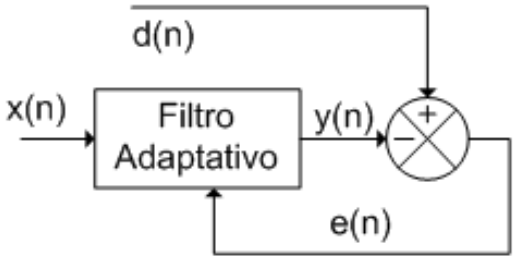

Figura nº 11 . Esquema general de filtro adaptativo.

1.1. Filtros Wiener.

Los filtros Wiener son los mejores filtros lineales de mínimos cuadrados, que pueden ser usados para predicción, estimación, interpolación, filtrado de señal y ruido, etc.

Para diseñarlos se necesita tener un conocimiento previo apropiado de las propiedades estadísticas de la señal de entrada. El problema reside en que este conocimiento generalmente no se puede obtener. En su lugar se usan filtros adaptativos, que hacen uso de los datos de entrada para aprender los datos estadísticos requeridos.

La teoría de Wiener es importante porque los filtros adaptativos que serán empleados convergen asintóticamente (en media) en la soluciones de Wiener.

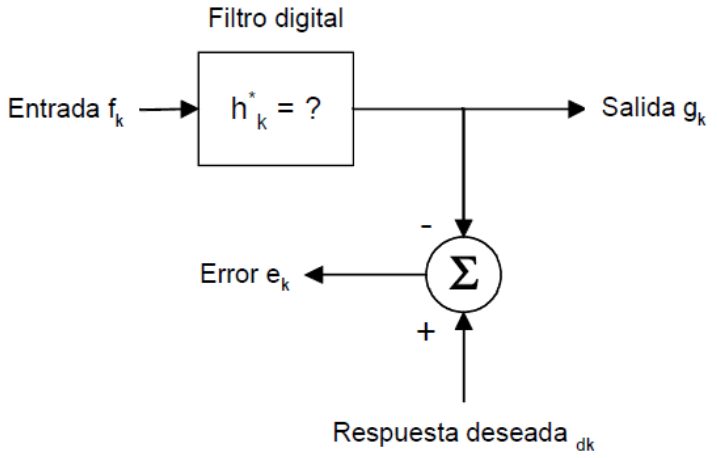

Figura nº 12 . Esquema general filtro Wiener digital.

El filtro digital tiene una señal de entrada y produce una señal de salida. Será un filtro de Wiener si su respuesta impulsiva se elige para minimizar el error cuadrático medio. El error se define como la diferencia entre la salida del filtro y la respuesta deseada:

$$e_k = d_k - g_k \tag{4.1}$$

Cuando se trabaja con filtros de Wiener, generalmente la respuesta deseada existe sólo de forma conceptual. Las propiedades estadísticas de la respuesta deseada y sus relaciones estadísticas con la señal de entrada al filtro asume que son conocidas por el diseñador. La situación es bastante diferente cuando se trata con filtros adaptativos. En éstos, la respuesta deseada existe como una señal que puede ser obtenida como entrada en tiempo real al algoritmo adaptativo, para conseguir aprender y adaptarse. Los filtros de Wiener no aprenden. Su diseño es fijo, basado en un primer conocimiento estadístico.

La respuesta impulsiva del filtro Wiener se obtiene encontrando una expresión para el error cuadrático medio y minimizándola con respecto a la respuesta impulsiva. Elevando al cuadrado en ambas partes, se obtiene:

$$e_k^2 = d_k^2 + g_k^2 - 2d_k g_k \tag{4.2}$$

Sabiendo que:

$$g_k = \sum_{l=0}^{\infty} f_{k-l} h_l \tag{4.3}$$

se puede sustituir:

$$e_k^2 = d_k^2 + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m f_{k-l} f_{k-m} - 2d_k \sum_{l=-\infty}^{\infty} f_{k-l} h_l \tag{4.4}$$

Si tomamos valor medio en ambos lados, encontramos una expresión para el error cuadrático medio (MSE, mean square error):

$$\begin{aligned} E[e_k^2] &= E[d_k^2] + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m E[f_{k-l} f_{k-m}] - 2 \sum_{l=-\infty}^{\infty} E[d_k f_{k-l}] h_l \\ &= \emptyset_{dd}(0) + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m \emptyset_{ff}(l-m) - 2 \sum_{l=-\infty}^{\infty} \emptyset_{fd}(l) h_l \end{aligned} \tag{4.4}$$

siendo $\emptyset_{dd}$la autocorrelación y $\emptyset_{fd}$ la correlación cruzada de dos señales f y d. Si derivamos con respecto a h, que es la respuesta impulsiva del filtro, e igualamos a cero para minimizar el error:

$$\sum_{l=-\infty}^{\infty} h_l^* \emptyset_{ff}(j-l) = \emptyset_{fd}(j) \tag{4.5}$$

Esta es la ecuación de Wiener-Hopf, que en forma de convolución queda:

$$h_l^* * \emptyset_{ff}(k) = \emptyset_{fd}(k) \tag{4.6}$$

Tomando transformada Z en ambas partes;

$$H^*(z)\mathbf{\Phi}_{ff}(z) = \mathbf{\Phi}_{fd}(z) \quad \text{ó} \quad H^*(z) = \frac{\mathbf{\Phi}_{fd}(z)}{\mathbf{\Phi}_{ff}(z)} \tag{4.7}$$

Con la solución de Wiener se puede encontrar la función de transferencia del filtro $H^*(z)$ a partir de la transformada Z de la función de autocorrelación de la señal de entrada, de la correlación cruzada de la señal de entrada y de la respuesta deseada. Si se sustituye esta ecuación en la expresión del error se obtiene el valor del mínimo MSE:

$$E[e_k^2]|_{min} = \emptyset_{dd}(0) - \sum_{l=-\infty}^{\infty} h_l^* \emptyset_{fd}(l)$$

(4.8)

En el presente trabajo se utilizará la medida NMSE (Normalized Mean Square Error) para presentar los resultados. Esta medida tiene la siguiente expresión.

$$NMSE = \frac{1}{M}\sum_{k=1}^{M} |e_k|^2$$

(4.9)

donde M representa el número total de muestras que contiene la señal.

3. *Los modelos de Volterra.*

Un modelo de Volterra de orden P de un sistema no lineal, de una entrada y una salida, causal, estable, con memoria que se desvanece y, por tanto, finita, es descrito con la ayuda de la relación entrada-salida siguiente:

$$y(k) = h_0 + \sum_{p=1}^{P} \sum_{m_1=1}^{M_p} \cdots \sum_{m_p=1}^{M_p} h_p(m_1, \dots, m_p) \prod_{i=1}^{P} u(k - m_i) = h_0 + \sum_{p=1}^{P} y_p(k)$$

(4.10)

donde $u(k)$ y $y(k)$ designan respectivamente las señales de entrada y de salida, P es el grado de no linealidad del modelo de Volterra, M_p es la memoria del término homogéneo de orden p representado por $y_p(k)$, y $h_p(m_1, \dots, m_p)$ es un coeficiente del núcleo de orden p. Este coeficiente es caracterizado por p índices, puede ser visto como un elemento de un tensor $\mathbb{H}_p \in \mathcal{K}^{M_p \times M_p \times \cdots \times M_p}$, de orden p, con $\mathcal{K} = \mathcal{R}$ o $\mathcal{C}$, según sean los coeficientes de los núcleos valores reales o complejos. El núcleo de orden p esta caracterizado por M_p^p coeficientes.

Como toda permutación de índices $m_1, \dots, m_p$ corresponde a un mismo producto $\prod_{i=1}^{P} u(k - m_i)$ de entradas retardadas, podemos convocar a todos los coeficientes asociados a estas permutaciones para obtener un núcleo simétrico definido como:

$$h_{p,sym}(m_1, \dots, m_p) = \frac{1}{p!} \sum_{\pi(.)} h_p(m_{\pi(1)}, \dots, m_{\pi(p)}) \tag{4.11}$$

donde $\pi(.)$ designa una permutación de índices $m_1, \dots, m_p$. El número de coeficientes independientes contenidos en el núcleo simétrico de orden p es igual a $C_p^{M_p+p-1}$. Así, por un nucleo cúbico de memoria $M_3 = 10$, el número de coeficientes es igual a $M_3^3 = 10^3$, en el caso no simétrico y $C_3^{12} = 220$ en el caso simétrico, que corresponde a una tasa de reducción de complejidad del 78%.

El núcleo de orden p se llama separable si se puede expresar como el producto de p núcleos de orden uno:

$$h_p(m_1, \dots, m_p) = \prod_{i=1}^{p} h_{m_i}^{(i)}$$

(4.12)

donde $h_{m_i}^{(i)}$ represeta el $m_i - \acute{e}simo$ elemento del núcleo de orden un $h^{(i)}$. Si el núcleo separable es también simétrico, entonces los p núcleos de orden un $h^{(i)}$ son idénticos y la expresión anterior se convierte en:

$$h_p(m_1, \dots, m_p) = \prod_{i=1}^{p} h_{m_i}^{(i)}$$

(4.13)

4. *Los modelos de Volterra-PARAFAC*

A continuación, vamos ante todo a recordar la descomposición PARAFAC. De un tensor, pues utilizaremos esta descomposición para desarrollar los núcleos de un modelo de Volterra, que nos dará el modelo Volterra-PARAFAC.

La descomposición PARAFAC de un tensor $\mathbb{H} \in \mathcal{K}^{M_1 \times M_2 \times \cdots \times M_p}$ de orden P y dimensiones $M_1 \times M_2 \times \cdots \times M_p$ se escribe bajo forma escalar como:

$$h_{m_1, m_2, \cdots, m_p} = \sum_{r=1}^{R} \prod_{p=1}^{P} a_{m_{p,r}}^{(p)} \quad m_p = 1, \dots, M_p$$

(4.14)

donde $a_{m_{p,r}}^{(p)}$ es un elemento del factor matricial $A^{(p)} \in \mathcal{K}^{M_p \times R}$, $p = 1, \dots, P$, y R es el rango del tensor.

La ecuación anterior puede ser reescrita como:

$$\mathbb{H} = \sum_{r=1}^{R} \boldsymbol{A}_{.r}^{(1)} \circ \boldsymbol{A}_{.r}^{(2)} \circ \cdots \circ \boldsymbol{A}_{.r}^{(P)} \tag{4.15}$$

donde $\circ$ designa el producto exterior y la $\boldsymbol{A}_{.r}^{(P)}$ *r-ésima* columna de la matriz $\boldsymbol{A}^{(P)}$. Esta expresión muestra que PARAFAC corresponde a una descomposición del tensor de orden P y de rango R, en una suma de R tensores de rango uno, es decir, una suma de R productos exteriores de P vectores.

Podemos indicar que para un tensor $\mathbb{H}$ de orden P de rango uno, (4.14) pasa a ser:

$$h_{m_1, m_2, \cdots, m_p} = \prod_{p=1}^{P} a_{m_p}^{(p)} \quad m_p = 1, \dots, M_p \tag{4.16}$$

donde $a_{m_p}^{(p)}$ es un elemento del factor vectorial $\boldsymbol{a}^{(p)} \in \mathcal{K}^{M_p \times R}$.

Comparando (4.7) con (4.3), podemos por tanto concluir que la descomposición PARAFAC de un núcleo de Volterra de rango uno equivale a un núcleo separable.

En el caso de un tensor $\mathbb{H}$ simétrico, de orden P, de rango R, los P factores matriciales $\boldsymbol{A}^{(P)}$, son idénticos e iguales a $\boldsymbol{A}$. De forma similar, para un tensor $\mathbb{H}$ simétrico, de orden P, de rango uno, los P factores vectoriales $\boldsymbol{a}^{(p)}$, $p = 1, \cdots, P$, son idénticos e iguales a $\boldsymbol{a}$.

La utilización de la descomposición PARAFAC de un núcleo de Volterra de orden p, de rango uno, simétrico permite reescribir el término homogéneo $y_p(k)$ de (4.10) como:

$$y_p(k) = \sum_{m_1=1}^{M_p} \cdots \sum_{m_p=1}^{M_p} h_p(m_1, \ldots, m_p) \prod_{i=1}^{P} u(k - m_i)$$
$$= \sum_{m_1=1}^{M_p} \cdots \sum_{m_p=1}^{M_p} \prod_{i=1}^{P} a_{m_i}^{(p)} u(k - m_i) = \left(\boldsymbol{u}^T(k)\boldsymbol{a}^{(p)}\right)^p \tag{4.17}$$

donde $\boldsymbol{u}^T(k) = \left[u(k-1) \cdots u(k-M_p)\right] \in \mathcal{K}^{1\times M_p}$ es el vector de regresión lineal asociado al núcleo de Volterra de orden p, de memoria M_p, y $\boldsymbol{a}^{(p)^T} = \left[a_1^{(p)} \cdots a_{M_p}^{(p)}\right] \in \mathcal{K}^{1\times M_p}$ está formado de coeficientes del factor vectorial (vector generador) de la descomposición PARAFAC de este núcleo. Nosotros deducimos que (4.17) puede ser visto como la salida de un modelo de Wiener obtenido concatenando un modelo lineal RIF, de memoria M_p, con uno no lineal sin memoria de grado p, como se ilustra con la ayuda de la Fig. 13.

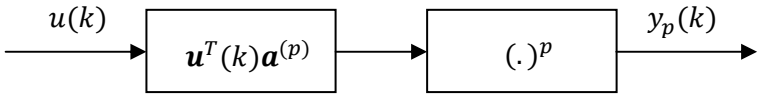

Figura nº 13. Realización del término homogéneo de orden p de un modelo de Volterra como un modelo de Wiener. Caso de un núcleo de rango uno.

Para un núcleo de Volterra de orden p, de rango r_p, simétrico, (4.17) pasa a:

$$y_p(k) = \sum_{r=1}^{r_p} \left(\boldsymbol{u}^T(k) \boldsymbol{A}_{.r}^{(p)} \right)^p$$

(4.18)

El término homogéneo de orden p puede, por tanto, estar realizado paralelizando r_p modelos de Wiener, cada uno de estos modelos siendo asociados con una columna del factor matricial de la descomposición PARAFAC del núcleo. Por consiguiente, la salida (4.10) del modelo de Volterra puede ser obtenida como la suma del término constante h_0, y de las salidas de $\sum_{p=1}^{P} r_p$ modelos de Wiener en paralelo, como se ilustra con la ayuda de la Fig. 2 para un modelo de Volterra cubica, donde $\boldsymbol{A}_{.r}^{(p)}$ designa la *r-ésima* columna del factor matricial de la descomposición PARAFAC del núcleo de orden p, $p = 2,3$, $\boldsymbol{A}_{.1}^{(1)} = [h_1(1) \cdots h_1(M_1)]^T$ y $\boldsymbol{u}_p^T(k) = \left[u(k-1) \cdots u(k-M_p)\right]$o $p = 1,2$ *y* 3.

Utilizando la descomposición PARAFAC de núcleos, la salida del modelo de Volterra-PARAFAC se escribe como:

$$y(k) = h_0 + \sum_{p=1}^{P} \sum_{r=1}^{r_p} \left(\boldsymbol{u}_p^T(k) \boldsymbol{A}_{.r}^{(p)} \right)^p$$

(4.19)

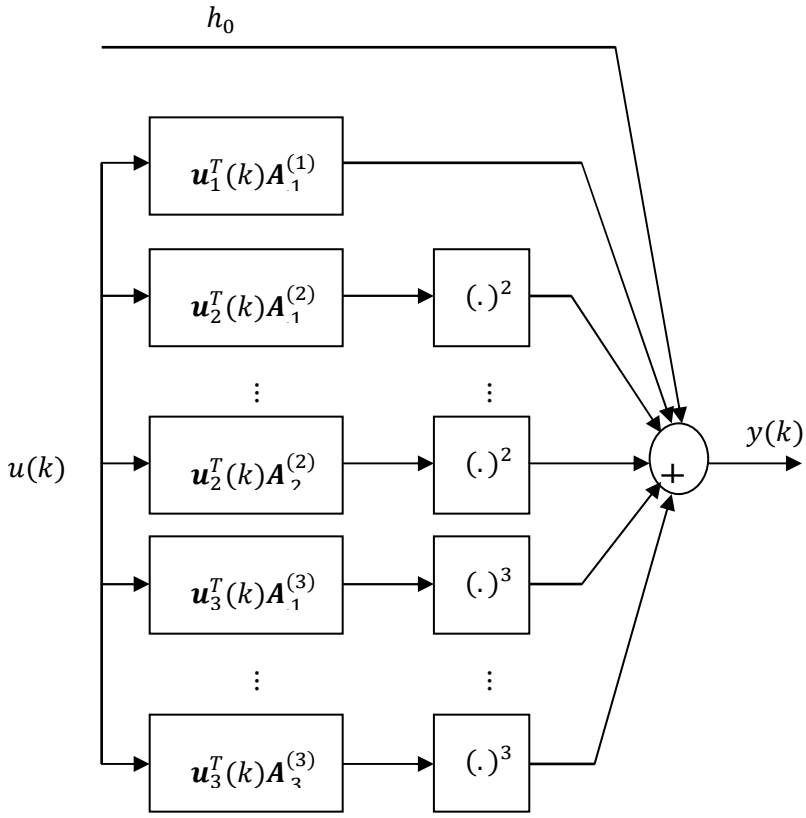

Figura nº 14. Realización de un modelo *de Volterra cubico como modelos de Wiener en paralelo.*

Por consiguiente, debido a este desarrollo la complejidad queda reducida y sólo se necesitará de un algoritmo recursivo para el cálculo de los diferentes parámetros haciendo que la salida sea cada vez lo más parecida a la entrada. Es de notar, que este desarrollo reduce complejidad pero, por otro lado, se introducen errores a la hora de simplificar el modelo. Este hecho se explicará en capítulos sucesivos.