

UNIVERSIDAD DE LA LAGUNA

**«Modelización racional en el
dominio de la frecuencia.
Una aplicación a series periódicas»**

**Autor: Sergio Fernando Alonso Rodríguez
Director: Dr. D^a. Concepción González Concepción**

Departamento de Economía Aplicada

CONCEPCIÓN GONZÁLEZ CONCEPCIÓN, CATEDRÁTICA DEL DEPARTAMENTO
DE ECONOMÍA APLICADA DE LA UNIVERSIDAD DE LA LAGUNA

CERTIFICO:

*Que la presente memoria titulada “**Modelización Racional en el Dominio de la Frecuencia. Una Aplicación a Series Periódicas**”, ha sido realizada bajo mi dirección por el Licenciado **D. Sergio Fernando Alonso Rodríguez**, y constituye su Tesis para optar al grado de Doctor por la Universidad de La Laguna.*

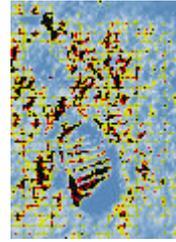
Y para que así conste, en cumplimiento de la legislación vigente y a los efectos que haya lugar, firmo el presente en La Laguna a 24 de marzo de 2003.

Quiero agradecer, en primer lugar, a la directora de este trabajo, la doctora Concepción González, su inmenso tesón y empuje. Ella, más que nadie, sabe que sin ellos esta memoria no habría visto la luz.

También quiero reconocer la labor de los compañeros de departamento que me preguntaban a cada momento y de forma insistente que cuándo iba a terminar la memoria. Gracias a Joaquín, David, Carlos, Miguel Ángel, Antonio, Carmen, Marcos, José Miguel y Macu, a los que he citado por orden de insistencia, y que me han dado un motivo más para acabarla de una vez.

A Cande y Celina, mi agradecimiento por su ánimo, por su interés. Y finalmente quiero agradecer a la empresa EMMASA, su colaboración para con este trabajo.

A mi familia



Índice

Introducción

i

Capítulo 1	<i>La aproximación lineal en la modelización de datos</i>	1
1.1	El entorno teórico: espacios vectoriales de Hilbert	3
1.2	En busca de la ‘mejor’ aproximación	6
1.3	Un primer ejemplo: aproximación en Espacios de Lebesgue, L^2	9
1.4	Aproximación de secuencias complejas finitas: la transformada discreta de Fourier	14
	<i>La transformada rápida de Fourier, FFT</i>	16
1.5	Aproximación lineal mediante retardos: modelización en l^2	19
1.6	El caso de los procesos estacionarios	22
1.7	Resumen final	27
Capítulo 2	<i>El modelo del filtro aplicado a la Transformada Discreta de Fourier</i>	29
2.1	Filtros Lineales	31
	<i>Definición y caracterización</i>	31
	<i>La z-transformada</i>	34
	<i>Estabilidad y causalidad</i>	36
2.2	<i>Predicción Lineal a través del Modelo del Filtro</i>	38
	<i>El Modelo del Filtro</i>	39
	<i>El Modelo del Filtro en \mathbf{C}^N</i>	44
	<i>Propiedades del Modelo del Filtro</i>	46
	<i>El Algoritmo de Levinson, (1947)</i>	50
2.3	<i>El Modelo del Filtro aplicado a la DFT</i>	54
	<i>Formulación del Modelo</i>	55
	<i>Una aplicación a una serie de datos simulados</i>	58
	<i>Interpretación del Modelo</i>	62
	<i>La aceleración del algoritmo</i>	68

Capítulo 3	<i>Un modelo de Padé-Laurent aplicado a la secuencia DFT de datos reales</i>	71
3.1	<i>Preliminares: álgebra de series de Laurent</i>	73
3.2	<i>Aproximantes de Padé</i>	76
	<i>Cálculo y propiedades de los coeficientes del aproximante</i>	77
3.3	<i>Aproximantes de Padé-Laurent</i>	80
3.4	<i>El Modelo del Filtro y los Aproximantes de Padé-Laurent</i>	82
	<i>El Algoritmo adaptado a la DFT de datos reales</i>	83
	<i>Ventajas e inconvenientes</i>	84
Capítulo 4	<i>Una aplicación práctica: series de consumo de agua en la ciudad de Santa Cruz de Tenerife</i>	87
4.1	<i>Las series de datos</i>	88
4.2	<i>Análisis de periodicidades</i>	90
4.3	<i>La periodicidad semanal</i>	104
4.4	<i>Estudio de los modelos</i>	106
	<i>Modelos para el depósito Alisios</i>	107
	<i>Modelos para el depósito Cueva Roja</i>	111
	<i>Conclusiones y cuestiones abiertas</i>	115
	<i>Bibliografía</i>	118

Introducción

En la literatura propia de series temporales, los modelos autorregresivos, desde un punto de vista estocástico, o los filtros lineales, desde un punto de vista determinístico, se han centrado, en última instancia, en términos de interpretar y ajustar datos procedentes de procesos estacionarios. La simplificación de éstos a través de su función de autocorrelación es la característica fundamental que posibilita este ajuste lineal. La motivación de este trabajo se inicia con el siguiente planteamiento: ¿es posible lograr un ajuste lineal por retardos de datos no necesariamente estacionarios?

Abordamos en este trabajo los modelos lineales clásicos pero en un contexto inusual. Para ello, nuestro primer empeño se centra en describir una estructura matemática que permita articular y englobar la aproximación lineal mínimo cuadrática. El entorno teórico de los espacios vectoriales de Hilbert es el punto de partida que nos permite tratar estos modelos tanto desde el punto de vista determinístico como estocástico. Considerando las secuencias de datos como vectores de un espacio de Hilbert, podemos definir formalmente la construcción de su mejor aproximación lineal.

Es en este momento donde el contexto vuelve novedoso el ajuste: la secuencia de datos no es cronológica sino que es el conjunto de pesos que la transformada de Fourier construye para las diversas frecuencias. El objetivo de este trabajo se centra pues en la modelización lineal de secuencias construidas por la transformada discreta de Fourier a partir de datos reales.

Dos son los aspectos que se desarrollan entorno a esta aproximación. Por un lado, la bondad del ajuste, teniendo como tal no sólo un error relativamente bajo, sino un razonablemente económico número de parámetros en el modelo. Y por otro, la conveniente interpretación del modelo construido.

Respecto al primer aspecto tenemos primero que volver a ese entorno global que presentamos como justificación de la aproximación: los espacios de Hilbert. Nada en su desarrollo limita el uso del producto escalar como medida de similitud: siempre es posible construir la 'mejor aproximación'. Sin embargo, ¿cuándo esta aproximación es 'buena'?, o, ¿cuánto de complejo ha de ser el modelo para que lo sea?

Sin embargo, es en el segundo aspecto donde hemos encontrado el más firme obstáculo. La intuitiva interpretación de los coeficientes que acompañan a los retardos en el dominio del tiempo, no es, de ninguna manera, trasladable al dominio de la frecuencia. Ni siquiera la operación 'retardo' sobre estos pesos, indexados por la frecuencia que representan, ofrece explicación inmediata. ¿Qué sentido tiene el peso de una periodicidad expresado como función lineal de los pesos de las periodicidades de índices menores?

Esta imposibilidad nos hizo variar el punto de vista con el que contemplábamos la aproximación y estudiarla desde dos aspectos bien distintos: por un lado, la interpretación que desde la geofísica se hace del algoritmo que la resuelve, el de Levinson (1947), y, por otro, los aproximantes de Padé-Laurent.

En las prospecciones petrolíferas, la localización de capas de gas natural o petróleo para su extracción, se realiza a través del estudio de las ondas sísmicas generadas a partir de una explosión controlada. Los ecos en la superficie de estas ondas sonoras son modelizadas a través de una estructura autorregresiva. En su construcción, lo realmente importante son los llamados 'coeficientes de reflexión' de las ondas respecto de cada capa, que indican su densidad relativa. Este inteligente uso del algoritmo de Levinson nos guió hacia una novedosa interpretación de la aproximación lineal del dominio de la frecuencia. En definitiva, no se considerarán en la interpretación los coeficientes lineales en los retardos, sino los de reflexión. Podemos destacar además, que la interpretación dada no sólo tiene importancia desde el punto de vista de las aplicaciones, sino que nos ha permitido proponer un método general de aceleración numérica del algoritmo de Levinson.

Los aproximantes de Padé-Laurent, presentes en la aproximación lineal, son el punto de partida del segundo modelo presentado en la memoria. Se trata de una modelización racional de

las secuencias de transformadas discretas de Fourier, pero no a través de filtros lineales, sino directamente usando los aproximantes mencionados.

Contamos para la aplicación real de los dos modelos que se construyen con datos reales de consumo urbano de agua en la ciudad de Santa Cruz de Tenerife. Estos datos, al ser periódicos, permiten un mejor aprovechamiento de la modelización en el dominio de la frecuencia, puesto que los modelos así construidos también lo son.

La presente memoria la hemos estructurado en cuatro capítulos. En el primero tratamos los espacios vectoriales de Hilbert y sus propiedades más relevantes, como entorno teórico unificado para la aproximación lineal. Estos resultados se aplican a varios espacios anotando las principales conclusiones.

El segundo capítulo desarrolla la aproximación lineal por retardos resuelta mediante filtros lineales, listando las condiciones para que ésta sea posible. Este desarrollo es luego aplicado a la secuencia de valores de una transformada discreta de Fourier, llegando a su formulación particular. El capítulo finaliza con la interpretación del modelo, y su aplicación a datos simulados.

El tercer capítulo introduce los aproximantes de Padé-Laurent y aquellos construidos por el algoritmo de Levinson. Su construcción sobre la secuencia de valores de una discreta de Fourier, construye un segundo modelo cerrado para el dominio de la frecuencia.

El último capítulo presenta los resultados de la aplicación de los modelos de los capítulos anteriores. Los datos son clasificados y modelizados tras la elección del número de parámetros necesario. El capítulo finaliza con una comparativa entre los errores de los modelos presentados y la transformada de Fourier truncada.

La memoria finaliza resumiendo conclusiones y aportando algunas cuestiones abiertas.

Capítulo 1 *La aproximación lineal en la modelización de datos*

Introducimos en este capítulo la clásica estructura de los espacios vectoriales de Hilbert y el teorema de la proyección, como entorno formal para la construcción de aproximaciones lineales de vectores.

Si bien, y de forma general, presentamos las definiciones propias de los espacios vectoriales de Hilbert y algunas de sus propiedades básicas, el objetivo es aplicar sobre convenientes espacios de Hilbert el teorema de la proyección. Se establecerán, primero de forma muy genérica, las ecuaciones de la aproximación lineal, y después, particularizando para cada espacio vectorial y cada conjunto de vectores generadores de la aproximación, el sistema de ecuaciones que la construye.

Pretendemos de esta forma establecer un marco uniforme para el desarrollo en esta memoria de tópicos a primera vista dispares, como son, de un lado, la construcción del dominio de la frecuencia o las series de Fourier, y de otro, la aproximación por retardos. La unificación de ambos bajo un mismo punto de vista, facilitará la definición de los modelos que planteamos en los capítulos 2 y 3.

Se inicia el capítulo con una sección introductoria que recuerda la definición y las herramientas básicas de los espacios de Hilbert, principalmente el producto escalar y el concepto de ortogonalidad. Tras enunciar el teorema de la proyección, se traslada a la siguiente sección su aplicación para la construcción de la aproximación lineal y la elaboración de las ecuaciones que la obtienen. Dichos sistemas de ecuaciones se aplican a los espacios vectoriales de Lebesgue L^2 , al de secuencias complejas acotadas l^2 , al N -dimensional complejo \mathbb{C}^N y, por último, a los procesos estacionarios, en las secciones posteriores. Acaba el capítulo con un resumen que, de forma esquemática, relaciona los espacios vectoriales de Hilbert sobre los que se ha trabajado a través de las aproximaciones construidas y de las metodologías aplicadas.

La bibliografía sobre espacios vectoriales de Hilbert es lógicamente abundante, incluso formando parte introductoria del análisis de Fourier desde donde es mencionado por sus propiedades geométricas. Para este último caso, se prefiere definir en el cuerpo de los complejos, \mathbb{C} , toda operación de escalares. Tomaremos también esta dirección en la presente memoria.

Puede ampliarse cualquier concepto de los resumidos aquí en [BroDav91], [Pri81], [Blo76] y [Wei90] si se les quiere relacionados con el estudio de series temporales. En un ámbito más general, debemos mencionar [RobMul87], [OppSch75] y [OppWilYou83].

1.1. El entorno teórico: espacios vectoriales de Hilbert

La característica inherente a un espacio de Hilbert es el producto escalar entre un par de vectores. Como veremos en el desarrollo posterior, usaremos este producto como *medida de similitud o aproximación* entre ellos.

Dado un espacio vectorial complejo H decimos que la aplicación $\langle \cdot, \cdot \rangle: H \times H \rightarrow \mathbb{C}$ es un producto escalar si se verifican las siguientes propiedades:

- (i) $\langle x, y \rangle = \overline{\langle y, x \rangle}, \quad \forall x, y \in H.$
- (ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \forall x, y, z \in H.$
- (iii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle, \quad \forall x, y \in H, \quad \forall \alpha \in \mathbb{C}.$
- (iv) $\langle x, x \rangle \geq 0, \quad \forall x \in H.$
- (v) $\langle x, x \rangle = 0, \quad \text{sii } x = 0.$

Un espacio vectorial complejo H , dotado de producto escalar, se dice que es un espacio de Hilbert si es completo, esto es, si toda secuencia de Cauchy en H converge en H .

En un espacio de Hilbert H , se define una norma de forma natural. Sea esta norma inducida la definida como,

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad \forall x \in H.$$

También en un espacio de Hilbert H , se redefine el producto escalar de la forma siguiente,

$$\langle x, y \rangle = \|x\| \|y\| \cos \angle xy, \quad \forall x, y \in H,$$

donde $\angle xy$ representa el ángulo formado por los vectores x e y . Esta expresión del producto escalar introduce por tanto el concepto de ángulo entre vectores, justificando el punto de vista geométrico que adorna la teoría de espacios de Hilbert.

Dado que dos vectores se dicen ‘ortogonales’ si el ángulo que forman es $\pm\pi/2$, llamaremos ortogonales a aquellos vectores cuyo producto escalar es cero, denotando este hecho con el símbolo \perp . Es decir, $x \perp y$ si $\langle x, y \rangle = 0$.

Dado un subconjunto $S \subseteq H$, se define su complemento ortogonal como el subespacio vectorial cerrado

$$S^\perp = \{y \in H / y \perp x \quad \forall x \in S\},$$

es decir, aquel conjunto de vectores, ortogonales a todos los elementos de S .

Con las anteriores definiciones ya puede ser formulado el siguiente resultado, el **teorema de la proyección**, en el que nos basaremos para definir la aproximación lineal. Se enuncia de la siguiente manera:

Dado un subespacio vectorial S del espacio de Hilbert H , y un vector $x \in H$, entonces,

- (i) existe un único elemento $\hat{x} \in S$ tal que $\|x - \hat{x}\| = \inf_{y \in S} \|x - y\|$.
- (ii) son equivalentes,
 - $\hat{x} \in S$ y $\|x - \hat{x}\| = \inf_{y \in S} \|x - y\|$
 - $\hat{x} \in S$ y $x - \hat{x} \in S^\perp$.

Si tomamos el sentido euclídeo de la norma como ‘distancia’, el teorema de la proyección afirma la existencia y unicidad del vector más ‘próximo’,

$$\exists \hat{x} \in S \text{ tal que } \|x - \hat{x}\| = \inf_{y \in S} \|x - y\|,$$

y resuelve la forma de hallarlo con una caracterización, como veremos, constructiva, $x - \hat{x} \in S^\perp$.

De esta última caracterización podemos deducir la siguiente descomposición $x = \hat{x} + x - \hat{x}$, donde el primer sumando \hat{x} es el más ‘próximo’ a x en S , y le denominaremos ‘aproximación en S ’, y el segundo, $x - \hat{x}$, está en el subespacio ortogonal S^\perp y nos referiremos a él como ‘error en la aproximación’.

Puesto que $x - \hat{x} \perp \hat{x}$, se comprueba fácilmente que,

$$\|x\|^2 = \|\hat{x}\|^2 + \|x - \hat{x}\|^2,$$

y en ese caso la aproximación \hat{x} minimiza el segundo sumando, esto es, la norma del error en la aproximación.

La descomposición indicada es única, pues si hubiera un par de vectores $y \in S$ y $z \in S^\perp$, tales que $x = y + z$, entonces $x - y = z \in S^\perp$, con lo que $y = \hat{x}$ por la unicidad planteada en el teorema de la proyección, y, finalmente, $z = x - \hat{x}$.

El teorema de la proyección permite definir la aplicación proyección $\Pi_S : H \rightarrow S$, de forma que $\Pi_S x = \hat{x}$, siendo $\|x - \hat{x}\| = \inf_{y \in S} \|x - y\|$, lo que nos facilitará la notación a partir de ahora.

1.2. En busca de la 'mejor' aproximación

El teorema de la proyección resuelve, como se ha indicado, la construcción de la aproximación en S , Π_S , mediante la caracterización $x - \Pi_S x \in S^\perp$. Simplemente aplicando la definición vista de subespacios ortogonales, se extrae el siguiente sistema de ecuaciones,

$$\langle x - \Pi_S x, y \rangle = 0, \quad \forall y \in S. \text{ Ecuación 1.1}$$

Nótese que la solución a este sistema existe y es única.

Aún así la metodología para su resolución permanece vaga debido al uso del cualificador $\forall y \in S$. Es precisamente bajo diversas elecciones del conjunto S cuando la resolución del sistema se hace posible a través de caracterizaciones de S^\perp .

Supongamos un subconjunto de vectores linealmente independientes $\{x^j / j \in J\}$ de H . Construimos a partir de ellos el menor subespacio vectorial cerrado que los contenga que denotaremos por $\overline{\text{gen}\{x^j / j \in J\}}$. Si el conjunto de índices J es finito, $\overline{\text{gen}\{x^1, x^2, \dots, x^p\}}$ está constituido por todas las posibles combinaciones lineales de los vectores $\{x^1, x^2, \dots, x^p\}$. Esto es

$$\overline{\text{gen}\{x^1, x^2, \dots, x^p\}} = \left\{ \sum_{j=1}^p \alpha^j x^j / \alpha^j \in \mathbb{C} \quad \forall j = 1, \dots, p \right\}.$$

En ese caso se verifica que,

$$y \in \overline{\text{gen}\{x^1, x^2, \dots, x^p\}}^\perp \text{ sii } \langle y, x^j \rangle = 0 \quad \forall j = 1, \dots, p,$$

por la forma en la que se expresan los vectores de $\overline{\text{gen}}\{x^1, x^2, \dots, x^p\}$ y la linealidad del producto escalar. Nótese que si buscamos la mejor aproximación, en este caso lineal, en $\overline{\text{gen}}\{x^1, x^2, \dots, x^p\}$, el sistema de descrito en *Ecuación 1.1* se transforma en otro de p ecuaciones,

$$\langle \Pi_{\overline{\text{gen}}\{x^1, x^2, \dots, x^p\}} x, x^j \rangle = \langle x, x^j \rangle, \quad \forall j = 1, \dots, p,$$

que podemos resumir diciendo que x y su aproximación tienen idénticos productos escalares con los vectores $\{x^1, x^2, \dots, x^p\}$. De ahí que habláramos del producto escalar como medida de similitud. La igualdad se puede desarrollar aún más, dado que

$$\Pi_{\overline{\text{gen}}\{x^1, x^2, \dots, x^p\}} x = \sum_{k=1}^p \alpha^k x^k,$$

para ciertos coeficientes complejos $\{\alpha^1, \alpha^2, \dots, \alpha^p\}$, que son buscados por las ecuaciones que verifiquen,

$$\langle \sum_{k=1}^p \alpha^k x^k, x^j \rangle = \sum_{k=1}^p \alpha^k \langle x^k, x^j \rangle = \langle x, x^j \rangle \quad \forall j = 1, 2, \dots, p. \quad \text{Ecuación 1.2.}$$

Éstas sí que constituyen un verdadero sistema de ecuaciones cuyas incógnitas son los coeficientes $\{\alpha^1, \alpha^2, \dots, \alpha^p\}$.

Si además el conjunto de vectores generadores $\{x^1, x^2, \dots, x^p\}$ es un sistema ortogonal, esto es, $\langle x^j, x^k \rangle = 0$, para $i, k = 1, \dots, p$, $i \neq k$, entonces las p ecuaciones 1.2 se resuelven por sí mismas quedando

$$\alpha^j = \frac{\langle x, x^j \rangle}{\|x^j\|^2}, \quad \text{para } j = 1, \dots, p.$$

Nótese que si el conjunto de vectores generadores es además ortonormal, esto es, $\langle x^j, x^k \rangle = \delta_{jk}$, (*delta de Kronecker*), la condición de similitud en la aproximación lineal se verifica para $\alpha^j = \langle x, x^j \rangle$ con $j = 1, \dots, p$. En ese caso,

$$\Pi_{\text{gen}\{x^1, x^2, \dots, x^p\}} x = \sum_{k=1}^p \alpha^k x^k. \text{ Ecuación 1.3.}$$

A los valores $\alpha^j = \langle x, x^j \rangle$ se les denominan *coeficientes de Fourier* de x respecto al conjunto ortonormal $\{x^1, x^2, \dots, x^p\}$.

Por último, si el espacio de Hilbert H tiene una base numerable, $H = \overline{\text{gen}\{x^j / j = 1, 2, 3, \dots\}}$ entonces

$$(i) \quad \forall x \in H, \quad x = \sum_j \langle x, x^j \rangle x^j$$

(Nótese que $\Pi_H x = x$ puesto que $H^\perp = \{0\}$)

$$(ii) \quad \forall x \in H, \quad \left\| x - \sum_{j=1}^n \langle x, x^j \rangle x^j \right\| \rightarrow 0, \text{ cuando } n \rightarrow \infty$$

(Convergencia de la aproximación lineal)

$$(iii) \quad \langle x, y \rangle = \sum_j \langle x, x^j \rangle \langle x^j, y \rangle, \quad \forall x, y \in H$$

(Identidad de Parseval)

Como hemos indicado, en las secciones siguientes aplicaremos sobre distintos espacios de Hilbert y convenientes conjuntos de vectores generadores los resultados expuestos.

1.3. Un primer ejemplo: aproximación en Espacios de Lebesgue L^2

Consideremos el espacio medible (Ω, F) , y sea μ una medida finita no nula¹. Denotamos por $L^2(\Omega, F, \mu)$ al conjunto de funciones complejas definidas sobre Ω a través de la integral de Lebesgue,

$$L^2(\Omega, F, \mu) = \{f / \int_{\Omega} |f|^2 d\mu < \infty\}.$$

El conjunto así definido es un espacio Hilbert con el producto escalar

$$\langle f, g \rangle = \int_{\Omega} f \bar{g} d\mu, \text{ con } f, g \in L^2(\Omega, F, \mu).$$

Nótese que por las propiedades de la integral, la igualdad entre funciones en $L^2(\Omega, F, \mu)$ no es la clásica igualdad para todo elemento en Ω . Esto se debe a que,

$$\langle f, f \rangle = \int_{\Omega} |f|^2 d\mu = 0,$$

no implica que $f = 0$, (*la función nula*).

En este caso si $\int_{\Omega} |f|^2 d\mu = 0$, entonces f es la función cero excepto en un conjunto de medida nula. Por lo tanto, podemos considerar que $L^2(\Omega, F, \mu)$ es el espacio cociente resultante de aplicar a las funciones la siguiente relación de equivalencia: $f \approx g$ sii $f = g$ *excepto en un conjunto de medida nula*.

¹ A DIFERENCIA DE LAS FUNCIONES DE PROBABILIDAD, UNA FUNCIÓN DE MEDIDA NO ESTÁ ACOTADA POR 1.

Un caso particular interesante es en el que $\Omega = [-\pi, \pi]$, F son los subconjuntos de Borel de $[-\pi, \pi]$ y μ es la medida uniforme sobre el intervalo, esto es, $d\mu = \frac{dx}{2\pi}$. El conjunto de funciones que denotaremos por $L^2[-\pi, \pi] = \{f / \int_{-\pi}^{\pi} |f(x)|^2 dx < \infty\}$ tiene como producto escalar entonces,

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{g(x)} dx.$$

Sean las funciones en $L^2[-\pi, \pi]$, $\{e_n(x) = e^{inx} / n \in \mathbb{Z}\}$. Este conjunto constituye un sistema ortonormal puesto que

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x} dx = \delta_{nm}.$$

Según los resultados expuestos, si aplicamos *Ecuación 1.3* para los coeficientes en la aproximación lineal sobre $\overline{\text{gen}\{e_n / |n| \leq N\}}$, dado $f \in L^2[-\pi, \pi]$, tenemos

$$\Pi_{\overline{\text{gen}\{e_n / |n| \leq N\}}} f = \sum_{k=-N}^N \langle f, e_k \rangle e_k,$$

donde los coeficientes de Fourier se obtienen como

$$\langle f, e_k \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx.$$

Esta aproximación tiene importantísimas propiedades asintóticas, esto es, cuando $N \rightarrow \infty$ en $\overline{\text{gen}\{e_n / |n| \leq N\}}$. Éstas son:

$$(i) \quad \left\| f - \sum_{k=-N}^N \langle f, e_k \rangle e_k \right\|^2 \rightarrow 0 \text{ cuando } N \rightarrow \infty.$$

$$(ii) \quad \Pi_{\overline{\text{gen}\{e_n/n \in \mathbb{Z}\}}} f = f$$

Por lo tanto $L^2[-\pi, \pi] = \overline{\text{gen}\{e_n/n \in \mathbb{Z}\}}$.

La igualdad $f(x) = \sum_k \langle f, e_k \rangle e_k(x)$ se establece en los términos mencionados previamente para espacios cocientes L^2 , esto es, igualdad punto a punto excepto en un conjunto de medida nula. La serie biinfinita $\sum_k \langle f, e_k \rangle e_k(x)$ se denomina serie de Fourier para $f(x) \in L^2[-\pi, \pi]$.

En general, si consideramos un intervalo cerrado arbitrario de longitud $T > 0$, cualquier función compleja f definida sobre él que verifique

$$\int_T |f(t)|^2 dt < \infty$$

puede ser aproximada en el intervalo mediante exponenciales. La idea resulta de aplicar un cambio de variable lineal entre los intervalos $[-\pi, \pi]$ y el de longitud T . Al mismo resultado se llega si, partiendo del esquema general de los espacios de Lebesgue L^2 , consideramos la medida

$$d\mu = \frac{1}{T} dt$$

y el sistema ortonormal

$$\{e^{i\frac{2\pi}{T}nt} / n \in \mathbb{Z}\}.$$

La serie de Fourier resultante sería de la forma $\sum_n \langle f, e_n \rangle e_n(t)$ donde

$$\langle f, e_n \rangle = \frac{1}{T} \int_T f(t) e^{-i \frac{2\pi}{T} nt} dt$$

siendo $f(t)$ la función mencionada.

La aproximación sobre intervalos sería finalista para un determinado tipo de funciones: las periódicas. Decimos que una función $p(t)$ definida para $t \in \mathbb{R}$ es periódica si existe una constante real $T > 0$ tal que $p(t) = p(t+T) \forall t \in \mathbb{R}$.

Al valor T_0 más pequeño con el que $p(t)$ verifica esta propiedad se denomina *periodo*. Si $p(t)$ es una función periódica de periodo T_0 , y representamos con el operador \int_{T_0} la integración sobre cualquier intervalo de longitud T_0 , la propiedad,

$$\int_{T_0} |p(t)|^2 dt < \infty$$

implica que existe y es convergente la serie de Fourier que la aproxima en dicho intervalo.

Nótese que las funciones usadas en la aproximación,

$$\{e^{i \frac{2\pi}{T_0} nt} / n \in \mathbb{Z}\},$$

son periódicas de periodo T_0 por lo que la serie de Fourier resultante de la aproximación también lo es al ser combinación lineal de las primeras. En ese caso,

$$p(t) = \sum_n \langle p, e_n \rangle e_n(t)$$

no sólo en el intervalo de aproximación sino $\forall t \in \mathbb{R}$.

En funciones periódicas, junto al valor del periodo T_0 , se define la frecuencia w_0 tal que $T_0 w_0 = 2\pi$. De ese modo las exponenciales usadas en la aproximación se reescriben como $\{e^{inw_0 t} / n \in \mathbb{Z}\}$ y se dice de ellas que tienen frecuencia nw_0 respectivamente.

Cada coeficiente de Fourier indica el peso que la correspondiente exponencial tiene en el sumatorio final. O lo que es lo mismo, el coeficiente $\langle p, e_n \rangle$ indica el peso de la exponencial $e^{\frac{i2\pi}{T_0} nt}$ en el sumatorio conjunto. De este modo, un valor casi nulo para $\langle p, e_n \rangle$ indicaría que en p no se registran comportamientos de periodo T_0/n (o frecuencia nw_0). Un alto valor para este coeficiente indicaría lo contrario.

Este análisis de las funciones periódicas de periodo T_0 a través de los distintos pesos de las exponenciales de periodicidad T_0/n aporta un punto de vista diferente que explica claramente comportamientos estacionales para frecuencias diversas. Los valores discretos de la secuencia $\{\langle f, e_n \rangle\}$ constituyen el denominado espectro de Fourier, o dominio de la frecuencia.

La aproximación lograda, sin embargo, se circunscribe en general a un intervalo de longitud arbitraria pero finita. Este hecho limitaría una aproximación efectiva global, excepto en el caso de funciones periódicas.

1.4. Aproximando secuencias finitas: la transformada discreta de Fourier

En el mismo entorno general de los espacios de Hilbert, queremos definir el dominio de la frecuencia desde el punto de vista de las secuencias de datos. Para ellos partiremos de una secuencia finita de valores complejos y usaremos para la aproximación en el mismo espacio vectorial secuencias generadas por exponenciales.

Sea $\{s_0, s_1, \dots, s_{N-1}\}$ una secuencia compleja finita de N términos. Podemos considerarla un vector del espacio vectorial complejo \mathbb{C}^N , que es de Hilbert con el producto escalar,

$$\forall \vec{x}, \vec{y} \in \mathbb{C}^N, \quad \langle \vec{x}, \vec{y} \rangle = \sum_{k=0}^{N-1} x_k \overline{y_k}, \quad \text{con } \vec{x} = (x_0, x_1, \dots, x_{N-1}) \text{ e } \vec{y} = (y_0, y_1, \dots, y_{N-1}).$$

En \mathbb{C}^N consideramos el siguiente conjunto de vectores, $\{\vec{W}_N^r / r = 0, 1, \dots, N-1\}$, donde se define el vector $\vec{W}_N^r = \frac{1}{\sqrt{N}} (1, W_N^r, W_N^{2r}, \dots, W_N^{(N-1)r})$, para valores $r = 0, 1, \dots, N-1$, y siendo $W_N^r = e^{i \frac{2\pi r}{N}}$.

Este sistema es ortonormal en \mathbb{C}^N , puesto que,

$$\langle \vec{W}_N^r, \vec{W}_N^s \rangle = \frac{1}{N} \sum_{k=0}^{N-1} W_N^{rk} \overline{W_N^{sk}} = \frac{1}{N} \sum_{k=0}^{N-1} W_N^{(r-s)k},$$

con lo que, si $r=s$, tenemos $\frac{1}{N} \sum_{k=0}^{N-1} 1 = 1$; en cambio, si $r \neq s$, la suma de

la serie geométrica $\frac{1}{N} \sum_{k=0}^{N-1} W_N^{(r-s)k} = \frac{1}{N} \frac{W_N^{(r-s)N} - 1}{W_N^{r-s} - 1} = 0$.

Los valores que construyen los vectores W_N^r son raíces N -ésimas de la unidad, por lo que están situadas sobre el círculo $|z|=1$, siendo su principal propiedad que son periódicas de periodo N . Más aún, si tenemos un divisor r de N la secuencia finita que construye W_N^r tiene como periodo N/r . Como el índice r recorre todos los valores posibles entre 0 y $N-1$, toma en alguna ocasión los valores de sus divisores.

Debido a que la dimensión de \mathbb{C}^N es N , el conjunto de vectores que hemos definido, $\{W_N^r / r=0,1,\dots,N-1\}$ constituye una base del espacio vectorial y por tanto $\mathbb{C}^N = \overline{\text{gen}\{W_N^r / r=0,1,\dots,N-1\}}$.

Aplicando las *Ecuación 1.3* tenemos que,

$$\forall \vec{x} \in \mathbb{C}^N, \quad \vec{x} = \sum_{r=0}^{N-1} \langle \vec{x}, W_N^r \rangle W_N^r,$$

siendo los coeficientes $\langle \vec{x}, W_N^r \rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k W_N^{-rk}$. Por las propiedades periódicas de los vectores de esta base construida a partir de exponenciales, este coeficiente que acompaña a W_N^r , para $r=0,1,\dots,N-1$, indica el peso o influencia que, en \vec{x} , tienen componentes de periodicidad N/r .

Estas ecuaciones, deducidas a través del teorema de proyección pueden escribirse también como las de cambio de base entre la cartesiana, según la cual $\vec{x} = (x_0, x_1, \dots, x_{N-1})$, y la que se ha definido mediante exponenciales que denotaremos por $\vec{X} = (X_0, X_1, \dots, X_{N-1})$. Las

fórmulas de conversión quedan referidas a las coordenadas de los vectores de la siguiente manera,

$$X_r = \langle \bar{x}, W_N^r \rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k W_N^{-rk}, \quad \text{Ecuación DFT}$$

$$x_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} X_r W_N^{rk}, \quad \text{Ecuación IDFT.}$$

Ambas expresiones, inversas la una de la otra, se denominan transformada discreta de Fourier, *DFT*, y transformada inversa de Fourier, *IDFT*, respectivamente.

Aún cuando en su construcción hemos preferido utilizar el entorno de los espacio de Hilbert y el teorema de la proyección es la introducción de estas relaciones discretas a través de acciones de muestreo sobre la transformada de Fourier, el desarrollo más usual en la bibliografía sobre el tema².

Desde el punto de vista computacional, las relaciones descritas han sido resueltas eficientemente mediante el algoritmo que describimos a continuación.

La transformada rápida de Fourier, FFT.

Este algoritmo calcula eficientemente la suma compleja:

$$X(r) = \sum_{k=0}^{N-1} A(k) W_N^{rk}.$$

² VÉASE POR EJEMPLO [ROBMUL87] Y [LAT92].

Esta suma, con adecuados valores para $A(n)$ y tratando $X(r)$, puede ser usada no sólo para el cálculo de la transformada discreta de Fourier y su inversa, sino de funciones de autocorrelación y convoluciones.

La transformada rápida de Fourier, *FFT* en adelante, fue propuesta por los investigadores James W. Cooley y John Tuckey en 1965 en un artículo cuyo título podemos traducir como “*Un algoritmo para el procesamiento de series complejas de Fourier*” [CooTuc65]. A este trabajo siguieron otros como [CocCoo67], [CooLewWel69] y [CooLewWel70] en los que se desarrolló la idea principal del método a otras aplicaciones. El algoritmo de la *FFT* supuso un gran avance por la reducción de tiempo de cálculo que supone y en la actualidad sigue siendo objeto de estudio.

La *FFT* se basa en la factorización de N . Supongamos, por ello, que es posible descomponerlo como $N = n_1 n_2$, con $n_1 \leq n_2$. En ese caso, los índices r y k cuyo producto es la potencia de la raíz N -ésima de la unidad se pueden expresar como

$$r = n_1 r_2 + r_1$$

$$k = k_1 n_2 + k_2,$$

donde $0 \leq r_j, k_j \leq n_j - 1$, y $j = 1, 2$. En ese sentido, diremos que r tiene una *codificación* \rightarrow de la forma $(r_2 r_1)$ y k tiene una *codificación* \leftarrow de la forma $(k_1 k_2)$. Veamos como en el siguiente ejemplo para $N = 2 \cdot 3 = 6$, serían las codificaciones en ambos sentidos de los números 0 al 5:

número	codificación \leftarrow	codificación \rightarrow
0	00	00
1	01	01
2	02	10

3	10	11
4	11	20
5	12	21

Tabla 1.1. Codificación por factorización

Codificados así los índices, y dado que las potencias son, en este caso, *congruentes* módulo N , tenemos que rk es:

$$rk = Nk_1r_2 + n_1r_2k_2 + r_1k_1n_2 + r_1k_2 \equiv k_2(n_1r_2 + r_1) + k_1n_2r_1 \pmod{N}.$$

Sustituyendo los índices por sus codificaciones obtenemos:

$$\begin{aligned} X(r_2r_1) &= \sum_{k_2=0}^{N-1} \sum_{k_1=0}^{N-1} A(k_1k_2) W_N^{k_2(n_1r_2+r_1)+k_1(n_2r_1)} = \\ &= \sum_{k_2=0}^{N-1} W_N^{k_2(n_1r_2+r_1)} \\ &\quad \sum_{k_1=0}^{N-1} A(k_1k_2) W_{n_1}^{k_1r_1} \end{aligned}$$

Si definimos la siguiente subserie,

$$A^1(r_1k_2) = \sum_{k_1=0}^{N-1} W_{n_1}^{k_1r_1} A(k_1k_2),$$

tenemos que $X(r_1r_2) = \sum_{k_2=0}^{N-1} A^1(r_1k_2) W_N^{k_2(n_1r_2+r_1)}$.

Sumando las operaciones, teniendo en cuenta que por ‘operación’ entendemos ‘una suma y un producto’ podemos resolver:

Suma directa

N operaciones por índice = **N^2 operaciones.**

A través de la factorización

Construcción de A^1 + Construcción de $X = n_1N + n_2N = (\mathbf{n}_1 + \mathbf{n}_2)N$.

En ese caso se ha pasado de las $N\mathbf{n}_1\mathbf{n}_2$ operaciones iniciales a $N(\mathbf{n}_1 + \mathbf{n}_2)$. Suponiendo que esta factorización se volviera a efectuar sobre divisores de n_2 y así sucesivamente, pasaríamos de $N\mathbf{n}_1\mathbf{n}_2 \cdots \mathbf{n}_m$ operaciones a $N(\mathbf{n}_1 + \mathbf{n}_2 + \dots + \mathbf{n}_m)$. Esto es, el orden de complejidad se vería rebajado entonces de $O(N^2)$ a $O(N \log N)$.

Si aplicamos el algoritmo mencionado a *DFT* e *IDFT*, según las ecuaciones respectivas,

a. Para la *DFT* basta tomar como $A(k)$ a $\frac{1}{\sqrt{N}} \overline{x_k}$, obteniendo en $X(r)$ a $\overline{X_r}$,

b. Para la *IDFT* tomaríamos como $A(r)$ a $\frac{1}{\sqrt{N}} X_r$ para obtener en $X(k)$ a x_k .

1.5. Aproximación lineal mediante retardos: modelización en l^2

En esta sección cambia el conjunto base de la aproximación puesto que no consideramos funciones o secuencias basadas en exponenciales. Se trata en este caso de obtener la modelización de datos como combinación lineal de información precedente, o más formalmente, aproximar secuencias complejas mediante sus retardos.

El conjunto de sucesiones complejas,

$$l^2 = \left\{ \{x\} / \sum_n |\{x\}_n|^2 < \infty \right\}$$

1. es un espacio vectorial complejo con las operaciones,

$\forall n \in \mathbb{Z}, (\{x\} + \{y\})_n = \{x\}_n + \{y\}_n$, dadas $\{x\}, \{y\} \in l^2$ (la suma de secuencias es la secuencia suma término a término),

$\forall n \in \mathbb{Z}, \forall \lambda \in \mathbb{C}, (\lambda \{x\})_n = \lambda \{x\}_n$, dada $\{x\} \in l^2$ (el producto de un escalar por la secuencia se resuelve multiplicándolo término a término). *Ecuaciones 1.4.*

2. y un espacio vectorial de Hilbert con el producto escalar,

$$\langle \{x\}, \{y\} \rangle = \sum_n \{x\}_n \overline{\{y\}_n}.$$

Sobre un vector $\{x\} \in l^2$, definimos el operador retardo de la siguiente manera,

$$z^k: l^2 \rightarrow l^2$$

$$\{x\} \mapsto (z^k \{x\})_n = \{x\}_{n-k} \quad \text{Ecuación 1.5.}$$

esto es, la sucesión $z^k \{x\}$ es aquella tal que su término n -ésimo es el $(n-k)$ -ésimo de $\{x\}$.

Dada una secuencia $\{x\} \in l^2$, sea $S^p = \overline{\text{gen}\{z\{x\}, z^2\{x\}, \dots, z^p\{x\}\}}$. Obtengamos la mejor aproximación lineal de $\{x\}$ en S^p , esto es, en sus p retardos.

Aplicando lo visto en la sección 1.2,

$$\Pi\{x\} = \sum_{j=1}^p \alpha^j z^j \{x\},$$

de donde se planteaba el sistema *Ecuaciones 1.2*,

$$\sum_{k=1}^p \alpha^k \langle z^k \{x\}, z^j \{x\} \rangle = \langle \{x\}, z^j \{x\} \rangle, \text{ para } j=1, 2, \dots, p.$$

En general,

$$\langle z^k \{x\}, z^j \{x\} \rangle = \langle z^{k-j} \{x\}, \{x\} \rangle,$$

con lo que se hace pertinente la siguiente definición: sea

$$r_j = \langle z^j \{x\}, \{x\} \rangle.$$

Nótese que, por las propiedades del producto escalar, $r_{-j} = \overline{r_j}$, igualdad que va a resultar clave para la construcción de la aproximación. Además al definirse a través de un producto escalar en l^2 se asegura la convergencia de cada término.

En todo caso, su definición permite la reescritura del sistema lineal anterior como,

$$\sum_{k=1}^p \alpha^k r_{k-j} = r_{-j}, \text{ para } j=1, 2, \dots, p,$$

cuya matriz de coeficientes es,

$$T^p = \begin{pmatrix} r_0 & r_1 & r_2 & \dots & r_{p-1} \\ r_{-1} & r_0 & r_0 & \dots & r_{p-2} \\ r_{-2} & r_{-1} & r_0 & \dots & r_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{1-p} & r_{2-p} & r_{3-p} & \dots & r_0 \end{pmatrix} = \begin{pmatrix} r_0 & r_1 & r_2 & \dots & r_{p-1} \\ \overline{r_1} & \overline{r_0} & \overline{r_0} & \dots & \overline{r_{p-2}} \\ \overline{r_2} & \overline{r_1} & \overline{r_0} & \dots & \overline{r_{p-3}} \\ \dots & \dots & \dots & \dots & \dots \\ \overline{r_{p-1}} & \overline{r_{p-2}} & \overline{r_{p-3}} & \dots & r_0 \end{pmatrix}$$

una matriz con estructura *Toeplitz*. Volveremos, por su importancia, en la sección siguiente sobre este tipo de sistemas, abordando brevemente algunos aspectos computacionales.

Finalmente anotamos que dado un vector $f \in L^2[-\pi, \pi]$, el conjunto de coeficientes de Fourier $\langle f, e_k \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx$ constituye una secuencia compleja acotada según

$$\|f\|^2 = \sum_k |\langle f, e_k \rangle|^2 < \infty, \text{ (identidad de Parseval).}$$

Por lo tanto $\{\langle f, e_k \rangle\}$ pertenece al espacio de Hilbert l^2 . Aún más, los dos espacios de Hilbert mencionados son isomorfos mediante la aplicación:

$$\begin{aligned} L^2[-\pi, \pi] &\rightarrow l^2 \\ f &\mapsto \{\langle f, e_k \rangle\} \end{aligned}$$

1.6. El caso de los procesos estacionarios

Desarrollaremos en esta sección una aplicación conocida del teorema de proyección sobre espacios de Hilbert a la aproximación lineal por retardos de procesos estacionarios.

Sea (Ω, F, P) un espacio de probabilidad y $L^2(\Omega, F, P)$, el conjunto de variables aleatorias complejas, ξ , verificando,

$$\int_{\Omega} |\xi(w)|^2 P(dw) < \infty.$$

Como ya se ha comentado en la sección 1.3, $L^2(\Omega, F, P)$ es un espacio vectorial de Hilbert con el producto escalar,

$$\langle \xi, \eta \rangle = E(\xi \bar{\eta}) = \int_{\Omega} \xi(w) \overline{\eta(w)} P(dw).$$

Consideremos el proceso en $L^2(\Omega, F, P)$, $\{\xi_t / t \in \mathbb{Z}\}$. Diremos que el proceso es *completamente estacionario* o *estacionario estricto* si, para cualquiera n instantes, t_1, t_2, \dots, t_n , y cualquier k , la distribución de probabilidad conjunta de $\{\xi_{t_1}, \xi_{t_2}, \dots, \xi_{t_n}\}$ es idéntica a la de $\{\xi_{t_1+k}, \xi_{t_2+k}, \dots, \xi_{t_n+k}\}$. Podemos resumir la anterior definición diciendo que *la estructura probabilística de un proceso completamente estacionario es invariante ante desplazamientos temporales* [Pri92].

Relajando esta definición, y haciendo intervenir sólo los momentos de orden diverso, se habla también de estacionariedad hasta cierto orden.

Un proceso $\{\xi_t / t \in \mathbb{Z}\}$ se dice *estacionario hasta orden m* si para cualesquiera n instantes t_1, t_2, \dots, t_n , y cualquier entero k , los momentos conjuntos de hasta orden m de $\{\xi_{t_1}, \xi_{t_2}, \dots, \xi_{t_n}\}$ existen y coinciden con los de $\{\xi_{t_1+k}, \xi_{t_2+k}, \dots, \xi_{t_n+k}\}$.

En ese caso y para los primeros valores de $m=1, 2$, la definición implica los siguientes resultados.

- Si $\{\xi_t / t \in \mathbb{Z}\}$ es estacionario hasta orden 1, entonces $E(\xi_t) = E(\xi_{t+k})$, $\forall k$, con lo que $E(\xi_t) = \mu$ (media del proceso), independiente del índice t .
- Si $\{\xi_t / t \in \mathbb{Z}\}$ es estacionario de hasta orden 2 verificará el punto anterior, media constante e independiente del índice, y además,

$E(|\xi_t|^2) = E(|\xi_{t+k}|^2)$, $\forall k$, con lo que la varianza σ^2 es constante e independiente del índice temporal.

$E(\xi_t \overline{\xi_s}) = E(\xi_{t+k} \overline{\xi_{s+k}}) = E(\xi_0 \overline{\xi_{s-t}})$, con lo que la autocovarianza sólo depende del incremento $t-s$.

A esta última definición para $m=2$, se le denomina *estacionariedad de hasta orden 2, débil, en sentido amplio o de covarianza*, siendo la que más comúnmente se toma como definición de estacionariedad por evidentes razones prácticas.

A la autocovarianza $cov(\xi_t \overline{\xi_s})$, por la propiedad mencionada para los procesos estacionarios, se denota por γ_l de forma que $\gamma_l = E(\xi_0 \overline{\xi_l}) - \mu^2$.

Dado un proceso estacionario $\{\xi_t / t \in \mathbb{Z}\}$, pretendemos aplicar el teorema de la proyección para obtener la mejor aproximación lineal sobre el siguiente conjunto de procesos $\overline{gen}\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}$, esto es, aquellos que se construyen por retardos de hasta orden p sobre el proceso estacionario que queremos aproximar. En ese caso, tenemos que

$$\Pi_{gen\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}} \xi_t = \sum_{k=1}^p \alpha^k \xi_{t-k},$$

es la aproximación que minimiza el error cuadrático medio,

$$\left\| \xi - \Pi_{gen\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}} \xi \right\|^2 = \int_{\Omega} \left| \xi - \Pi_{gen\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}} \xi \right|^2 P(dw) = E \left| \xi - \Pi_{gen\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}} \xi \right|^2.$$

Las ecuaciones de predicción en este caso son,

$$\sum_{k=1}^p \alpha^k \langle \xi_{t-k}, \xi_{t-j} \rangle = \langle \xi_t, \xi_{t-j} \rangle, \text{ para } j=1, 2, \dots, p.$$

Sustituyendo la expresión del producto escalar tenemos, usando el operador *esperanza*,

$$\sum_{k=1}^p \alpha^k E(\xi_{t-k} \overline{\xi_{t-j}}) = E(\xi_t \overline{\xi_{t-j}}), \text{ para } j=1, 2, \dots, p.$$

Supongamos que, sin pérdida de generalidad, el proceso estacionario que modelizamos tiene media cero. Esta suposición se basa en que, si definimos el proceso auxiliar también estacionario, $\xi_t^0 = \xi_t - E(\xi_t)$, éste tiene igual función de autocovarianza y media cero, y además,

$$\Pi_{gen\{\xi_{t-1}, \xi_{t-2}, \dots, \xi_{t-p}\}} \xi = E(\xi) + \Pi_{gen\{\xi_{t-1}^0, \xi_{t-2}^0, \dots, \xi_{t-p}^0\}} \xi^0, \text{ [BroDav91]}.$$

En ese caso, la función de autocovarianza coincide con el autoproducto escalar $\gamma_l = \langle \xi_0, \xi_l \rangle = E(\xi_0 \overline{\xi_l})$ y permite redefinir el sistema *Ecuación 1.2* como,

$$\sum_{k=1}^p \alpha^k \gamma_{k-j} = \gamma_{-j}, \text{ para } j=1, 2, \dots, p.$$

para calcular la mejor aproximación lineal por retardos del proceso.

La matriz de coeficientes tiene estructura Toeplitz al igual que ocurre con la aproximación en l^2 vista en la sección anterior,

$$T^p = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_{-1} & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_{-2} & \gamma_{-1} & \gamma_0 & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{1-p} & \gamma_{2-p} & \gamma_{3-p} & \cdots & \gamma_0 \end{pmatrix} = \begin{pmatrix} \overline{\gamma_0} & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \overline{\gamma_1} & \overline{\gamma_0} & \gamma_1 & \cdots & \gamma_{p-2} \\ \overline{\gamma_2} & \overline{\gamma_1} & \overline{\gamma_0} & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \overline{\gamma_{p-1}} & \overline{\gamma_{p-2}} & \overline{\gamma_{p-3}} & \cdots & \gamma_0 \end{pmatrix}.$$

Supongamos ahora una realización de tamaño N del proceso estacionario de media cero $\{\xi_t / t \in \mathbb{Z}\}$ que denotaremos por $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$. Si usamos el estimador sesgado de Parzen [1961], para la autocovarianza podemos estimarla como

$$\hat{\gamma}_l = \frac{1}{N} \sum_{n=1}^{N-|l|} \varepsilon_n \varepsilon_{n+l},$$

empleando dicho valor en la construcción de la aproximación.

En ese caso las ecuaciones para el cálculo de los coeficientes de la aproximación serían

$$\frac{1}{N} \sum_{k=1}^p \alpha^k \sum_{n=1}^{N-|k-j|} \varepsilon_n \varepsilon_{n+k-j} = \frac{1}{N} \sum_{n=1}^{N-j} \varepsilon_n \varepsilon_{n-j}, \text{ para } j=1, 2, \dots, p,$$

esto es,

$$\sum_{k=1}^p \alpha^k \hat{\gamma}_{k-j} = \hat{\gamma}_{-j}, \text{ para } j=1, 2, \dots, p,$$

de donde se obtiene la aproximación sobre l^2 de la secuencia,

$$\{\varepsilon_n\} = \begin{cases} \varepsilon_n & \text{si } 1 \leq n \leq N \\ 0 & \text{eoc} \end{cases}.$$

Nótese además que el estimador insesgado de la autocovarianza es igualmente válido, puesto que su constante normalizadora, $1/(N-|l|)$, también se eliminaría. Por la misma razón, también se pueden sustituir los valores de estimación de la autocovarianza, por los valores respectivos de la autocorrelación,

$$\rho_l = \frac{\gamma_l}{\gamma_0}.$$

Este sistema de ecuaciones, conocido como de *Yule-Walker*, sería resuelto de forma ordinaria por algoritmos generales que solucionan sistemas lineales con un orden de complejidad $O(p^3)$. Sin embargo debido a la estructura *Toeplitz* del mismo puede ser resuelto más eficientemente con una complejidad $O(p^2)$. por el *algoritmo de Levinson*. Este procedimiento, que tiene importantes propiedades por la forma en la que está estructurado, es fundamental en el siguiente capítulo.

1.7. Resumen final

Hemos dispuesto como sección final, un resumen de las distintas aproximaciones lineales expuestas en este capítulo. Podemos simplificar lo desarrollado diciendo que dos grandes tipos de aproximaciones han sido tratadas. Por un lado, las que utilizan como vectores generadores de ellas las exponenciales complejas tanto en su versión continua (sección 1.3) como en su versión

discreta (sección 1.4), y por otro, las que utilizan ‘retardos’ de secuencias como los casos de l^2 (sección 1.5) y los procesos estacionarios (sección 1.6).

Ya se ha puesto de manifiesto que los espacios vectoriales $L^2[-\pi,\pi]$ y l^2 son isomorfos, siendo la serie de Fourier el elemento unificador que nos permite trasladarnos de un conjunto a otro. Por último, el conjunto de vectores en \mathbb{C}^N puede considerarse como un subespacio vectorial de l^2 si consideramos la transformación que asocia a cada vector $\vec{x} = (x_0, x_1, \dots, x_{N-1}) \in \mathbb{C}^N$, la secuencia acotada

$$\{x\}_n = \begin{cases} x_n & \text{si } 0 \leq n \leq N-1 \\ 0 & \text{eoc} \end{cases}.$$

Finalmente ha quedado patente que la metodología usada en la construcción de los modelos autorregresivos para procesos estacionarios es trasladable a la aproximación de secuencias en l^2 . En el siguiente capítulo se usa dicha construcción para la modelización de vectores reales en \mathbb{C}^N usando las coordenadas que poseen en el dominio de la frecuencia.

Capítulo 2 *El Modelo del Filtro aplicado a la transformada discreta de Fourier*

Introducimos en este capítulo los esquemas de predicción lineal sobre l^2 en el entorno de los filtros lineales, o en terminología de *Box y Jenkins*, [BoxJen76], a través de funciones de transferencia. El modelo del filtro es presentado a partir de las relaciones generales que sobre predicción lineal fueron deducidas en el primer capítulo, planteando las hipótesis bajo las cuales la modelización citada construye una buena aproximación.

La novedad del modelo propuesto reside en la secuencia sobre la que se aplica la modelización lineal por retardos: la primera mitad de la *DFT* de una muestra de datos reales. Por las propiedades de la *DFT*, esta primera mitad contiene toda la información acerca de las frecuencias con pesos relevantes por lo que es suficiente una aproximación de la subsecuencia para la construcción de un modelo en el dominio temporal. El hecho de que la aproximación lineal se establezca sobre datos cuya indexación no es cronológica, justifica la generalidad teórica con la que se abordó en el capítulo anterior, y que continúa en el presente prefiriendo usar el concepto *autoproducto escalar sobre retardos* al término *autocorrelación*.

El capítulo está dividido en tres grandes secciones, la primera de las cuales presenta las propiedades básicas de los filtros lineales. La segunda parte describe el modelo del filtro clásico desde los diversos puntos de vista donde ha sido desarrollado, y enumera sus propiedades así como el algoritmo que lo resuelve, el de *Levinson*. En la segunda parte, se aplica sobre la secuencia de pesos de la *DFT* de datos reales, formulando el modelo en ambos dominios (tiempo y frecuencia). Una serie simulada de datos se usa para exponer las principales propiedades de la aproximación que deducirán un criterio único para la elección del orden del modelo. También los resultados de la simulación completarán la interpretación que de la modelización se puede establecer tomando los aspectos de la teoría de la dispersión que intervienen en una formulación particular del algoritmo de Levinson.

2.1. Filtros lineales

Consideraremos en esta sección las secuencias complejas del espacio vectorial de Hilbert l^2 . Debido a que, como se ha visto, \mathbb{C}^N es subespacio vectorial de l^2 , son aplicables las definiciones y resultados aquí expuestos. Notacionalmente indicaremos el paso de un espacio vectorial a otro usando paréntesis, $()$, en el caso de \mathbb{C}^N y llaves, $\{ \}$, en el caso de l^2 , aún cuando podrá deducirse por el contexto fácilmente.

El uso de sistemas filtros lineales es un tópico propio de las Ciencias Físicas y las Ingenierías. Hemos adaptado la terminología al espacio l^2 hablando de secuencias en vez de señales discretas. En esta introducción incluimos junto a las definiciones y resultados iniciales, el uso de la z -transformada para la expresión de estos sistemas así como las condiciones para su estabilidad.

Definición y caracterización

Teniendo en cuenta la definición de la secuencia pulso discreto, $\{\delta\}$, cuyo término general es,

$$\{\delta\}_n = \begin{cases} 1 & \text{si } n=0 \\ 0 & \text{eoc} \end{cases},$$

cualquier secuencia de l^2 puede ser expresada a través de ella y de las tres operaciones que hemos referenciado con la expresión,

$$\{x\} = \sum_k \{x\}_k z^k \{\delta\}, \text{ Ecuación 2.1,}$$

puesto que las secuencias que construyen el sumatorio son,

$$\left\{ \{x\}_k z^k \{\delta\} \right\}_n = \begin{cases} \{x\}_k & \text{si } n = k \\ 0 & \text{eoc} \end{cases}.$$

Un sistema es una aplicación entre secuencias y, como tal, puede ser definido según la notación,

$$\begin{array}{ccc} T: I^2 & \longrightarrow & I^2 \\ \{x\} & \mapsto & \{Tx\} \end{array}.$$

Exigiremos a los sistemas dos propiedades que nos van a permitir expresarlos como si de secuencias se tratara. Diremos que un sistema T es *lineal* si verifica para cualquier combinación lineal de secuencias, $\alpha\{x\} + \beta\{y\}$, $T(\alpha\{x\} + \beta\{y\}) = \alpha\{Tx\} + \beta\{Ty\}$. Por otro lado, diremos que es *invariante ante retardos* si conmuta con el operador retardo, esto es, $z^k \{Tx\} = Tz^k \{x\}$.

Un sistema lineal no se ve afectado por las operaciones de suma y producto por un escalar en el sentido que podemos deshacer la combinación lineal de secuencias, hacer actuar el sistema sobre cada una de ellas, y rehacer las operaciones para obtener la secuencia pedida.

Un sistema invariante ante retardos no se ve afectado por el índice de la secuencia en el sentido de que si la transformación mediante un sistema de este tipo de una secuencia $\{x\}$ es $\{Tx\}$, entonces la imagen de la secuencia retardada $z^k \{x\}$ mediante T es la secuencia $\{Tx\}$ también retardada, esto es, $z^k \{Tx\}$.

Si aplicamos un sistema lineal invariante ante retardos, *LSI*¹ de ahora en adelante, T sobre una secuencia expresada según la *Ecua-ción 2.1* tenemos que,

$$\{Tx\} = T \sum_k \{x\}_k z^k \{\delta\} = \sum_k \{x\}_k z^k \{T\delta\}.$$

Por tanto, sólo es necesario conocer la secuencia $\{T\delta\}$, también denominada *secuencia respuesta al impulso* del sistema *LSI*, T , para construir la transformación de cualquier otra secuencia. En ese caso, si denotamos por $\{h\}$ a la secuencia respuesta al impulso del sistema *LSI*, T , tenemos que,

$$\{Tx\} = \sum_k \{x\}_k z^k \{h\},$$

con lo que el término n -ésimo viene dado por,

$$\{Tx\}_n = \sum_k \{x\}_k \{h\}_{n-k} = \{x * h\}_n$$

definiendo de esta manera una nueva operación entre secuencias, la convolución, cuyo operador hemos denotado por $*$.

En definitiva un sistema *LSI* puede ser identificado con una secuencia, la de respuesta al impulso, de forma que la actuación del sistema sobre las secuencias se traduce a la operación de convolución. Y al contrario, dada una secuencia, podemos construir un sistema *LSI* usando la operación de la convolución para hacerlo actuar.

En el siguiente apartado, se sustituirá la operación convolución por un producto de polinomios, facilitando con ello la manipulación de los filtros lineales.

¹ LSI, LINEAR SHIFT INVARIANT, (*LINEAL E INVARIANTE ANTE TRASLACIONES*), EN INGLÉS.

La z-transformada

Dada una secuencia en l^2 , $\{x\}$, definimos su z -transformada como el polinomio en variable compleja z ,

$$X(z) = \sum_k \{x\}_k z^k.$$

que tiene sentido para aquellos valores de la variable independiente donde el sumatorio converge. El conjunto de valores donde esto se verifica se denomina región de convergencia, *RDC*, y viene normalmente especificado a través de inecuaciones sobre la variable z^1 .

En literatura propia de Ciencias Físicas e Ingenierías es mucho más común expresar la z -transformada de una secuencia en potencias de z^{-1} , esto es, $X(z) = \sum_k \{x\}_k z^{-k}$. Hemos preferido, sin embargo, seguir la notación matemática ya que, de cualquier forma, los resultados son trasladables con sólo efectuar e interpretar un sencillo cambio de variable².

La razón de expresar la z -transformada a través de polinomios en la variable z , (el mismo símbolo que tomamos para denotar el operador retardo), es que se verifica trivialmente que si $X(z)$ es la z -transformada de la secuencia $\{x\}$, entonces $z^k X(z)$ es la z -transformada de la secuencia retardada, $z^k \{x\}$. La elección de la notación para el operador retardo facilita por tanto el paso de l^2 a su expresión mediante z -transformadas y viceversa.

¹ VÉASE [OPPSCH75], ENTRE OTROS

² PODEMOS IMAGINAR QUE, SI TENEMOS UN GRÁFICO DE LA SECUENCIA COMO FUNCIÓN DE SU ÍNDICE SOBRE UN CRISTAL TRANSPARENTE, LO ÚNICO QUE VA A OCURRIR ES QUE UN FÍSICO LA OBSERVARÍA PERO DESDE EL OTRO LADO, ENFRENTA DE NOSOTROS. Y SIN EMBARGO, SIGUE SIENDO LA MISMA SECUENCIA.

Una segunda propiedad de la z -transformada es la de simplificar la acción de la operación convolución convirtiéndola en un producto. En efecto, se comprueba que, si $\{c\} = \{a * b\}$, entonces sus respectivas z -transformadas, $C(z)$, $A(z)$ y $B(z)$ verifican $C(z) = A(z)B(z)$. Esta expresión, más simple en cuanto a operatoria, aísla el efecto de cada secuencia interviniente pudiendo en ese caso invertir el proceso y afirmar, por ejemplo que, $B(z) = C(z)/A(z)$.

Una tercera propiedad es su relación con la *ecuación DFT*. Si tomamos una secuencia de tamaño N , $\{x_0, x_1, \dots, x_{N-1}\}$ su z -transformada

$$X(z) = \sum_{k=0}^{N-1} \{x\}_k z^k,$$

verifica

$$\frac{X(W_N^{-r})}{\sqrt{N}} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \{x\}_k W_N^{-rk} = X_r, \text{ Ecuación 2.2}$$

donde X_r es el término r -ésimo de la *DFT*. Luego la evaluación de la z -transformada en el círculo unidad nos traslada al dominio de la frecuencia.

Estas tres propiedades demuestran la operatividad de la construcción de la z -transformada para la manipulación de sistemas *LSI*, y su relación con la *DFT*.

Denotaremos a partir de ahora la acción de un sistema *LSI* sobre una secuencia a través de sus z -transformadas, denominando a la z -transformada del sistema *LSI*, su *función de transferencia*. Supondremos además que la forma general de una función de transferencia

corresponde a un sistema *LSI* de orden finito, esto es, el definido mediante un cociente de polinomios en $\mathbb{C}[z]$ como,

$$H(z) = \frac{b_0 + b_1 z + b_2 z^2 + \dots + b_q z^q}{1 + a_1 z + a_2 z^2 + \dots + a_p z^p}.$$

Estabilidad y Causalidad

Tras los apartados anteriores quedan aún por resolver dos cuestiones importantes que involucran la buena definición de un sistema *LSI*. En primer lugar, se trata de asegurar que dada una secuencia de partida $\{x\} \in l^2$, el sistema *LSI*, T , sea tal que $\{Tx\}$ pertenezca también a l^2 , es decir, que sea un sistema denominado estable, o más formalmente, l^2 -estable.

En este sentido tomaremos, sin embargo una condición más relajada, la *BIBO estabilidad*¹ o estabilidad acotada que aseguraría que el sumatorio convolución que construye cada término de $\{Tx\}$ converge. Si definimos el espacio normado,

$$l^\infty = \left\{ \{y\} \mid \|\{y\}\|^\infty = \sup_k |y_k| < \infty \right\},$$

la *BIBO estabilidad* se establece cuando de $\{y\} \in l^\infty$ podemos concluir que $\{Ty\} \in l^\infty$, por lo que también se le denomina l^∞ -estabilidad.

Buscando la acotación del sumatorio convolución para cada término tenemos que,

¹ BIBO, BOUNDED INPUT BOUNDED OUTPUT, EN INGLÉS

$$|\{Ty\}_k| = |\{h * y\}_k| = \sum_m \{h\}_m \{y\}_{k-m} \leq |\{y\}|^\infty \sum_m \{h\}_m,$$

con lo que basta que este último sumatorio esté acotado para asegurar la *BIBO estabilidad*.

La segunda cuestión tiene que ver con el cálculo de la secuencia $\{Tx\}$, para un sistema *LSI*, T . Sería deseable que, para el cálculo del término k -ésimo de $\{Tx\}$, no intervinieran los valores $\{x\}_n$, con $n > k$. Dicho de otro modo, si el índice de la secuencia $\{x\}$ fuera cronológico nos referiríamos a que valores presentes dependieran sólo de valores pasados y no de valores futuros. Esta propiedad denominada *causalidad* establece un comportamiento que va a ser necesario exigir en el modelo que se plantea en este capítulo.

Decimos por tanto que un sistema *LSI*, T , es causal si dadas dos secuencias $\{x\}$ y $\{y\}$, tales que $\{x\}_n = \{y\}_n$, $\forall n < k$, entonces $\{Tx\}_n = \{Ty\}_n$ también $\forall n < k$.

De nuevo referimos la propiedad de causalidad de un sistema *LSI* a condiciones sobre la secuencia de respuesta al impulso de dicho filtro. Se demuestra fácilmente que un sistema *LSI*, T , es causal, sii $\{T\delta\}_n = 0$, $\forall n < 0$. Por esa razón, a las secuencias que verifican la anterior propiedad se les denomina causales.

Si consideramos la z -transformada de un sistema *LSI* de orden finito de la forma ya establecida en la sección anterior,

$$H(z) = \frac{b_0 + b_1 z + b_2 z^2 + \dots + b_q z^q}{1 + a_1 z + a_2 z^2 + \dots + a_p z^p}$$

las condiciones sobre causalidad y estabilidad se establecerán sobre la secuencia respuesta al impulso $\{h\} = \{T\delta\}$. Supuesta causalidad de

la secuencia, la condición necesaria y suficiente para lograr la *BIBO estabilidad* se establece sobre las raíces del polinomio $A_p(z) = 1 + a_1z + a_2z^2 + \dots + a_pz^p$, denominado polinomio característico del sistema.

Teorema 2.0.¹- Un sistema LSI con función de transferencia dada por,

$$H(z) = \frac{b_0 + b_1z + b_2z^2 + \dots + b_qz^q}{1 + a_1z + a_2z^2 + \dots + a_pz^p}$$

es BIBO estable sii las raíces de su polinomio característico están fuera del círculo unidad, esto es, en $|z| > 1$.

2.2. Predicción lineal a través del modelo del filtro

Aunque la aproximación lineal mínimo cuadrática podemos situarla ya en los trabajos del matemático Carl Friedrich Gauss en 1795, la primera aplicación del término específico predicción lineal data del libro de N. Wiener de 1949, [Wie66] cuyo capítulo segundo fue titulado '*Un predictor lineal para una serie temporal simple*'. Varios nombres y formulaciones diversas han revisitado su utilización según se iba adaptando a distintos campos de aplicación.

De tratamiento general del filtro de Wiener citamos el fundamental trabajo de Levinson [Lev47] y el de [RobTre67] y los relacionados con su aplicación a la geofísica, [PeaTre69] y [RobTre77]. Una recomendable visión histórica de la predicción lineal puede consultarse en [Kai74], [Mak75] y [MarGra76].

¹ LA DEMOSTRACIÓN DE ESTE TEOREMA PUEDE VERSE, POR EJEMPLO, EN [ROBMUL92] Y [ALO94].

En la presente sección, consideraremos la forma general determinística del modelo de predicción lineal a través de un filtro en dos etapas: análisis y síntesis. Derivaremos el mismo a partir de los resultados de predicción lineal por retardos introducidos en el primer capítulo. La formulación, junto a la expresión de su solución mediante el algoritmo de Levinson, es la que se desarrolla dentro de la teoría de los polinomios ortogonales de Szegő, [Sze67]. Su compilación proviene de los textos de A. Bultheel y M. Van Barel [BulBar93] que hemos adaptado para la modelización que pretendemos.

El modelo del filtro

La relación entre los sistemas *LSI* o filtros lineales y la predicción lineal se formula a través del modelo del filtro. El problema que se plantea es la construcción de un sistema *LSI* cuya función respuesta al impulso sea la secuencia que se quiere aproximar. Otra forma equivalente de plantearlo sería la construcción de un filtro cuya función de transferencia sea la *z*-transformada de la secuencia que se quiere aproximar.

La predicción lineal presentada en la sección 1.5 planteaba la aproximación de una secuencia en l^2 , $\{x\}$, sobre sus retardos de hasta orden p que minimizara la norma del vector error. Esto es, buscaba

$$\Pi\{x\} = \sum_{j=1}^p \alpha^j z^j \{x\}$$

o lo que es equivalente,

$$\Pi\{x\} = F_p(z)\{x\}, \text{ Ecuación 2.3}$$

haciendo intervenir el polinomio en variable z ,

$$F_p(z) = \sum_{j=1}^p \alpha_j z^j.$$

de forma que el vector error de la aproximación, definido como

$$\{e^p\} = \{x\} - \Pi\{x\},$$

tuviera norma mínima. La norma del vector error se obtiene desde,

$$\|\{e^p\}\|^2 = \langle \{x\} - \Pi\{x\}, \{x\} - \Pi\{x\} \rangle = \langle \{x\}, \{x\} \rangle - \langle \{x\}, \Pi\{x\} \rangle - \langle \Pi\{x\}, \{x\} - \Pi\{x\} \rangle.$$

Dado que $\Pi\{x\}$ y $\{x\} - \Pi\{x\}$ son ortogonales según la caracterización del teorema de la proyección, el último sumando desaparece con lo que, denotando por E_p la norma del vector error para una aproximación de hasta orden p , tenemos que,

$$E_p = \langle \{x\}, \{x\} \rangle - \sum_{j=1}^p \alpha_j \langle \{x\}, z^j \{x\} \rangle = r_0 - \sum_{j=1}^p \alpha_j r_{-j}.$$

La secuencia error $\{e^p\}$ puede también ser expresada, teniendo en cuenta la *Ecuación 2.3*, como

$$\{e^p\} = \{x\} - F_p(z)\{x\} = (1 - F_p(z))\{x\} = A_p(z)\{x\}$$

o, a través de las z -transformadas,

$$E_p(z) = A_p(z)X(z) \text{ Ecuación 2.4}$$

en la que ya interviene el polinomio que denominaremos predictor y que denotaremos de la forma

$$A_p(z) = 1 - F_p(z) = \sum_{k=0}^p a_k z^k \text{ con } a_0 = 1 \text{ y } a_k = -\alpha^k \text{ para } 1 \leq k \leq p.$$

La *Ecuación 2.4* formula, a través de la aproximación lineal por retardos, un filtro lineal que, teniendo como función de transferencia el polinomio predictor, actúa sobre la secuencia de z -transformada, $X(z)$, para devolver la secuencia de z -transformada $E_p(z)$. En otros términos, el filtro definido transforma la secuencia que queremos aproximar, $\{x\}$, en la secuencia error, $\{e^p\}$.

En ese caso y bajo condiciones de invertibilidad que serán expuestas en una sección posterior, podemos expresar el filtro en la dirección opuesta,

$$X(z) = \frac{E_p(z)}{A_p(z)}$$

en la que la secuencia ahora obtenida es, sin embargo, la de partida en el proceso de aproximación.

Este proceso debe ser tal que en el polinomio predictor quede la información relevante de la secuencia de partida $\{x\}$. O expresado más formalmente, la z -transformada $A_p(z)$ absorbe de la secuencia $\{x\}$ aquella parte de ella expresable como combinación lineal de sus retardos y deja en la secuencia error la restante información. El modelo del filtro mantiene como hipótesis de trabajo que, en la secuencia error, no existe más que ruido que corresponde a una secuencia de z -transformada $\hat{E}_p(z) = G_p$, con G_p una constante real positiva.

De ese modo y tras la aproximación, $E_p(z)$ se ve sustituida por $\hat{E}_p(z)$ para obtener la modelización de $X(z)$ como,

$$X_p(z) = \frac{G_p}{A_p(z)}, \text{ Ecuación 2.5.}$$

La constante real positiva G_p , llamada *ganancia del filtro*, se escoge de forma que la norma de la secuencia error de la aproximación lineal de orden p , $\{e^p\}$, y la norma de la secuencia por la que ha sido sustituida, $G_p\{\delta\}$, coincidan. En ese caso, tenemos que,

$$G_p^2 = E_p = r_0 - \sum_{j=1}^p \alpha^j r_{-j} = \sum_{k=0}^p a_k r_{-k},$$

interviniendo en la última igualdad los coeficientes del polinomio predictor tal y como fueron definidos.

El resto de las ecuaciones, esto es, aquellas que calculan los coeficientes de la aproximación de orden p , son también reescritas en función de estos nuevos coeficientes como

$$-\sum_{k=1}^p \alpha^k r_{k-j} = \sum_{k=1}^p a_k r_{k-j} = -r_{-j}, \text{ para } j=1, 2, \dots, p,$$

que es equivalente al sistema de ecuaciones,

$$\sum_{k=0}^p a_k r_{k-j} = 0, \text{ para } j=1, 2, \dots, p,$$

Por tanto, en conjunto, los $p+1$ parámetros requeridos para la construcción del filtro, $\{G_p, a_1, a_2, \dots, a_p\}$, se obtienen a través del sistema lineal de expresión matricial,

$$(1 \ a_1 \ a_2 \ \dots \ a_p) \begin{pmatrix} r_0 & r_1 & r_2 & \dots & r_p \\ r_{-1} & r_0 & r_1 & \dots & r_{p-1} \\ r_{-2} & r_{-1} & r_0 & \dots & r_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ r_{-p} & r_{-p+1} & r_{-p+2} & \dots & r_0 \end{pmatrix} = (G_p^2 \ 0 \ 0 \ \dots \ 0)$$

Ecuación 2.6

Tal y como se ha indicado, podemos establecer dos fases en la construcción del modelo del filtro para una secuencia dada $\{x\}$. En una primera, y a través de la aproximación lineal por retardos, un *filtro de síntesis* que, tomando como función de transferencia el polinomio predictor, es alimentado por $\{x\}$ para construir la secuencia error $\{e^p\}$. El proceso en este momento hace una hipótesis de trabajo: la aproximación ha blanqueado la secuencia de forma que el error puede ser sustituido por una secuencia de z-transformada plana. Esta primera fase también es denominada por ello *filtro de blanqueo*. Invirtiendo el sentido del sistema, se construye un filtro cuya secuencia respuesta al impulso es una modelización de la secuencia de partida. A dicho filtro se le denomina *filtro de análisis* y conforma la segunda fase del proceso.

Obteniendo la secuencia de z-transformada $X(z)$ por simple división polinómica suponiendo causalidad, la expresamos término a término como,

$$\{x\}_n = \begin{cases} 0 & \text{si } n < 0 \\ G_p & \text{si } n = 0, \\ -\sum_{k=1}^p a_k \{x\}_{n-k} & \text{si } n > 0 \end{cases}$$

o a través del álgebra de secuencias,

$$\{\hat{x}\} = -\sum_{k=1}^p a_k z^k \{\hat{x}\} + G_p \{\delta\}, \text{ Ecuación 2.7.}$$

El modelo del filtro en \mathbb{C}^N

El cálculo de la matriz de autoproductos escalares T_p debe resolver sumatorios infinitos sobre los valores de las secuencias en l^2 que, si bien sabemos que convergen, presenta lógicos problemas prácticos. Además, en general trabajaremos sobre secuencias construidas a través de muestreos de tamaño finito, N , esto es, secuencias que denotaremos por $\{x_0, x_1, \dots, x_{N-1}\}$, que en l^2 son de la forma ya expresada en el primer capítulo como,

$$\{\mathbf{x}\}_n = \begin{cases} x_n & \text{si } 0 \leq n \leq N-1 \\ 0 & \text{eoc} \end{cases},$$

pretexto que usábamos para hablar de \mathbb{C}^N como subespacio vectorial de l^2 .

Diversos trabajos han formulado las distintas maneras de establecer la matriz T_p suponiendo que de la secuencia biinfinita sólo *podemos observar* ciertos N términos consecutivos. Los cuatro métodos más comunes en la literatura se pueden exponer definiendo una cierta matriz $C_{p,m,M}$ de orden $(p+1) \times (M-m+1)$ para ciertos valores de M y m según el esquema siguiente,

$$C_{p,m,M} = \begin{pmatrix} x_0 & x_1 & \dots & x_p & \dots & x_{N-1} & 0 & \dots & 0 \\ 0 & x_0 & \dots & x_{p-1} & \dots & x_{N-2} & x_{N-1} & \dots & 0 \\ 0 & 0 & \dots & x_{p-2} & \dots & x_{N-3} & x_{N-2} & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & x_0 & \dots & x_{N-p-1} & x_{N-p} & \dots & 0 \end{pmatrix}$$

$m=0$ $m=p$ $M=N-1$ $M=N+p-1$

La matriz T_p se define a través del producto de matrices $C_{p,m,M} \overline{C_{p,m,M}}^t$, indicando en cada caso m la columna inicial y M la columna final según el gráfico anterior, t la matriz traspuesta y $\bar{}$ la conjugada compleja.

La *Tabla 2.1* muestra la denominación de los diferentes métodos según la elección de los valores iniciales y finales.

Columna inicial	Columna Final	Denominación del método
$m=0$	$M=N+p-1$	Método de la autocorrelación
$m=0$	$M=N-1$	prewindowed
$m=p$	$M=N+p-1$	postwindowed
$m=p$	$M=N-1$	Método de la covarianza

Tabla 2.1 Métodos para la construcción de T_p

De los cuatro casos, sólo el primero corresponde a una matriz *Toeplitz* obtenida a través de un autoproducto escalar. Por esa razón

es una matriz de *Gram*¹ y semidefinida positiva. La estructura *Toeplitz* le va a permitir una resolución eficiente mediante el algoritmo de Levinson. La segunda propiedad implica que es definida positiva en el caso de que el sistema de retardos sea linealmente independiente, lo que asegurará la estabilidad del filtro invertido empleado en la fase de *síntesis* del modelo.

El resto de los métodos no construyen una matriz T_p con estas propiedades con lo que no podemos asegurar, en el peor caso, estabilidad para el filtro sintetizado.

Propiedades del modelo de filtro

El sistema expresado en *Ecuación 2.6* construye, a través del modelo del filtro, una secuencia $\{x\}$ según la igualdad de la *Ecuación 2.7*. La secuencia modelo construida y el sistema lineal verifican importantes propiedades que comentaremos a continuación.

Hemos denotado por r_l los valores de los autoprodutos escalares a retardos de longitud l para la secuencia $\{x\}$. Denotaremos de forma análoga por \hat{r}_l dicha operación para la secuencia $\{\hat{x}\}$. Usando la expresión dada de $\{\hat{x}\}$ en la *Ecuación 2.7* obtenemos el siguiente resultado para \hat{r}_l ,

$$\hat{r}_0 = G_p^2 - \sum_{k=1}^p a_k \hat{r}_{-k}$$

$$\hat{r}_l = -\sum_{k=1}^p a_k \hat{r}_{l-k}, \text{ para valores de } 1 \leq l \leq p.$$

¹ VÉASE POR EJEMPLO [LANTIS85]

Esto es, los valores $\{\hat{r}_0, \hat{r}_1, \hat{r}_2, \dots, \hat{r}_p\}$ verifican las mismas relaciones lineales que los valores $\{r_0, r_1, r_2, \dots, r_p\}$ con respecto a los coeficientes del polinomio predictor. Por las propiedades de la inversa de la matriz de *Toeplitz* T_p^{-1} se demuestra que $r_0 = \hat{r}_0$ y que además $r_l = \hat{r}_l$ para valores de $1 \leq l \leq p$.

Nótese que en ese caso, si $r_0 = \hat{r}_0$, tenemos que la norma de la secuencia que se modeliza $\{x\}$ y la construida por el modelo $\{\hat{x}\}$ coinciden. Además el problema resuelto en la modelización se puede replantear de la siguiente forma: *búsquese una secuencia $\{x\}$ con igual secuencia de autoproducos hasta orden p que la dada $\{x\}$.*

Si consideramos la expresión de la z -transformada de los valores r_l denotándola como,

$$R(z) = \sum_l z^l r_l$$

se puede expresar como el producto de z -transformadas

$$R(z) = X(z)\tilde{X}(z),$$

donde, recordamos, $X(z)$, es la z -transformada de la secuencia $\{x\}$, y definimos la acción del operador *recíproco*, \sim , como

$$\tilde{X}(z) = \overline{X\left(\frac{1}{z}\right)}, \text{ Ecuación 2.8.}$$

Puesto que el modelo del filtro construye la z -transformada de la secuencia modelizada en forma cerrada, tenemos, por la *Ecuación 2.5* y el resultado previo, que

¹ FÓRMULA DE GOHBERG-SEMENCUL, VÉASE [BULBAR93]

$$\hat{R}(z) = \sum_I z^I \hat{r}_I = \hat{X}(z) \tilde{X}(z) = \frac{G_p^2}{A_p(z) \tilde{A}_p(z)}$$

con lo que un nuevo enfoque puede darse al problema de modelización resuelto. Éste resulta de comprobar que $\hat{R}(z)$ así expresado es un cociente de polinomios en z de grado 0 en el numerador que ajusta la serie biinfinita $R(z)$ hasta el grado p . Esto es, es un aproximante de *Laurent-Padé* $[0/p]$. Esta propiedad será el punto de partida del capítulo siguiente.

Ya han sido expuestas en una sección anterior las dos fases de las que consta la modelización expuesta. Nótese que, para la elaboración del polinomio predictor, la fuente de información parte realmente de los autoprodutos escalares y no de la secuencia $\{x\}$ inicial. Podríase, en ese caso, plantearse la resolución de este problema simplemente dando la secuencia r_I y trabajando sobre ella. Es por lo que se establece una relación entre dos conjuntos de parámetros: por un lado la secuencia comentada $\{r_0, r_1, \dots, r_p\}$ y por otro $\{G_p, a_1, \dots, a_p\}$. Para estos conjuntos se definen los siguientes problemas, inverso el uno del otro:

- I. Dado $\{r_0, r_1, \dots, r_p\}$ calcúlese $\{G_p, a_1, \dots, a_p\}$ verificando,

$$(1 \ a_1 \ a_2 \ \dots \ a_p) T_p = (G_p \ 0 \ 0 \ \dots \ 0), \text{ Ecuación 2.9}$$
 que reconocemos como la fase de *análisis o blanqueo* ya comentada.

- II. Dado $\{G_p, a_1, \dots, a_p\}$ calcúlese $\{r_0, r_1, \dots, r_p\}$ verificando,

$$(r_0 \ r_1 \ r_2 \ \dots \ r_p) Q_p = (G_p \ 0 \ 0 \ \dots \ 0), \text{ Ecuación 2.10}$$
 siendo Q_p la matriz de Jury, [Jur64],

$$Q_p = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_p & 0 & 0 & \dots & 0 \\ a_{p-1} & a_p & 0 & \dots & 0 \\ \dots & a_{p-1} & a_p & \dots & 0 \\ a_2 & \dots & a_{p-1} & \dots & a_p \\ a_1 & a_2 & \dots & \dots & a_{p-1} \\ 1 & a_1 & a_2 & \dots & \dots \\ 0 & 1 & a_1 & \dots & a_2 \\ 0 & 0 & 1 & \dots & a_1 \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

que permite construir el vector de autoproductos escalares. Reconocemos en este caso la fase de *síntesis* del modelo del filtro, también comentada anteriormente.

Para ambos sistemas lineales, equivalentes por su formulación, establecemos los siguientes teoremas, que dan las condiciones para asegurar existencia y unicidad en sus soluciones, así como buena definición de los parámetros hallados. Nos referimos con *buena definición* a las siguientes propiedades:

a) Para el conjunto de parámetros $\{r_0, r_1, \dots, r_p\}$: que construya una matriz Toeplitz T_p definida positiva. Esto es, la matriz *Gram* de un sistema de vectores linealmente independiente¹.

b) Para el conjunto de parámetros $\{G_p, a_1, \dots, a_p\}$: que construya un filtro lineal estable. Esto es, que el polinomio predictor $A_p(z)$ al pasar del numerador al denominador posea sus raíces fuera del círculo unidad. Esta propiedad también se denomina *invertibilidad* o *condición de mínima fase*².

¹ VÉASE EN [BRODAV91], CONDICIONES NECESARIAS Y SUFICIENTES PARA QUE $\{r_0, r_1, \dots, r_p\}$ SEA UNA FUNCIÓN DE AUTOCOVARIANZA COMPLEJA.

² VÉASE EN [OPPSHA75]

Los teoremas mencionados¹ son:

Teorema 2.1 El sistema lineal de la *Ecuación 2.9* tiene solución única, $G_p > 0$ y el filtro que construye es estable si la matriz Toeplitz T_p es definida positiva.

Teorema 2.2 El sistema lineal de la *Ecuación 2.10* tiene solución única, y la matriz T_p que se genera es definida positiva si las raíces del polinomio predictor $1 + \sum_{k=1}^p a_k z^k$ están situadas fuera del círculo unidad.

En definitiva, basta con que el sistema de vectores construidos por acción del operador retardo sobre la secuencia que se quiere aproximar sea linealmente independiente, para que el filtro que genera su aproximación para un determinado orden sea único respecto a los autoproducidos escalares y estable.

El algoritmo de Levinson, (1947)

El sistema lineal de la *Ecuación 2.6* expresa las relaciones lineales de los parámetros necesarios para la construcción de un filtro lineal de orden p con las propiedades vistas en la sección anterior.

Para cualquier sistema lineal, los algoritmos eficientes basados en métodos de eliminación de Gauss o triangularización de Cholesky necesitan del orden de p^3 operaciones para su resolución. Sin em-

¹ CONSÚLTASE LA DEMOSTRACIÓN EN [ROBMUL87], Y, ACERCA DE LA ESTABILIDAD DE LIAPUNOV EN [LANTIS85].

bargo, la propiedad de estructura *Toeplitz* de la matriz T_p del sistema de ecuaciones reduce el orden de su resolución a $O(p^2)$.

El algoritmo clásico de Levinson, que llamamos así para distinguirlo de otras versiones más actuales, fue presentado en el trabajo de este autor, *N. Levinson*, titulado "*The Wiener RMS error criterion in filter design and prediction*" en 1947.

Dados los valores $\{r_0, r_1, r_2, \dots\}$, y el orden p , la expresión vectorial del algoritmo es la siguiente,

$$E_0 = r_0;$$

$$A_0 = (1);$$

Para $n=1$ hasta p hacer

Si $E_{n-1} \neq 0$ entonces

$$\rho_n = \frac{A_{n-1}(r_n \ r_{n-1} \ \dots \ r_2 \ r_1)}{E_{n-1}}$$

$$A_n = (A_{n-1} \ 0) - \rho_n (0 \ \tilde{A}_{n-1}) \text{ siendo } \tilde{A}_j = (\overline{a_j} \ \overline{a_{j-1}} \ \dots \ \overline{a_1} \ 1)$$

$$E_n = (1 - |\rho_n|^2) E_{n-1}$$

Algoritmo clásico de Levinson

Estableciéndose la salida del algoritmo en la matriz final $A_p = (1 \ a_1 \ a_2 \ \dots \ a_p)$ y en el valor $G = \sqrt{E_p}$.

La comprobación del método se realiza a través de la verificación del invariante del bucle, es decir,

$$A_n T_n = (1 \ a_1 \ a_2 \ \dots \ a_n) T_n = (E_n \ 0 \ 0 \ \dots \ 0)$$

por inducción sobre n , usando la propiedad de que $T_p^t = \overline{T_p}$, es decir, su transpuesta es su conjugada compleja.

Nótese que para el cálculo de los coeficientes del polinomio predictor $A_p(z)$ y de la ganancia del filtro G_p es necesario calcular los polinomios y ganancias de orden inferior. Es más, el polinomio de grado n se construye operando sobre los polinomios de grado $n-1$, según la expresión polinomial,

$$A_n(z) = A_{n-1}(z) - \rho_n z^{k+1} \tilde{A}_{n-1}(z),$$

que utiliza el coeficiente ρ_n para una ponderación de los coeficientes.

En este sentido el algoritmo puede verse como la actualización desde el grado $n-1$ hasta el n a través del coeficiente ρ_n . Éste recibe varios nombres según el campo de aplicación de la predicción lineal, como pueden ser, *coeficiente de correlación parcial -PARCOR-* (en su interpretación estocástica), *coeficientes de Schur o Szegő (teoría de polinomios ortogonales)*, o *coeficientes de reflexión (teoría de la dispersión)*... Es precisamente esta última aplicación la que escogemos para establecer la interpretación del modelo que se presenta en una sección posterior.

La secuencia de errores, E_n , también se obtiene a partir de los coeficientes $|\rho_n|$ mediante la expresión,

$$E_n = (1 - |\rho_n|^2) E_{n-1}.$$

Si consideramos la unicidad del predictor y el objetivo de minimizar la norma de la secuencia error, precisamente E_n , tenemos que se establece el orden

$$r_0 = E_0 \geq E_1 \geq \dots \geq E_p \geq \dots \geq 0,$$

con lo que se verifica que $|\rho_n| < 1$ para $n = 1, 2, 3, \dots$. En particular, según se afirma en [Bul92], 'si $|\rho_n| = 1$ entonces $E_n = 0$ y en ese caso el modelo encaja perfectamente'. Sin embargo, nótese que en ese caso el modelo del filtro no puede ser aplicado pues $G_n = 0$.

En el lado opuesto, si los coeficientes son de módulo próximo a cero, $|\rho_n| \cong 0$, la norma del vector error se mantiene prácticamente constante, esto es, la adición de un nuevo retardo más no mejora significativamente la aproximación.

<i>Autor</i>	<i>Función penalty</i>
<i>Akaike, 1970</i>	$E_p \frac{N+p-1}{N-p-1}$
<i>Akaike, 1974</i>	$\ln(E_p) + \frac{2p}{N}$
<i>Parzen, 1974</i>	$\left(\frac{1}{N} \sum_{j=1}^p \frac{1}{E_j} \right) - \frac{1}{E_p}$

Tabla 2.2

Identificación del orden del polinomio predictor

La elección del orden de la aproximación p está obstaculizada por el objetivo de la minimización. Éste ha sido elegido de forma que un parámetro más mejora, aunque sea de forma inapreciable, el modelo y hace más pequeño el error cometido por éste. Se han de conjugar entonces dos criterios para la elec-

ción de este orden: por un lado, el buen ajuste, y de otro, evitar la sobreparametrización del modelo. Se plantea por lo tanto una optimización bicriterio que ha sido resuelta mediante la minimización de funciones uniobjetivo por diversos autores. Citamos en la *Tabla 2.2*

los trabajos de Akaike y Parzen que aparecen junto a otros métodos de identificación de modelos ARMA en [Cho92].

Queremos recoger finalmente la aportación que diversos autores han hecho para la mejora computacional en la resolución de sistemas lineales cuya matriz es *Toeplitz*. Una versión más *rápida* del algoritmo de Levinson clásico, denominada versión *split* puede encontrarse en la serie de trabajos de los autores P. Desarte y Y. Genin en [DelGen86] y [DelGen91]. Estos trabajos presentan una económica relación de recurrencia de las partes simétricas de los polinomios predictores eliminando de este modo cierta redundancia de la versión clásica. Además, la justificación teórica de la versión *split* ha abierto, en la teoría de los polinomios ortogonales de Szegő nuevas líneas de estudio. La versión clásica, sin embargo, es suficiente para la descripción de la interpretación que realizaremos del modelo planteado por lo que hemos preferido su uso al de la reciente versión. Finalmente también queremos citar el trabajo [AmmGra88] que rebaja la complejidad a $O(p \log_2^2 p)$ en la resolución de este tipo de sistemas, usando propiedades de la *FFT*.

Por último queremos incluir las conclusiones que, sobre la estabilidad del algoritmo clásico de Levinson, establece G. Cybenko en [Cyb80]. Este autor, que comenta las diversas controversias en numerosos trabajos sobre el particular que habían llegado a aconsejar la no utilización del mencionado método, concluye asegurando su estabilidad numérica, comparable incluso al método de Cholesky.

2.3. El modelo del filtro aplicado a la DFT

Presentamos en esta sección una modelización basada en el modelo del filtro sobre la transformada discreta de Fourier de una

muestra de tamaño N de datos reales. Las propiedad de conjugada par de la DFT en este caso simplificará la modelización a los $N/2$ primeros pesos, logrando con el modelo del filtro una buena aproximación para la z -transformada.

El caso que se modeliza - cuando la muestra finita es de valores reales - es la más común en series de datos cronológicos por lo que hablamos de un caso muy general. Además, esta redundancia de la información en la DFT de la muestra es el punto de partida de la transformada de Hartley¹ que usa la parte real y la imaginaria de una secuencia compleja para evitarla.

Tras la formulación del modelo propuesto, se presentarán algunas propiedades que relacionan la secuencia de errores con la convergencia del modelo en una aplicación simulada y, a partir de los coeficientes de reflexión, se aportará una interpretación del modelo.

Formulación del modelo

Las Ecuaciones DFT e IDFT del capítulo 1 mostraban las relaciones de cambio de base entre las coordenadas cartesianas y las de la base de vectores exponenciales en \mathbb{C}^N . Dado un vector $(x_0, x_1, \dots, x_{N-1}) \in \mathbb{C}^N$ son

$$X_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k W_N^{-rk}, \text{ Ecuación DFT,}$$

$$x_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} X_r W_N^{rk}, \text{ Ecuación IDFT.}$$

¹ VÉASE, [BRA86]

Si el vector es real, esto es, $(x_0, x_1, \dots, x_{N-1}) \in \mathbb{R}^N \subset \mathbb{C}^N$, entonces la *DFT* es conjugada par, es decir, $X_r = \overline{X_{N-r}}$. De esta forma, podemos obtener toda la secuencia refiriéndola sólo a los $N/2$ primeros pesos ya que $X_r = \overline{X_{N-r}}$ si $r > N \text{div} 2$, siendo *div* la división entera.

Combinando este último resultado con la expresión de la *Ecuación IDFT* tenemos que,

$$\sqrt{N}x_k = \sum_{r=0}^{N \text{div} 2} X_r W_N^{rk} + \sum_{r=N \text{div} 2+1}^{N-1} \overline{X_{N-r}} W_N^{rk},$$

que, tras un cambio de índices queda como,

$$\sqrt{N}x_k = 2 \operatorname{Re} \left(\sum_{r=0}^{N \text{div} 2} X_r W_N^{rk} \right) - X_0 + (N \bmod 2 - 1) \overline{X_{N \text{div} 2}} W_N^{(N \text{div} 2)k}, \text{ Ecuación 2.11,}$$

siendo *Re* la función parte real de un complejo y *mod* el resto de la división entera.

Consideremos ahora la serie finita de datos, $\{X_0, X_1, \dots, X_{N \text{div} 2}\}$ y apliquemos sobre ella el modelo del filtro. Obtenemos entonces la aproximación $\{\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{N \text{div} 2}, \dots\}$ en dos formas equivalentes,

mediante la z-transformada

$$\sum_{r=0}^{\infty} \hat{X}_r z^r = \frac{G_p}{A_p(z)},$$

y en forma recursiva causal,

$$\hat{X}_r = \begin{cases} 0 & \text{si } r < 0 \\ G_p & \text{si } r = 0, \\ -\sum_{j=1}^p a_j \hat{X}_{r-j} & \text{si } r > 0 \end{cases}$$

En ambos casos los parámetros $\{G_p, a_1, a_2, \dots, a_p\}$ son los calculados en la *Ecuación 2.6* del modelo del filtro, y $a_0 = 1$.

La *Ecuación 2.11* va a ser aproximada en las siguientes dos fases:

1 Sustitución del sumatorio $\sum_{r=0}^{N_{div} 2} X_r W_N^{rk}$ por $\sum_{r=0}^{N_{div} 2} \hat{X}_r W_N^{rk}$, justificándose tal cambio en el ajuste de la modelización de $\{X_0, X_1, \dots, X_{N_{div} 2}\}$ por $\{\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{N_{div} 2}\}$, obtenida ésta última por aplicación de la forma recursiva del modelo.

2 Sustitución del sumatorio finito $\sum_{r=0}^{N_{div} 2} \hat{X}_r W_N^{rk}$ por el infinito $\sum_{r=0}^{\infty} \hat{X}_r W_N^{rk}$ justificándose el cambio por la convergencia $|X_r| \xrightarrow{r \rightarrow \infty} 0$ dado que $\|\{X\}\| = \|\{X\}\| < \infty$. En este caso, los valores del sumatorio infinito se obtienen mediante su forma cerrada $\frac{G_p}{A_p(z)}$.

Además, sustituyendo el valor X_0 por su aproximación \hat{X}_0 , o lo que es lo mismo, G_p y eliminando el residuo $X_{N_{div} 2}$ debido a su mínima relevancia, queda finalmente expresado el modelo que presentamos en base a los parámetros del filtro, $\{G_p, A_p(z)\}$ como,

$$\hat{x}_k = \frac{2G_p}{\sqrt{N}} \operatorname{Re} \left(\frac{1}{A_p(W_N^k)} \right) - \frac{G_p}{\sqrt{N}}, \text{ Ecuación 2.12.}$$

Finalizamos la presente sección anotando tres comentarios acerca de la formulación del modelo:

1. *Sobre la naturaleza periódica del modelo.* En efecto, ya que $W_N^{k+N} = W_N^k$, la secuencia \hat{x}_k es periódica discreta de periodo N . Es una consecuencia lógica de basarse en una modelización de la *DFT* que también posee esta propiedad. En ese caso, la modelización será un ajuste en el intervalo muestral en general, y para todo índice, si la secuencia de partida es periódica.

2. *Sobre el sistema generador de la aproximación.* Los p retardos de $(X_0, X_1, \dots, X_{N_{div}2})$, a saber,

$$\begin{aligned} &(X_1, X_2, \dots, X_{N_{div}2}, 0) \\ &(X_2, X_3, \dots, X_{N_{div}2}, 0, 0) \\ &\dots \\ &(X_p, X_{p+1}, \dots, X_{N_{div}2}, 0, \dots, 0) \end{aligned}$$

son, en general linealmente independientes. En ese caso, T_p es una matriz de *Gram* y en consecuencia definida positiva, con lo que se asegura la estabilidad e invertibilidad del filtro.

3. *Sobre la computación eficiente del polinomio $A_p(z)$* para valores de z en el círculo unidad, $z = W_N^k$. Nótese que el sumatorio finito $\sum_{r=0}^p a_p W_N^{rk}$ puede abordarse como una aplicación de la inversa de la transformada discreta de Fourier a una secuencia de longitud N de la forma $(a_0 a_1 \dots a_p 0 0 \dots 0)$. En ese caso el algoritmo de la *FFT* (véase capítulo 1) es el apropiado para su cálculo.

Una aplicación a una serie de datos simulados

Presentamos en esta sección los resultados de la aplicación de la modelización propuesta a una secuencia simulada de datos. Puesto que el modelo del filtro se aplica en el dominio de la frecuencia y son las propiedades de la *DFT* las que se usan para pasarlo al dominio temporal, la aplicación parte de la simulación de una serie de Fourier.

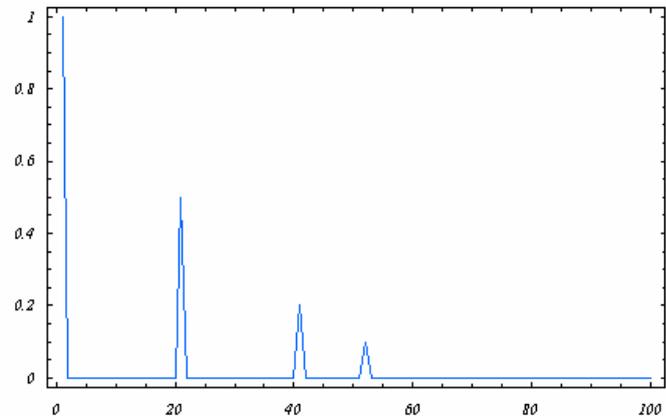


Gráfico 2.1.
Secuencia $(X_0, X_1, \dots, X_{N_{div}2})$ de la simulación.

Se trata de una secuencia que correspondería a la primera mitad de una *DFT* de una serie de datos reales de un tamaño $N=100$. Hemos supuesto que las frecuencias con peso apreciable se sitúan en los índices 0, 20, 40 y 51.

La secuencia generada, que se muestra en el *Gráfico 2.1*, es típica para las secuencias periódicas, y, si hemos preferido situar sólo 4 frecuencias con pesos relevantes, es a causa de la simplificación en el seguimiento de la modelización y su interpretación pero sin pérdida de generalidad.

Una primera apreciación que hacemos es que la secuencia de los auto-productos escalares de los retardos, imita en su módulo el comportamiento

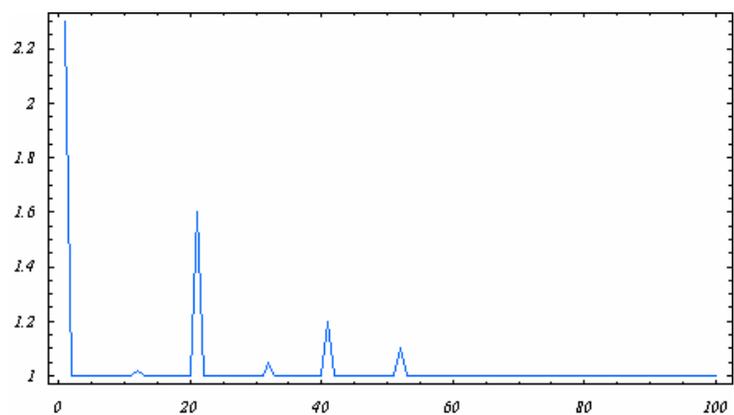


Gráfico 2.2.
Autoproducidos escalares para los diferentes retardos.

gráfico de la *DFT*. Esta propiedad resulta muy importante si tenemos en cuenta que tanto $(X_0, X_1, \dots, X_{N_{div}2})$ como $\{\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{N_{div}2}, \dots\}$, tienen coincidentes autoproductos hasta el orden de la aproximación, esto es, $r_l = \hat{r}_l$, para valores $l=0, 1, \dots, p$. Los autoproductos de los retardos de la secuencia $(X_0, X_1, \dots, X_{N_{div}2})$ para la serie simulada aparecen en el *Gráfico 2.2*.

Debido a esta similitud podemos plantear el ajuste del modelo en cuanto a estos autoproductos escalares, de manera que el modelo y la secuencia coincidan en ellos hasta los índices relevantes. Nótese que en este caso además de r_0 están r_{20} , r_{40} y r_{51} siendo el resto inapreciable y por lo tanto necesitamos que el modelo tenga un orden al menos $p=51$.

Para $l > 51$ los autoproductos del modelo, \hat{r}_l , tenderán hacia 0 cuando $l \rightarrow \infty$, con lo que obtenemos un buen ajuste de autoproductos que es suficiente para la bondad del modelo. El resultado del ajuste se muestra en el *Gráfico 2.3* interpretándose con la siguiente leyenda: en la parte superior y desplazada 1 unidad de su valor, la secuencia original; en la parte inferior, el ajuste para el orden mencionado. Comprobamos que el modelo se comporta de forma excelente hasta el orden elegido, tras el cual aporta pesos falsos aunque irrelevantes en las frecuencias de índice 60, 71, 80 y 91. Nótese que estos índices con pesos espúreos se pueden expresar de la siguiente forma: $60=20+40$, $71=51+20$, $80=40+40$ y $91=51+40$, esto es, sumas de los índices con peso relevante en la simulación de datos. Este efecto se produce al usar los autoproductos escalares de los retardos de distinto orden como medida de similitud.

Otra forma de abordar la bondad del modelo y, en este caso, la determinación de su orden, es a través de un seguimiento de los errores. La norma del error, E_p , interviene en el modelo del filtro de forma que la ganancia de éste es $G_p = \sqrt{E_p}$. Por otro lado y según la Ecuación 2.7 $X_0 = G_p$, por lo que, el grado de ajuste del modelo se puede medir por el grado de convergencia de G_p a X_0 .

En este sentido, cabe añadir que se verifica la siguiente igualdad entre los determinantes de la matriz T_p ,

$$|T_n| = E_n |T_{n-1}|,$$

con lo que, de forma iterativa,

$$|T_n| = E_0 E_1 E_2 \dots E_n.$$

Esta propiedad es importante puesto que deduce el comportamiento numérico del algoritmo de Levinson en esta modelización. Con $|T_p|$ tendiendo a cero, se produciría un mal condicionamiento por *under-flow* en los cálculos intervinientes. Este hecho no se produce en general en la modelización presentada si $X_0 > 0$.

Queremos hacer notar que, según las hipótesis del modelo, la ganancia del filtro recoge la información de la secuencia no expresable como combinación lineal de

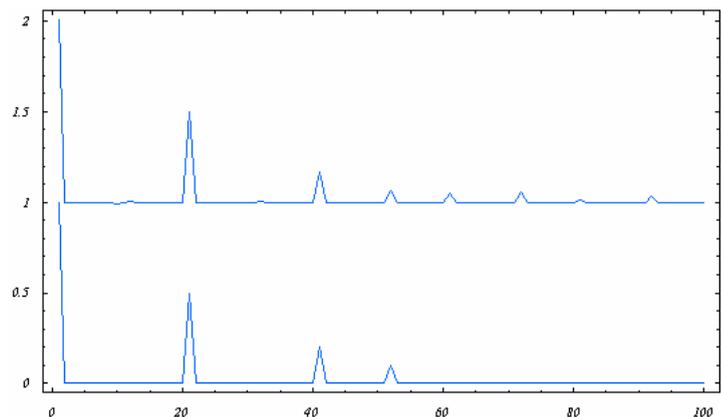


Gráfico 2.3.
Ajuste para $p=51$ de la secuencia simulada.

retardos. Este ‘error’ por lo tanto no es tal, puesto que su anulación impediría la fase de síntesis.

Conviene no confundir la ganancia del filtro con el error de la modelización, que, en nuestro caso, tiene dos vertientes,

error del modelo de la DFT, $X_r - \hat{X}_r$, para $r = 0, 1, \dots, N \text{div} 2$.

error del modelo en el dominio del tiempo, $x_k - \hat{x}_k$, para $k = 0, 1, \dots, N-1$.

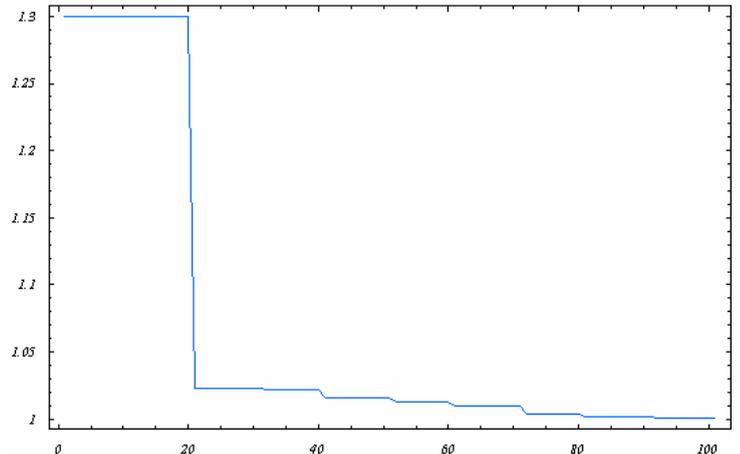


Gráfico 2.4
Valores de E_p para los distintos órdenes del modelo

Usaremos estos últimos en una comparativa con la *DFT* truncada de igual número de parámetros, en todos los modelos desarrollados en el capítulo 4.

La convergencia de la ganancia del filtro será usada, en cambio, para el establecimiento del orden del modelo.

En el *Gráfico 2.4* podemos constatar la convergencia en forma de escalones de los valores de E_p al valor $\sqrt{X_0} = 1$. Los puntos en donde se produce el cambio de tendencia acelerando la aproximación son los índices 20, 40 y 51 que justamente corresponden a los índices para los que los valores de la *DFT* de la serie que analizamos tienen relevancia. Este hecho justifica la interpretación del modelo que realizamos en la sección siguiente basándonos en la relación del algoritmo de Levinson con la teoría de la dispersión de la energía.

Interpretación del modelo

La modelización por retardos en el dominio del tiempo tiene de forma estructural una interpretación clara: el establecimiento de valores *actuales* de una variable por ponderación lineal de sus valores *pasados*. Asimismo los pesos que aparecen en dicha ponderación hablan de la relevancia de ciertos retardos sobre otros.

En cambio, en el dominio de la frecuencia, el hecho de establecer una relación entre el peso de una frecuencia como combinación lineal de los pesos de las frecuencias de índice inferior no se justifica de esa manera intuitiva. Presentamos aquí una interpretación diferente.

Como ya se ha indicado al comienzo de la sección 1.2 la predicción lineal destaca por el numeroso y variado campo de aplicabilidad que ha desarrollado desde su formulación. El algoritmo que la resuelve y construye a través del modelo del filtro, el de Levinson, sobresale a su vez por sus diferentes expresiones. Una de ellas, es usada en geofísica para el estudio de las distintas capas que componen el interior de la Tierra.

El polinomio predictor $A_p(z)$ y su recíproco $z^p \tilde{A}_p(z)$, ya definidos, se relacionan en el algoritmo de Levinson de la siguiente forma matricial, [RobTre77],

$$\begin{pmatrix} z^n \tilde{A}_n(z) \\ A_n(z) \end{pmatrix} = \begin{pmatrix} 1 & -\overline{r_n} \\ -r_n & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z^{n-1} A_{n-1}(z) \\ A_{n-1}(z) \end{pmatrix}.$$

usando los coeficientes, ρ_n a los que nos hemos referido como de reflexión.

Si bien el polinomio predictor era introducido como sistema que construía la secuencia error,

$$\{e^p\} = A_p(z)\{x\},$$

su recíproco definirá otra secuencia error de forma análoga,

$$\{f^p\} = z^p \tilde{A}_p(z)\{x\} = \sum_{k=0}^p \overline{a_k} z^{p-k} \{x\},$$

con lo que su término n -ésimo

$$\{f^p\}_n = \sum_{k=0}^p \overline{a_k} \{x\}_{n-p+k} = \{x\}_{n-p} + \overline{a_1} \{x\}_{n-p+1} + \overline{a_2} \{x\}_{n-p+2} + \dots + \overline{a_p} \{x\}_n$$

expresa el error cometido al escribir $\{x\}_{n-p}$ en función de valores de índices superiores. Al verificarse que $E_p = \|\{f^p\}\|$, podemos concluir que el cálculo de los coeficientes $\{a_1, a_2, \dots, a_p\}$ minimiza la norma de ambas secuencias error.

Si trabajamos con los polinomios normalizados por la ganancia del filtro, esto es,

$$A'_n(z) = \frac{A_n(z)}{G_n} \quad \tilde{A}'_n(z) = \frac{\tilde{A}_n(z)}{G_n}$$

la relación matricial es, en cambio,

$$\begin{pmatrix} z^n \tilde{A}'_n(z) \\ A'_n(z) \end{pmatrix} = \frac{1}{\sqrt{1-|r_n|^2}} \begin{pmatrix} 1 & -\overline{r_n} \\ -r_n & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z^{n-1} \tilde{A}'_{n-1}(z) \\ A'_{n-1}(z) \end{pmatrix} \quad \text{Ecuación 2.13}$$

que permite comprobar que, conocido el conjunto de coeficientes de reflexión, éste puede ser usado para construir de forma iterativa am-

bos polinomios normalizados y con ello el filtro lineal correspondiente.

Un sistema dinámico multicapa al ser atravesado por una onda, presenta el aspecto que se ha esquematizado en el *Gráfico 2.5*, [RobTre77] y [BulBar93]. Las hipótesis que se toman para la simplificación del medio son las siguientes,

- Una onda al llegar al límite de una capa responde al cambio de densidad dividiéndose en una ascendente (*reflexión*) y otra incidente (refracción) que continúa su camino hacia el interior.
- Los estratos son homogéneos en cuanto a densidad a lo largo de la capa, y no absorben ni añaden energía a la onda.

Se aplica por lo tanto la ley de conservación de la energía en la subdivisión de ondas en el paso de capas, obteniendo una relación recursiva entre las ondas descendentes de la capa i -ésima, $\{f^i\}$ y las ascendentes de la capa i -ésima $\{e^i\}$, esto es,

$$\begin{pmatrix} \{f^i\}_n \\ \{e^i\}_n \end{pmatrix} = \frac{1}{\sqrt{1-|\rho_n|^2}} \begin{pmatrix} 1 & -\bar{\rho}_n \\ -\rho_n & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \{f^i\}_{n-1} \\ \{e^i\}_{n-1} \end{pmatrix}$$

Ecuación 2.14

donde el coeficiente ρ_n representa un valor directamente proporcional a la densidad del medio.

La equivalencia de las expresiones 2.13 y 2.14, justifica la siguiente aplicación al estudio de la densidad de las capas que conforman el interior de la tierra. Supongamos que en la superficie de la tierra se efectúa una ex-

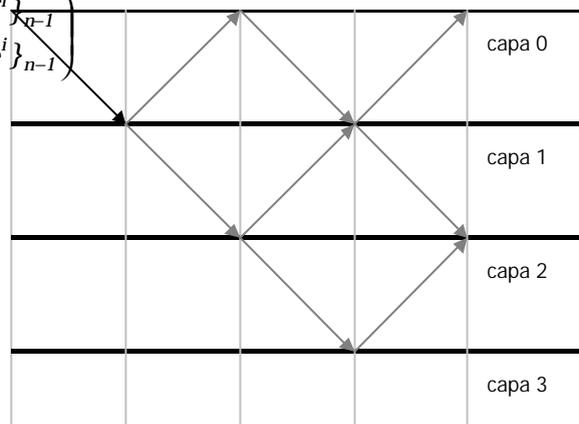


Gráfico 2.5
Esquema del medio estratificado

plosión controlada, midiéndose a continuación los ecos de las ondas reflejadas. Aplicando el modelo estratificado, el algoritmo de Levinson irá calculando los coeficientes de reflexión de ondas de cada capa, ρ_n , (de ahí su nombre) que indicarán medidas relativas de la densidad de cada una de ellas. Por lo tanto, pueden detectarse capas de gas o petróleo sin acudir a las excavaciones.

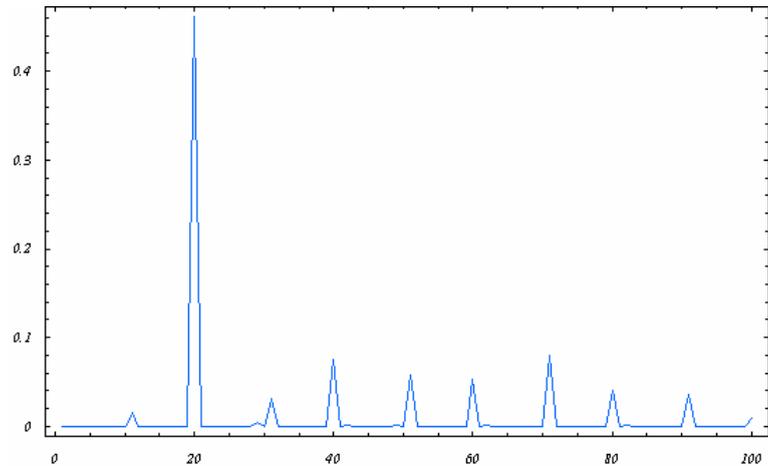


Gráfico 2.6
Valores del parámetro $|\rho_n|^2$ para la secuencia simulada.

En nuestro caso, la observación del *Gráfico 2.4* de evolución de los valores de la ganancia del filtro/error del modelo descubre el comportamiento de los distintos coeficientes de reflexión. Como ya se ha mencionado entre las observaciones de las expresiones del algoritmo de Levinson en dicha sección, valores constantes de E_n implican el valor nulo para el coeficiente ρ_n . Y al contrario, una alteración de valores en el paso de E_n a E_{n+1} , implica un valor no nulo de ρ_{n+1} ¹.

El *Gráfico 2.6* muestra los valores del módulo al cuadrado del coeficiente de reflexión en la modelización de la secuencia de datos que se ha simulado en la sección anterior. Volvemos a comprobar que en los índices 20, 40 y 51 el parámetro tiene un valor no nulo.

¹ RECUÉRDASE QUE $E_n = (1 - |\rho_n|^2)E_{n-1}$

Por lo tanto el modelo lineal interpreta el dominio de la frecuencia como un medio estratificado en el que cada capa representa un índice de frecuencia cuya densidad varía dependiendo de la relevancia de su peso. Un peso mayor se traduciría en una capa muy densa, mientras que índices de frecuencias no relevantes darían lugar a capas con baja densidad.

Mostramos en el *Gráfico 2.7* de forma esquemática el medio multicapa que construiría el modelo del filtro aplicado a la secuencia de datos simulados de la sección anterior, usando un color más oscuro para indicar mayor densidad.

Queremos referir en esta última sección varias notas que creemos son importantes. La primera de ellas es referente al uso del algoritmo de Levinson para secuencias de números complejos. Es decir, si su uso clásico estaba destinado a la resolución de la aproximación lineal de realización de procesos estacionarios reales, tanto los autoproductos escalares como el resto de su enunciado se han adaptado al uso más general.

Y en segundo lugar queremos hacer hincapié en la verificación de la conservación de las normas (*energías*) no sólo entre los dominios (*identidad de Parseval*), sino con la norma del modelo. Esto es,

$$|\{x\}|^2 = |\{X\}|^2 = 2\|\{X\}\|^2 - |X_0|^2,$$

usando la norma en l^2 .



Gráfico 2.7

Como una consecuencia de la interpretación de los parámetros del modelo (nos referimos a los coeficientes de reflexión y no a los coeficientes de los polinomios predictores) que se ha descrito, se puede lograr acelerar el cálculo del mismo.

La aceleración del algoritmo

Trataremos en esta sección de predecir los posibles valores nulos de los coeficientes de reflexión que no tienen porqué ser calculados. Dicha predicción se establece a través de las magnitudes de las frecuencias no relevantes.

Si asumimos un valor nulo para los coeficientes de reflexión de los índices de las frecuencias no relevantes, podemos ahorrarnos su cálculo y por tanto el uso de la expresión iterativa. Ésta última se va a ver substituida en el algoritmo clásico de Levinson por asignaciones más sencillas. Se trata del siguiente cambio,

$$\rho_n = \frac{A_{n-1}(r_n \ r_{n-1} \ \dots \ r_2 \ r_1)}{E_{n-1}}$$

$$A_n = (A_{n-1} \ 0) - r_n(0 \ \tilde{A}_{n-1}) \text{ siendo } \tilde{A}_j = (\overline{a_j} \ \overline{a_{j-1}} \ \dots \ \overline{a_1} \ 1)$$

$$E_n = (1 - |\rho_n|^2) E_{n-1}$$

por la asignación directa,

$$\rho_n = 0$$

$$A_n = (A_{n-1} \ 0)$$

$$E_n = E_{n-1}$$

En ese caso, tras un estudio de las magnitudes de las frecuencias, se deberá construir un subconjunto con los índices representativos de las mismas. Dicho conjunto, que denotaremos por S también

puede ser construido mediante una regla de corte como por ejemplo, $n \in S \Leftrightarrow |X_n| > 0.05$ o similar. Nótese que, en realidad, la complejidad del algoritmo dependerá de la cardinalidad del conjunto de índices relevantes.

El algoritmo acelerado queda escrito finalmente de la siguiente manera,

$$E_0 = r_0;$$

$$A_0 = (1);$$

Para $n = 1$ hasta $\max\{S\}$ hacer

Si $E_{n-1} \neq 0$ y $n \in S$ entonces

$$\rho_n = \frac{A_{n-1}(r_n \ r_{n-1} \ \dots \ r_2 \ r_1)}{E_{n-1}}$$

$$A_n = (A_{n-1} \ 0) - \rho_n (0 \ \tilde{A}_{n-1}) \text{ siendo } \tilde{A}_j = (\overline{a_j} \ \overline{a_{j-1}} \ \dots \ \overline{a_1} \ 1)$$

$$E_n = (1 - |\rho_n|^2) E_{n-1}$$

si $n \notin S$ entonces

$$\rho_n = 0$$

$$A_n = (A_{n-1} \ 0)$$

$$E_n = E_{n-1}$$

Algoritmo acelerado de Levinson

En el *Gráfico 2.8* se muestra la acción del algoritmo acelerado sobre la serie que hemos simulado. No se observan diferencias sustanciales entre los resultados de ambos algoritmos, y sin embargo sólo ha hecho falta calcular 5 coeficientes frente a los 51 del método clásico.

La forma cerrada del modelo se simplifica a la forma siguiente,

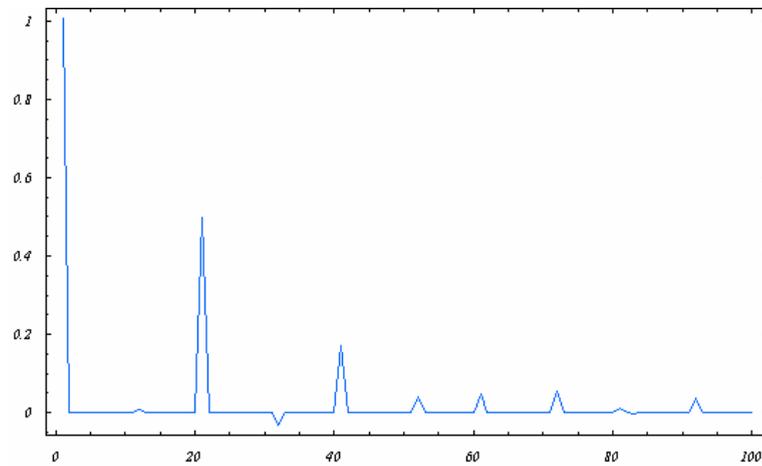


Gráfico 2.8
Secuencia simulada modelizada por el algoritmo acelerado.

$$\tilde{X}(z) = \frac{1.00574}{1 - 0.0057z^{11} - 0.4962z^{20} + 0.0374z^{31} + 0.0752z^{40} - 0.0754z^{51}}$$

en la que la mayoría de los coeficientes del polinomio predictor del denominador se han anulado completamente.

Debido a la relación ya expuesta entre el modelo del filtro y la aproximación de *Padé-Laurent*, hemos creído conveniente introducir en el siguiente capítulo, un modelo basado en dicha aproximación para la secuencia *DFT* de datos reales. Ello nos permitirá ofrecer en el último capítulo resultados comparativos para una aplicación real.

Capítulo 3 *Un modelo de Padé-Laurent aplicado a la secuencia DFT de datos reales*

En este capítulo retomamos la perspectiva de la modelización lineal a través del modelo del filtro como una caso particular de aproximación racional de series biinfinitas utilizando *aproximantes de Padé-Laurent*. Si este resultado fue ya comentado en el capítulo anterior, estudiamos, tras unos preliminares algebraicos útiles para enfrentarnos a las series biinfinitas o formales, la generalización a estas series de los *aproximantes clásicos de Padé*, [BakGra95].

El planteamiento de un nuevo modelo nace de la misma similitud, trabajada en el capítulo anterior, entre la secuencia de pesos de la *DFT* de datos reales y sus autoprodutos escalares. Sin embargo, la idea tiene su origen en la lectura de las cuestiones abiertas de la Tesis Doctoral de M. Jiménez, [Jim91], en las que se plantea la falta en la literatura de aproximaciones racionales en el dominio de la frecuencia. En este caso, nos centramos en usar los aproximantes de *Padé-Laurent* directamente sobre la secuencia de pesos de la *DFT* y obtener de este modo un ajuste equivalente, e incluso más sencillo al ser trasladado al dominio del tiempo.

Construimos, por tanto, un modelo alternativo basado en un *aproximante de Padé-Laurent* de la serie *DFT* de una secuencia de datos reales usando para ello el algoritmo de Levinson y sus propiedades en la construcción del modelo del filtro.

3.1. Preliminares: álgebra de series de Laurent¹

Dada una secuencia biinfinita de números complejos, $\{f_k\}_{k=-\infty}^{\infty}$, le asociamos la serie formal de Laurent, (*sfl*, de ahora en adelante), de la forma siguiente,

$$F(z) = \sum_{k=-\infty}^{\infty} f_k z^k.$$

Sobre las *sfl* definimos el operador proyección como,

$$\prod_{m:n} F(z) = \sum_{k=m}^n f_k z^k,$$

siendo, en general, m y n enteros positivos tales que $m \leq n$. Se generaliza de forma natural esta definición permitiendo también que m represente $-\infty$ y n sea ∞ con lo que las subseries así construidas pueden poseer un infinito número de términos.

Para valores finitos de estos índices, denominamos *polinomios de Laurent* al resultado de la aplicación del operador proyección. Así podemos denotar un polinomio de Laurent como $f(z) = f_m z^m + f_{m+1} z^{m+1} + \dots + f_{n-1} z^{n-1} + f_n z^n$, siendo su grado el entero positivo $\max\{|k| / f_k \neq 0\}$.

La necesidad de trabajar con subramas infinitas de *sfl* exige una extensión de la definición usual de infinitésimo. Así decimos que,

¹ LA NOTACIÓN QUE SEGUIMOS ES, BÁSICAMENTE, LA DEL LIBRO DE A. BULTHEEL, [BUL87].

$$F(z) \in O_+(z^m) \text{ sii } \prod_{-\infty:m-1} F(z) = 0$$

$$F(z) \in O_-(z^n) \text{ sii } \prod_{n:\infty} F(z) = 0.$$

Usando esta notación, las series formales de potencias serían las del conjunto $O_+(z^0)$.

Puesto que trivialmente $O_+(z^m) \subset O_+(z^{m-1})$ y $O_+(z^n) \subset O_+(z^{n+1})$ para valores enteros finitos cualesquiera m y n , se justifica hablar de dos órdenes para las *sfl*. Definimos,

$$\delta^+ F(z) = \sup\{m / F(z) \in O_+(z^m)\}$$

$$\delta^- F(z) = \inf\{n / F(z) \in O_-(z^n)\}.$$

Como propiedad deducible de esta definición podemos indicar que los *polinomios de Laurent* poseen los órdenes δ^+ y δ^- finitos. El poseer al menos uno de ellos no finito deduce la existencia de una subsecuencia infinita de coeficientes no nulos.

Debido a la naturaleza biinfinita de las *sfl*, su álgebra operacional debe tratarse con cuidado. La suma de *sfl* no ofrece dificultad si se define de forma natural atendiendo al grado de los monomios.

El producto, tratado como una generalización de la forma usual entre polinomios, no ofrece tampoco mayor dificultad en los casos en los que los factores tienen ambos δ^+ finitos, ambo δ^- finitos, o uno de ellos es un polinomio de Laurent. Estos casos son suficientes para el desarrollo que nos ocupa.

El tratamiento del cociente está ligado al producto si estudiamos la inversa de una *sfl*. El siguiente resultado para series con δ^+ finito o δ^- finito resuelve de forma constructiva la existencia y unicidad de la *sfl* inversa.

Proposición 3.1 Si $F(z)$ es una *sfl* con orden $\delta^+(F(z))=m$ finito, entonces existe una única *sfl* $G(z)$ con $\delta^+(G(z))=-m$ tal que $F(z)G(z)=1$.

Para la demostración basta con usar en el cálculo de $G(z)$ el algoritmo de Euclides para la división de polinomios repetido reiteradamente. Además, una proposición análoga deduce un resultado equivalente cuando se trata de una *sfl* con orden δ^- finito.

Dada una *sfl* $F(z)$ con infinitos términos no nulos y grados δ^+ ó δ^- finitos, su inversa se define como resultado de las proposiciones anteriores como $F^{-1}(z)=G(z)$.

Si $F(z)$ es un polinomio de Laurent, la división reiterada puede construir su inversa como una *sfl* de grado δ^+ ó δ^- finito según se aplique la expansión en potencias de z ó de z^{-1} .

En ese caso denotaremos,

$L_+(F^{-1}(z))$ a la expansión en potencias de z ,

$L_-(F^{-1}(z))$ a la expansión en potencias de z^{-1} .

Concluimos entonces definiendo el cociente de dos *sfl* $F(z)$ y $G(z)$ como el producto $F(z)G^{-1}(z)$ en los casos en los que éste sea posible.

Para aquellas ocasiones en las que sean ambos *polinomios de Laurent*, construimos las expansiones usando las siguientes igualdades,

$$L_+(F(z)/G(z))=F(z)L_+(G^{-1}(z))$$

$$L_-(F(z)/G(z))=F(z)L_-(G^{-1}(z)).$$

3.2. Aproximantes de Padé

Dada la serie de potencias $F(z)$, y los polinomios $A_m(z)$ y $B_n(z)$ de grado m y n respectivamente, decimos que el cociente $A_m(z)/B_n(z)$ es un *aproximante de Padé de tipo (m/n)* si se verifica,

$$F(z) - L_+(A_m(z)/B_n(z)) \in O_+(m+n+1) \text{ con } B_n(0) \neq 0, \text{ Ecuación 3.1.}$$

Sin embargo, para su generalización a *sfl* necesitamos de la siguiente ecuación, en la que se obvia el cociente de polinomios,

$$F(z)B_n(z) - A_m(z) \in O_+(m+n+1) \text{ con } B_n(0) \neq 0, \text{ Ecuación 3.2.}$$

Comprobamos que la exigencia de que $B_n(0) \neq 0$ implica que el término independiente del polinomio $B_n(z)$ es no nulo. De este modo, al multiplicar la expresión de la *Ecuación 3.1* por el polinomio en cuestión obtenemos la expresión de la *Ecuación 3.2* al no alterarse el grado de los polinomios. El camino inverso lo logramos al multiplicar *3.2* por la expansión $L_+(1/B_n(z)) = L_+(B_n^{-1}(z))$. Por lo tanto, ambas ecuaciones son equivalentes.

Podemos deducir de la *Ecuación 3.2* que, si $F(z)$ es una *sfl* y $B_n(z)$ un polinomio, no puede existir, en general, polinomio $A_m(z)$ que verifique tal propiedad. Para que esto se pueda lograr, $A_m(z)$ ha de ser también una *sfl* tal que $A_m(z) \in O_-(z^m)$. En este escenario vamos a deducir las ecuaciones que calculan los coeficientes de $A_m(z)$ y $B_n(z)$.

Cálculo y propiedades de los coeficientes del aproximante

Dada la *sfl* $F(z)$, trabajaremos con la matriz *Toeplitz* que tienen asociadas las *sfl*, $T_n^{(m)} = (f_{m+i-j})_{i,j=0}^n$. Esto es,

$$T_n^{(m)} = \begin{pmatrix} f_m & f_{m-1} & f_{m-2} & \cdots & f_{m-n} \\ f_{m+1} & f_m & f_{m-1} & \cdots & f_{m-n+1} \\ f_{m+2} & f_{m+1} & f_m & \cdots & f_{m-n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{m+n} & f_{m+n-1} & f_{m+n-2} & \cdots & f_m \end{pmatrix}.$$

Decimos que una *sfl* $F(z)$ es *normal* si su matriz *Toeplitz* $T_n^{(m)}$ es no singular para valores arbitrarios de sus parámetros m y n . La siguiente proposición justifica esta exigencia.

Proposición 3.2 Si $F(z)$ es una *sfl normal*, todos los *aproximantes* $A_m(z)/B_n(z)$ de tipo (m/n) existen, siendo los órdenes $\delta^-(A_m(z)) = m$, $\delta^+(B_n(z)) = n$ y el orden del residuo $\delta^+(R(z)) = \delta^+(F(z)B_n(z) - A_m(z)) = m+n+1$. Además el par $(A_m(z)|B_n(z))$ es único salvo constante multiplicativa.

La comprobación de esta proposición pone de manifiesto la necesidad de la propiedad de *normalidad*. Denotaremos a los polinomios del aproximante como, $A_m(z) = \sum_{k=-\infty}^m a_k z^k$ y $B_n(z) = \sum_{k=0}^n b_k z^k$.

Igualando coeficientes según la *Ecuación 3.2* hasta el grado $m+n+1$ obtenemos los siguientes 3 bloques de ecuaciones, en donde denotamos por r_i al i -ésimo coeficiente del residuo $R(z) = \sum_{k=m+n+1}^{\infty} r_k z^k$.

$$\forall i = m, m-1, m-2, \dots \quad \sum_{k=0}^n f_{i-k} b_k = a_i \quad \text{Bloque (a)}$$

$$\forall i = m+1, m+2, \dots, m+n \quad \sum_{k=0}^n f_{i-k} b_k = 0 \quad \text{Bloque (b)}$$

$$\forall i = m+n+1, m+n+2, \dots, \quad \sum_{k=0}^n f_{i-k} b_k = r_i \quad \text{Bloque (r)}$$

Ecuaciones 3.3

La estrategia de resolución sería la siguiente: el bloque (b) es el punto de partida para el cálculo de los n coeficientes de $B_n(z)$ ($b_0 = 1$, con lo que no es coeficiente a calcular). Con dichos valores y las ecuaciones del bloque (a) se calculan los coeficientes del polinomio $A_m(z)$ y con las del bloque (r) se calculan los coeficientes del residuo. En ese caso, fijémonos inicialmente en el bloque (b),

$$\sum_{k=0}^n f_{i-k} b_k = 0, \text{ para valores del índice } i = m+1, m+2, \dots, m+n.$$

$$\sum_{k=0}^n f_{m+i-k} b_k = 0, \text{ donde, tras el cambio de índice, } i = 1, 2, \dots, n.$$

$$\sum_{k=1}^n f_{m+i-k} b_k = -f_{m+i}, \text{ también para } i = 1, 2, \dots, n, \text{ ya que el coeficiente}$$

b_0 se normaliza al valor 1.

Por lo tanto, hemos reducido el bloque (b) a un sistema lineal que tiene a $T_{n-1}^{(m)}$ por matriz de coeficientes y, dada su no singularidad, es posible entonces obtener una solución única para el conjunto de coeficientes $\{b_1, b_2, \dots, b_n\}$.

Para comprobar que el grado de $B_n(z)$ es n , esto es, que $b_n \neq 0$ supondremos por reducción al absurdo lo contrario. En ese caso, su valor nulo se obtendría del numerador de un cociente de determinantes siguiendo el clásico método de resolución de *Cramer*. Así dicho determinante tendría la forma,

$$\begin{vmatrix} f_m & f_{m-1} & f_{m-2} & \dots & f_{m-n+2} & -f_{m+1} \\ f_{m+1} & f_m & f_{m-1} & \dots & f_{m-n+3} & -f_{m+2} \\ f_{m+2} & f_{m+1} & f_m & \dots & f_{m-n+4} & -f_{m+3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_{m+n-1} & f_{m+n-2} & f_{m+n-3} & \dots & f_{m+1} & -f_{m+n} \end{vmatrix} = \begin{vmatrix} f_{m+1} & f_m & f_{m-1} & f_{m-2} & \dots & f_{m-n+2} \\ f_{m+2} & f_{m+1} & f_m & f_{m-1} & \dots & f_{m-n+3} \\ f_{m+3} & f_{m+2} & f_{m+1} & f_m & \dots & f_{m-n+4} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_{m+n} & f_{m+n-1} & f_{m+n-2} & f_{m+n-3} & \dots & f_{m+1} \end{vmatrix} = \left| T_{n-1}^{(m+1)} \right|,$$

que no puede ser nulo por la propiedad de normalidad de la *sfl*.

De igual forma se comprueba que $d^-(A_m(z)) = m$, esto es, $a_m \neq 0$. Para ello escogemos de los bloques de la *Ecuación 3.3* las siguientes igualdades,

$$\sum_{k=0}^n f_{m-k} b_k = a_m, \text{ la primera ecuación del bloque (a)}$$

$$\sum_{k=0}^n f_{m+i-k} b_k = 0, \text{ para } i=1,2,\dots,n, \text{ las } n \text{ ecuaciones del bloque (b).}$$

y construimos el sistema lineal de expresión matricial, $T_n^{(m)}(b_0 \ b_1 \ b_2 \ \dots \ b_n)^t = (a_m \ 0 \ 0 \ \dots \ 0)^t$. Si suponemos que $a_m = 0$, el anterior sistema es homogéneo y, al ser su matriz regular, solo existe una solución: $b_k = 0$ para $k=0,1,\dots,n$, lo cual es absurdo pues ya hemos comprobado que $b_n \neq 0$. Por lo tanto, también $a_m \neq 0$.

Finalmente, y para comprobar que el $d^+(R(z))$ es $m+n+1$, esto es, $r_{m+n+1} \neq 0$, escogemos de forma similar las siguientes $n+1$ ecuaciones,

$$\sum_{k=0}^n f_{m+i-k} b_k = 0, \text{ para } i=1,2,\dots,n, \text{ las } n \text{ ecuaciones del bloque (b).}$$

$$\sum_{k=0}^n f_{m+n+1-k} b_k = r_{m+n+1}, \text{ la primera ecuación del bloque (r).}$$

que construyen el sistema lineal, $T_n^{(m+1)}(b_0 \ b_1 \ b_2 \ \dots \ b_n)^t = (0 \ 0 \ 0 \ \dots \ r_{m+n+1})$. Si por reducción al absurdo consideramos que $r_{m+n+1} = 0$, el sistema homogéneo resultante solo posee la solución trivial, con lo que $b_n = 0$, lo cual no es cierto. Por lo tanto, $r_{m+n+1} \neq 0$.

Esta generalización aparece en [Bul87] y es la que más fácilmente se traslada de una definición clásica de *aproximantes de Padé* (Ecuación 3.1). En la práctica conlleva, sin embargo, alguna dificultad pues uno de los polinomios que intervienen, $A_m(z)$, es, en realidad, una *sfl* cuyos infinitos coeficientes no se calculan de forma cerrada sino construyendo reiteradamente el sistema lineal adecuado. Este obstáculo práctico se salva en la propuesta de definición para estos aproximantes de los trabajos [Gil95] y [GilGon97] que consideran, en una segunda fase, *la aproximación clásica de Padé* para el polinomio $A_m(z)$.

Existen otras maneras de aproximar racionalmente las *sfl* tomando también como punto de partida los aproximantes de Padé. En particular, la forma más interesante de ellas para nuestros objetivos, se presenta en la sección siguiente.

3.3. Aproximantes de Padé-Laurent

Para la construcción de los aproximantes de Padé-Laurent de una *sfl*, $F(z) = \sum_{k=-\infty}^{\infty} f_k z^k$, se construyen las siguientes subseries o ramas,

$$Z^+(z) = \frac{1}{2}f_0 + \sum_{k=1}^{\infty} f_k z^k$$

$$Z^-(z) = \frac{1}{2}f_0 + \sum_{k=1}^{\infty} f_{-k} z^{-k},$$

tales que, trivialmente, $F(z) = Z^-(z) + Z^+(z)$.

A partir de éstas, dados los enteros positivos m y n , calcularemos los cocientes de polinomios de Laurent, $P^+(z)/Q^+(z)$ y $P^-(z)/Q^-(z)$ verificando las siguientes propiedades,

- $Q^+(z)$ es un polinomio en z y $Q^-(z)$ es un polinomio en z^{-1} , ambos, como mucho, de grado n .
- $Z^+(z) - L_+(P^+(z)/Q^+(z)) \in O_+(z^{m+n+1})$.
- $Z^-(z) - L_-(P^-(z)/Q^-(z)) \in O_-(z^{-(m+n+1)})$.
- La expresión $\frac{P^+(z)}{Q^+(z)} + \frac{P^-(z)}{Q^-(z)}$ se deriva hacia un cociente de polinomios de Laurent, $\frac{A_m(z)}{B_n(z)}$, de grados, como mucho, m y n respectivamente.

Denominaremos en ese caso al par $(P^+(z)/Q^+(z), P^-(z)/Q^-(z))$ un aproximante de Padé-Laurent de la serie *sfl* $F(z)$. Nótese que la razón de definir tal aproximante como un par de cocientes es que con un simple cálculo a partir de éstos se puede obtener el cociente $A_m(z)/B_n(z)$ mientras que la dirección contraria no es única.

Así definido obtenemos la suma de dos expansiones independientes, $L_+(P^+(z)/Q^+(z)) + L_-(P^-(z)/Q^-(z))$, que construyen una *sfl* cuyos coeficientes son idénticos a los de la *sfl* de partida $F(z)$ para todo índice $|k| \leq m+n$, obteniendo una forma cerrada de la expresión al construir el cociente $A_m(z)/B_n(z)$.

La propiedad de no singularidad de las *sfl* normales también es suficiente en este caso para asegurar la existencia y la unicidad salvo constante multiplicativa del aproximante de *Padé-Laurent* para valores arbitrarios de m y n . La comprobación de este hecho es similar a la estrategia usada en el caso anterior.

3.4. El modelo del filtro y los aproximantes de *Padé-Laurent*

Ya se ha mencionado en la *Sección 2.2*, la relación entre el modelo del filtro y los aproximantes de *Padé-Laurent* de la serie de autoproductos escalares, que denotábamos por,

$$R(z) = \sum_l z^l r_l.$$

Dada la expresión en forma cerrada del filtro lineal, los autoproductos escalares de la secuencia del modelo atendían a la forma de cociente de polinomios de Laurent,

$$\hat{R}(z) = \sum_l z^l \hat{r}_l = \frac{G_p^2}{A_p(z) \tilde{A}_p(z)}.$$

Dado que $R(z)$ y $\hat{R}(z) = G_p^2 / A_p(z) \tilde{A}_p(z)$ poseen coincidentes valores para los índices consecutivos desde $l=-p$ hasta $l=p$, $\hat{R}(z)$ es

un aproximante de *Padé-Laurent*[0/p] según la definición vista en la sección anterior.

Por esta razón se puede usar el algoritmo de Levinson para la construcción de aproximantes de *Padé-Laurent*[0/p] para un determinado tipo de series biinfinitas, $\sum_k f_k z^k$: aquellas que verifiquen que $f_{-k} = \overline{f_k}$ y que se denominan *hermíticas*.

Nótese que esta propiedad es verificada tanto por la secuencia de autoproductos escalares como por la secuencia de la *DFT* de datos reales. *Usaremos esto último para construir aproximantes de Padé-Laurent de órdenes [0/p] mediante el algoritmo de Levinson al que únicamente ha de variarse la entrada y la especificación de salida.* La generalización de este algoritmo para la construcción de aproximantes de *Padé-Laurent* de órdenes arbitrarios se encuentra en [Bul87].

El algoritmo adaptado a la DFT de datos reales

La modificación del algoritmo para su nuevo objetivo no va a variar ninguna de las operaciones internas del mismo, lo que cambiarán serán la entrada y la salida del mismo.

Si bien cuando construíamos el filtro lineal teníamos como entrada la secuencia de autoproductos y el orden del polinomio, y dábamos como salida la ganancia del filtro y los coeficientes del polinomio predictor, en este caso la entrada será la secuencia de pesos de la *DFT* de datos reales y el orden del aproximante, y la salida será el aproximante pedido. El algoritmo queda de la siguiente forma,

Entrada $\{X_0, X_1, \dots, X_p, p\}$

$$E_0 = r_0;$$

$$A_0 = (1);$$

Para $n=1$ *hasta* p *hacer*

Si $E_{n-1} \neq 0$ *entonces*

$$\rho_n = \frac{A_{n-1}(X_n \ X_{n-1} \ \dots \ X_2 \ X_1)}{E_{n-1}}$$

$$A_n = (A_{n-1} \ 0) - \rho_n (0 \ \tilde{A}_{n-1}) \text{ siendo } \tilde{A}_j = (\overline{a_j} \ \overline{a_{j-1}} \ \dots \ \overline{a_1} \ 1)$$

$$E_n = (1 - |\rho_n|^2) E_{n-1}$$

$$A_p(z) = (1 \ a_1 \ a_2 \ \dots \ a_p)(1 \ z \ z^2 \ \dots \ z^p)^t$$

$$G_p = \sqrt{E_p}$$

$$\text{Salida } \frac{G_p^2}{A_p(z)\tilde{A}_p(z)}$$

Algoritmo de Levinson para un aproximante de Padé-Laurent[0|p] de la serie de pesos de la DFT de datos reales

Ventajas e inconvenientes

La forma cerrada que describe el algoritmo anterior puede sustituir con ventaja a los cálculos de la transformada discreta inversa de Fourier,

$$x_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} X_r W_N^{rk}, \text{ (Ecuación IDFT, capítulo 1).}$$

Esto es, si dada la secuencia de pesos $\{X_0, X_1, \dots, X_{N-1}\}$, tenemos una aproximación racional de la serie,

$$\sum_{r=0}^{N-1} X_r z^r,$$

mediante,

$$\frac{E_p}{A_p(z)\tilde{A}_p(z)},$$

en las sucesivas sustituciones de la variable z por las potencias W_N obtendremos una aproximación de la secuencia de datos reales que genera la *DFT*. Por esta razón, la comparativa de errores se realiza en el dominio del tiempo y no en el de la frecuencia.

Por otro lado, ya fue definida $\tilde{A}_p(z) = \overline{A_p\left(\frac{1}{z}\right)}$. En ese caso, al tomar valores la variable z en las sucesivas potencias de $W_N = e^{\frac{2\pi}{N}}$ tenemos que,

$$\tilde{A}_p(W_N^k) = \overline{A_p\left(\frac{1}{W_N^k}\right)} = \overline{A_p\left(W_N^{-k}\right)} = \overline{A_p(W_N^k)},$$

con lo que, $A_p(W_N^k)\tilde{A}_p(W_N^k) = |A_p(W_N^k)|^2$. En ese caso, el cociente tiene como expresión final,

$$\frac{G_p^2}{|A_p(W_N^k)|^2}.$$

Nótese además que en el denominador aparece el módulo del número complejo $A_p(W_N^k)$ cuyo cálculo obedece, como se vio en el

capítulo anterior, a una *FFT* aplicada sobre una secuencia construida con los coeficientes del polinomio predictor.

Una desventaja al hecho de usar directamente la secuencia en la construcción del aproximante y no la secuencia de autoprodutos correspondientes es que no es aplicable el desarrollo lineal que ha ocupado buena parte de esta memoria. Para resolver temas como la determinación del grado de polinomio predictor debemos, por tanto, circunscribirnos a la aproximación racional.

En este sentido, y como se verá en la aplicación práctica, la elección del orden p del aproximante vendrá determinada por *el índice de la más alta frecuencia que, por relevante, queramos incorporar al modelo*. Esto es, si son coincidentes secuencia y aproximante hasta el coeficiente de índice p , p será el índice del peso de más alta frecuencia que queramos conservar en el modelo. Omitimos una parte final con una simulación de datos puesto que no añade ningún resultado ni interpretación relevante.

En el siguiente capítulo se desarrolla una aplicación de los modelos presentados sobre series que cuantifican el consumo de agua potable en la ciudad de Santa Cruz de Tenerife.

Capítulo 4 *Una aplicación práctica: series de consumo de agua en la ciudad de Santa Cruz de Tenerife*

El presente capítulo centra sus objetivos en el análisis y modelización de las series de datos de consumo de agua en la ciudad de Santa Cruz de Tenerife. Las características periódicas de su comportamiento justifican el empleo del análisis de Fourier y la modelización en el dominio de la frecuencia.

Son dos las partes en las que hemos dividido el capítulo. En una primera parte efectuamos una revisión de las frecuencias que de forma más relevante forman parte del comportamiento global de cada serie. Este primer estudio nos permitirá, por un lado, clasificar en cierta medida los depósitos según la aparición o no de ciertas periodicidades, y de otro, conocer el número de parámetros necesarios para la definición del modelo.

La segunda parte atiende a la construcción de modelos periódicos para cada una de las series usando las dos metodologías expuestas en los capítulos 2 y 3 que mejoran, mediante la consideración de una forma polinómica cerrada, las modelizaciones basadas en series de truncadas de Fourier.

4.1. Las series de datos

Las series de datos objeto de nuestro análisis y modelización son las obtenidas desde los distintos depósitos que la empresa responsable de la distribución de agua en Santa Cruz de Tenerife, EMMASA, tiene localizados en distintos puntos de la ciudad.

El viaje del agua puede esquematizarse centrándolo en los depósitos: llega y se almacena en ellos procedente de diversos canales de suministro y sale de ellos a la red de agua potable. La acción de almacenamiento en estos lugares permite por tanto independizar el flujo de entrada del de salida dando por supuesta la verificación de la siguiente restricción: el agua que se almacena debe ser superior a una cota mínima que permita asegurar el abastecimiento, e inferior a una cota máxima que evite el desaprovechamiento.

Cada depósito es controlado electrónicamente desde la central mediante sensores que informan de incidentes en torno a su funcionamiento. Uno de ellos controla el caudal de salida del depósito, medido en litros por segundo, y calcula la media por hora de este valor para construir la serie con la que trabajamos. Las cantidades de nuestra secuencia corresponden al intervalo entre las 0 horas del día 11 de octubre de 1996 hasta las 9 horas del 11 de noviembre de 1996, es decir, 754 datos por depósito. Como veremos, las conclusiones sobre los pesos de las periodicidades variarán notablemente el tamaño de la serie a modelizar.

La relación entre la salida de agua de los depósitos y el consumo real no es, sin embargo, equivalente. Esto es debido al efecto de fugas o pérdidas que se producen lo largo de la red de

agua potable. No obstante, sí que existe una relación directamente proporcional entre el nivel de pérdidas y el consumo. El motivo es el siguiente: si sube el nivel de consumo, aumenta la presión del agua en la red y con él el nivel de fugas; al bajar ocurre lo contrario. Podría decirse entonces que las pérdidas pueden asumirse como un coste añadido y que, en ese caso, los patrones de comportamiento en la salida de agua de los depósitos son trasladables al consumo real. Es por lo que nos referiremos a las series obtenidas como series de consumo de agua.

Una segunda hipótesis de trabajo se refiere a la ausencia de tendencia en las series de datos. Esta suposición la basamos en dos razones:

- Las zonas urbanas en las que los depósitos dividen la ciudad tienen una población estabilizada, esto es, con crecimiento muy bajo.
- Aún cuando pudiera producirse un incremento poblacional en alguna zona, éste no sería apreciable en el intervalo de tiempo en el que se han obtenido los datos.

En definitiva, la demarcación de la ciudad en zonas atendidas por los depósitos eliminan componentes tendenciales al dividir también éstas. Este fenómeno no se produciría en series de consumo con una única central distribuidora, que acumularía incrementos de consumidores a niveles apreciables. Un ejemplo de esto último lo encontraríamos en series de consumo eléctrico.

No todos los depósitos tienen instalado los dispositivos de control de gestión electrónicos. Además, su dependencia del suministro eléctrico hace que, en numerosas ocasiones al faltar éste queden incompletas o falseadas secuencias de valores medios. Por este motivo, han sido apartados del estudio los depósitos de Plaza de

Toros y Sobradillo al acumular demasiados fallos que hacen irre recuperable la serie. En el *Gráfico 4.1* mostramos como ejemplo al serie de Plaza de Toros.

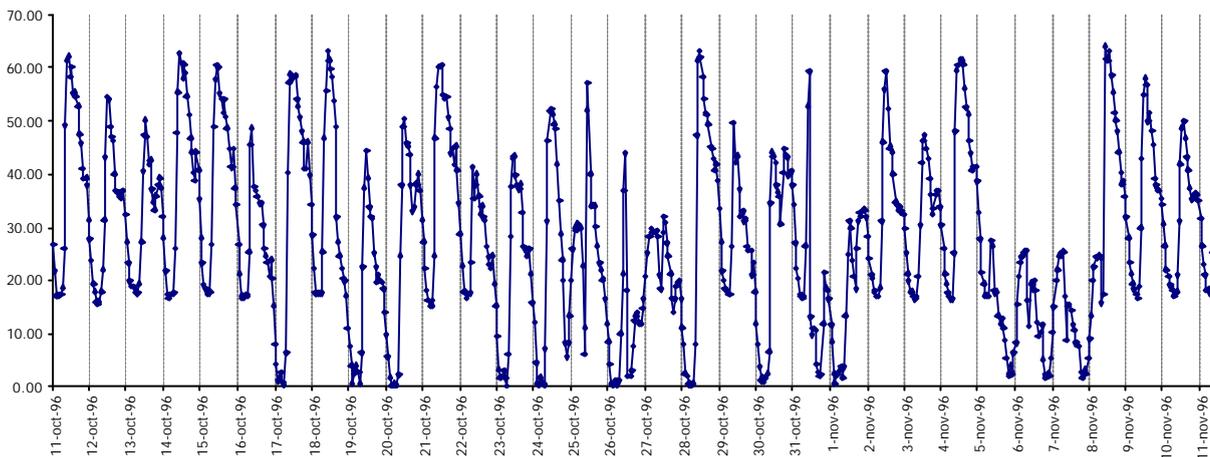


Gráfico 4.1. Serie de Consumo de Agua del depósito de Plaza de Toros

Los 10 depósitos que han centrado nuestro trabajo han sido, por orden alfabético: Alisios, Barranco Grande, Cueva Roja, Ofra, Polígono Ofra, Salamanca, Salud, San Andrés, Tío Pino y Vuelta Grande. Para cada uno de ellos, presentamos en la sección siguiente, los gráficos de sus transformadas y las tablas de las frecuencias más importantes.

4.2. Análisis de Periodicidades

Mostraremos en esta sección y para cada uno de los 10 depósitos, los gráficos correspondientes a la *DFT* de las series de consumo de agua. Recogeremos en el índice r de la secuencia resultante, el peso de la periodicidad N/r horas, siendo N el número

de datos, 754. Así, el peso del índice 31 mostrará la influencia de la periodicidad $754/31=24.32$ horas, esto es, la fuerza de la periodicidad diaria. Como veremos, esta periodicidad es la más fuerte y la que es común a todas las series. Con ello concluimos que el patrón o rutina diaria es el más importante.

Debido a que los datos de las secuencias son reales, su transformada discreta va a ser conjugada par, es decir, el peso del índice r es conjugado complejo del índice $N-r$. Por esa razón, los módulos de los pesos que son los mostrados en los gráficos son simétricos respecto del eje vertical, OY . Sólo se dibujarán por tanto, la secuencia de los $N/2$ primeros pesos.

Junto a los gráficos y también para cada uno de los depósitos, incluimos en forma de tabla un listado que muestra las periodicidades más relevantes ordenadas de mayor a menor influencia. Las columnas que aparecen en las tablas corresponden, respectivamente:

Índice de la frecuencia, r : que muestra un valor en el rango 0 a $N-1$ para referenciar a la senoide de periodo N/r .

Periodicidad en horas, p_r : donde se indica el periodo en horas según corresponda al índice.

Peso, w_r : que muestra para cada índice el peso o influencia de la senoide de periodo $754/r$ según la *DFT*.

Energía, e_r %: que calcula el tanto por ciento de energía para la periodicidad de índice r . Ésta se calcula sobre la energía total, atendiendo por tanto a la relación

$$e_r \% = \frac{2|w_r|^2}{\sum_{k=0}^{N-1} |w_k|^2} \cdot 100$$

Duplicamos el peso del índice para recoger también el valor en el índice $N-r$, de igual módulo, que también es una senoide de igual periodicidad.

Acumulado de Energía. $\sum e_r \%$: en el que recogemos la suma de energías aportadas al total hasta el índice r . Corresponde por tanto a la ecuación

$$\sum e_r \% = \frac{\sum_{k=0}^r 2|w_r|^2}{\sum_{k=0}^{N-1} |w_r|^2} \cdot 100.$$

Energía en componentes periódicas. $e_r \%$, y acumulado de energías periódicas. $\sum e_r \%$: que elimina el índice 0 , es decir, la componente constante, de los cálculos. Las relaciones por tanto son las siguientes:

$$\varepsilon_r \% = \frac{2|w_r|^2}{\sum_{k=1}^{N-1} |w_k|^2} \cdot 100 \qquad \sum \varepsilon_r \% = \frac{\sum_{k=1}^r 2|w_r|^2}{\sum_{k=1}^{N-1} |w_r|^2} \cdot 100.$$

Nuestro interés por mostrar las frecuencias más relevantes está dificultado por la indefinición de 'relevancia'. Los criterios que parecen más interesantes pasan por el establecimiento de cotas para las diferentes columnas de las tablas: energía o energía acumulada, con o sin el efecto de la constante. Ejemplos de dichas cotas serían, por ejemplo,

- a. listar hasta que $e_r \% < 0.1\%$,
- b. listar hasta que $\sum e_r \% > 95\%$,
- c. listas hasta que $\sum \varepsilon_r \% < 80\%$.

Sin embargo, cualquiera de estas cotas provoca un heterogéneo comportamiento en cuanto a la longitud de las tablas para los distintos depósitos. Por ello hemos preferido listar un número fijo de filas, 15, pudiéndose en este caso apreciarse, como varían los distintos valores para las columnas.

Indicamos con la siguiente leyenda los pesos correspondientes a las periodicidades diaria y semanal. La primera por ser la más relevante, y la segunda porque, como veremos, nos permitirá una cierta clasificación de los depósitos y una considerable simplificación en el conjunto de datos a modelizar.

<input checked="" type="checkbox"/>	Diaria
<input type="radio"/>	Semanal

Para cada depósito, indicamos también en la tabla el peso relevante de mayor índice sombreando su fila, y, en el gráfico posterior, los dos valores periódicos más relevantes, indicando su índice y su magnitud.

Depósito Alisios

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	443.94	86.24%	86.24%		
☑24.32	31	78.05	5.33%	91.58%	38.76%	38.76%
11.97	63	45.87	1.84%	93.42%	13.39%	52.15%
☑23.56	32	44.82	1.76%	95.18%	12.78%	64.93%
☑25.13	30	23.16	0.47%	95.65%	3.41%	68.35%
☑27.93	27	16.60	0.24%	95.89%	1.75%	70.10%
12.16	62	15.99	0.22%	96.11%	1.63%	71.72%
☑22.85	33	13.74	0.17%	96.28%	1.20%	72.93%
6.03	125	12.60	0.14%	96.41%	1.01%	73.94%
☑22.18	34	12.07	0.13%	96.54%	0.93%	74.86%
11.78	64	11.24	0.11%	96.65%	0.80%	75.67%
☑26.93	28	10.95	0.10%	96.76%	0.76%	76.43%
12.57	60	10.85	0.10%	96.86%	0.75%	77.18%
3.44	219	10.35	0.09%	96.95%	0.68%	77.86%
5.98	126	10.14	0.09%	97.04%	0.65%	78.51%

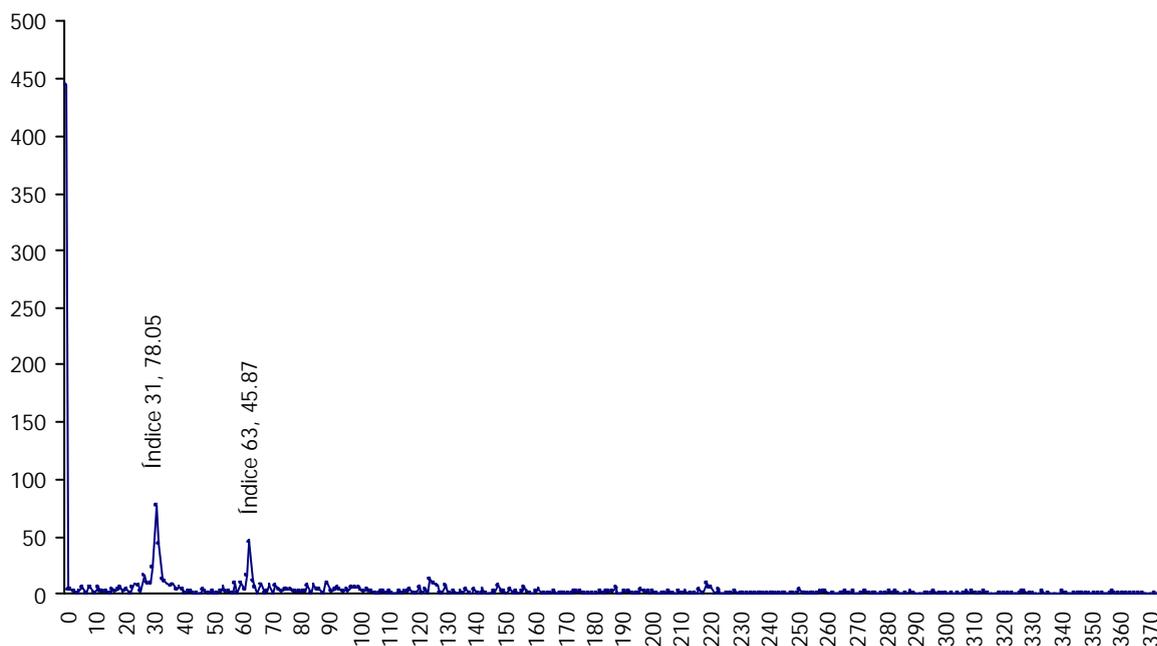


Gráfico 4.2. Transformada Discreta de Fourier de la serie del depósito Alisios.

Depósito Barranco Grande

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	63.08	83.58%	83.58%		
☑24.32	31	11.46	5.52%	89.10%	33.61%	33.61%
8.02	94	7.60	2.42%	91.53%	14.77%	48.38%
☑23.56	32	7.26	2.21%	93.74%	13.49%	61.87%
☑25.13	30	3.83	0.62%	94.36%	3.76%	65.63%
11.97	63	3.00	0.38%	94.74%	2.31%	67.94%
20.94	36	2.61	0.29%	95.02%	1.74%	69.68%
☑27.93	27	2.48	0.26%	95.28%	1.57%	71.26%
6.03	125	2.18	0.20%	95.48%	1.22%	72.48%
26.00	29	1.97	0.16%	95.65%	0.99%	73.47%
☑22.85	33	1.96	0.16%	95.81%	0.99%	74.46%
11.25	67	1.93	0.16%	95.96%	0.95%	75.41%
7.94	95	1.89	0.15%	96.11%	0.92%	76.32%
5.98	126	1.79	0.13%	96.25%	0.82%	77.14%
☑22.18	34	1.73	0.13%	96.37%	0.76%	77.91%

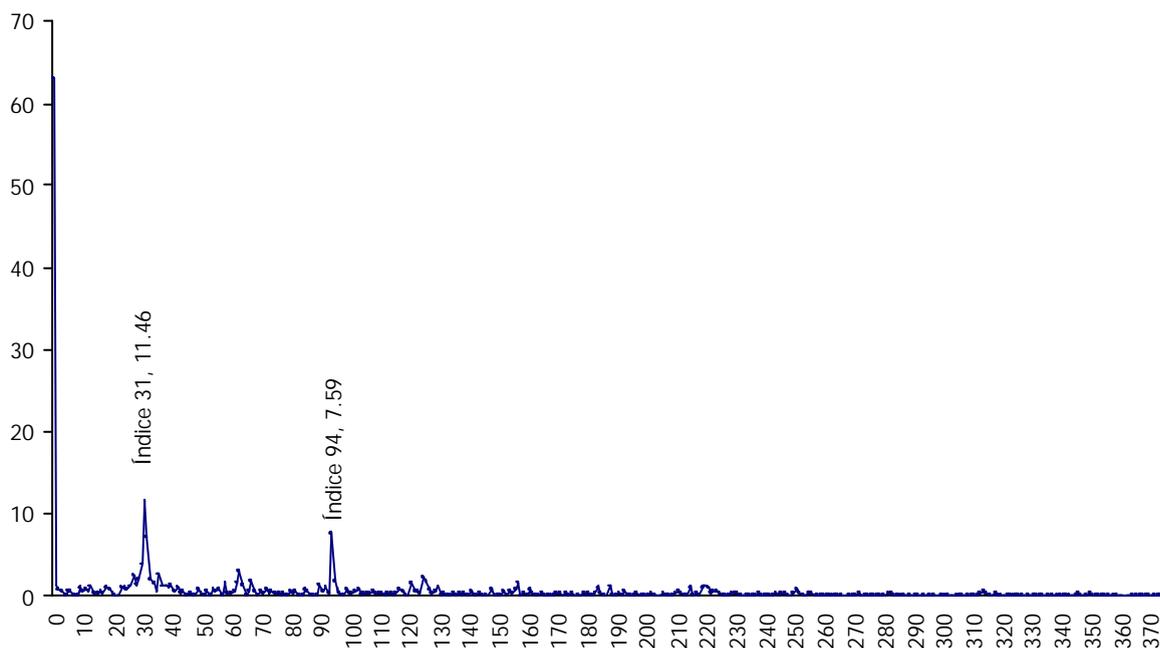


Gráfico 4.3. Transformada Discreta de Fourier de la serie del depósito Barranco Grande.

Depósito Cueva Roja

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0					
	0	2007.33	81.05%	81.05%		
☒24.32	31	444.97	7.97%	89.02%	42.03%	42.03%
11.97	63	278.21	3.11%	92.13%	16.43%	58.46%
☒23.56	32	255.39	2.62%	94.75%	13.85%	72.31%
12.16	62	110.03	0.49%	95.24%	2.57%	74.88%
☒22.85	33	109.39	0.48%	95.72%	2.54%	77.42%
☒25.13	30	107.83	0.47%	96.19%	2.47%	79.89%
☒26.00	29	79.56	0.25%	96.44%	1.34%	81.23%
11.78	64	72.98	0.21%	96.66%	1.13%	82.36%
6.03	125	66.48	0.18%	96.84%	0.94%	83.30%
⊙188.50	4	59.63	0.14%	96.98%	0.75%	84.06%
8.02	94	57.34	0.13%	97.11%	0.70%	84.76%
5.98	126	54.84	0.12%	97.23%	0.64%	85.39%
☒22.18	34	54.42	0.12%	97.35%	0.63%	86.02%
27.93	27	50.04	0.10%	97.45%	0.53%	86.55%

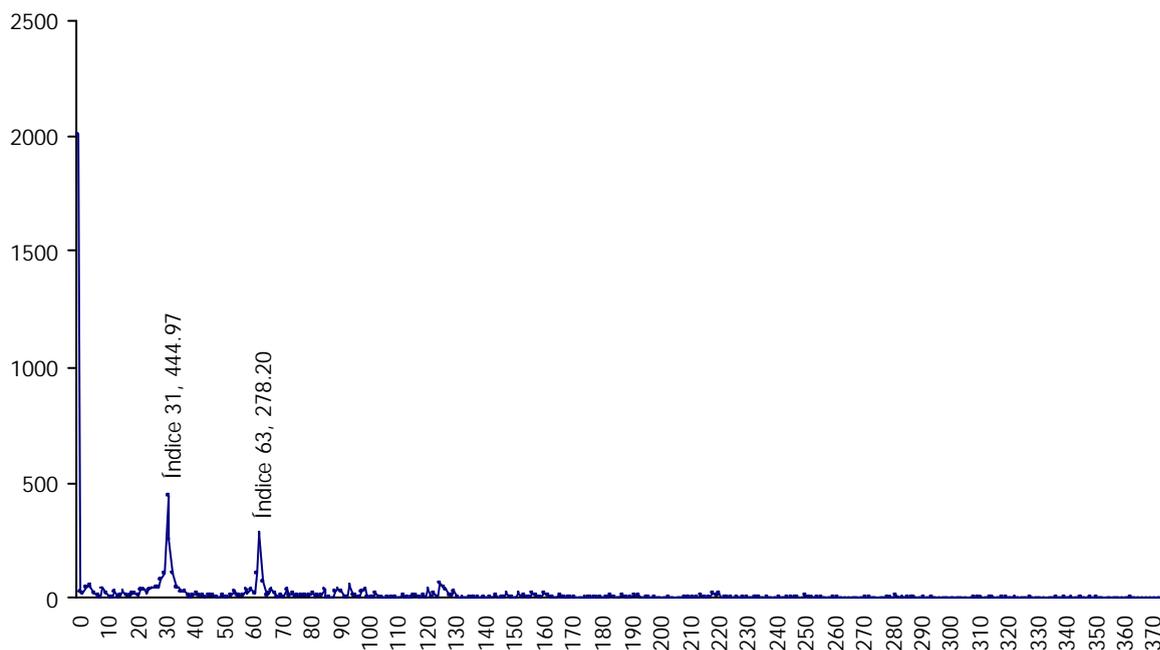


Gráfico 4.4. Transformada Discreta de Fourier de la serie del depósito Cueva Roja.

Depósito Ofra

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	1582.31	84.02%	84.02%		
☒25.13	30	334.39	7.50%	91.52%	46.95%	46.95%
☒24.32	31	197.30	2.61%	94.13%	16.34%	63.29%
12.36	61	173.99	2.03%	96.16%	12.71%	76.01%
☒26.00	29	83.06	0.46%	96.63%	2.90%	78.90%
☒23.56	32	76.10	0.39%	97.02%	2.43%	81.34%
12.57	60	66.14	0.29%	97.31%	1.84%	83.17%
26.93	28	61.42	0.25%	97.56%	1.58%	84.76%
☒22.85	33	47.51	0.15%	97.71%	0.95%	85.70%
⊙150.80	5	46.76	0.15%	97.86%	0.92%	86.62%
⊙251.33	3	43.11	0.12%	97.99%	0.78%	87.40%
⊙188.50	4	42.09	0.12%	98.11%	0.74%	88.15%
12.16	62	39.87	0.11%	98.21%	0.67%	88.81%
6.23	121	38.73	0.10%	98.31%	0.63%	89.44%
27.93	27	36.44	0.09%	98.40%	0.56%	90.00%

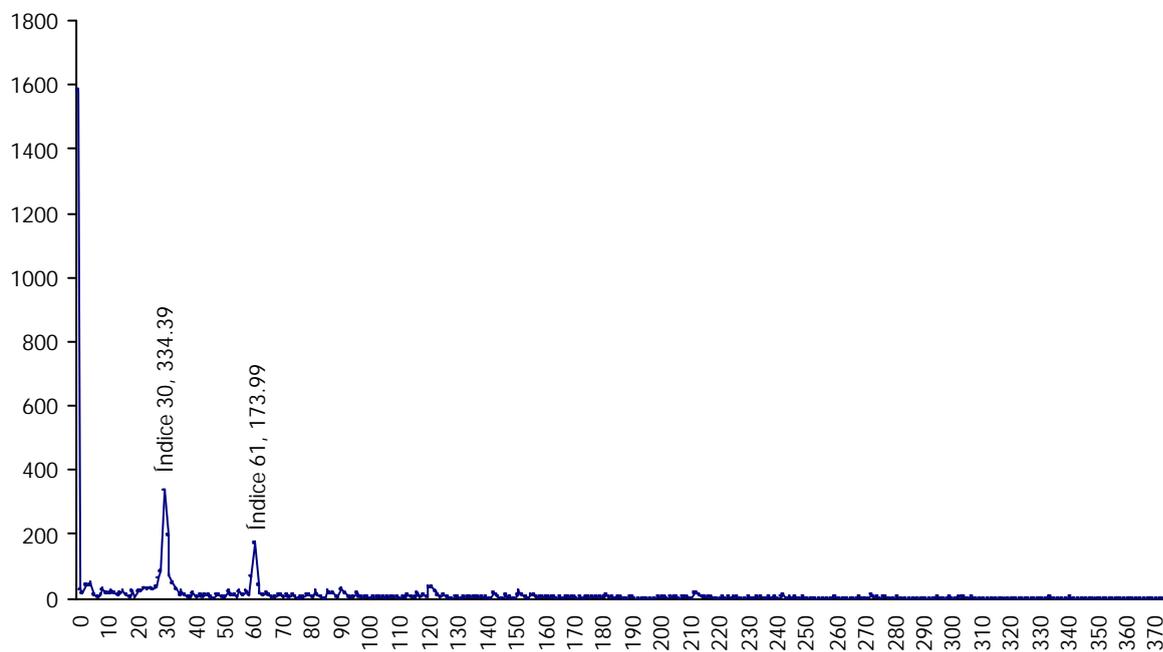


Gráfico 4.5. Transformada Discreta de Fourier de la serie del depósito Ofra.

Depósito Polígono Ofra

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	450.15	73.62%	73.62%		
☒24.32	31	125.22	11.39%	85.01%	43.18%	43.18%
☒23.56	32	70.18	3.58%	88.59%	13.56%	56.74%
11.97	63	68.33	3.39%	91.98%	12.86%	69.60%
☒25.13	30	35.63	0.92%	92.90%	3.50%	73.10%
☒22.85	33	28.42	0.59%	93.49%	2.22%	75.32%
☒26.00	29	26.74	0.52%	94.01%	1.97%	77.29%
12.16	62	26.49	0.51%	94.52%	1.93%	79.23%
☒22.18	34	19.37	0.27%	94.79%	1.03%	80.26%
⊙251.33	3	19.36	0.27%	95.06%	1.03%	81.29%
26.93	28	18.88	0.26%	95.32%	0.98%	82.27%
11.78	64	17.89	0.23%	95.56%	0.88%	83.15%
377.00	2	16.96	0.21%	95.76%	0.79%	83.95%
27.93	27	16.63	0.20%	95.97%	0.76%	84.71%
6.03	125	16.45	0.20%	96.16%	0.75%	85.45%

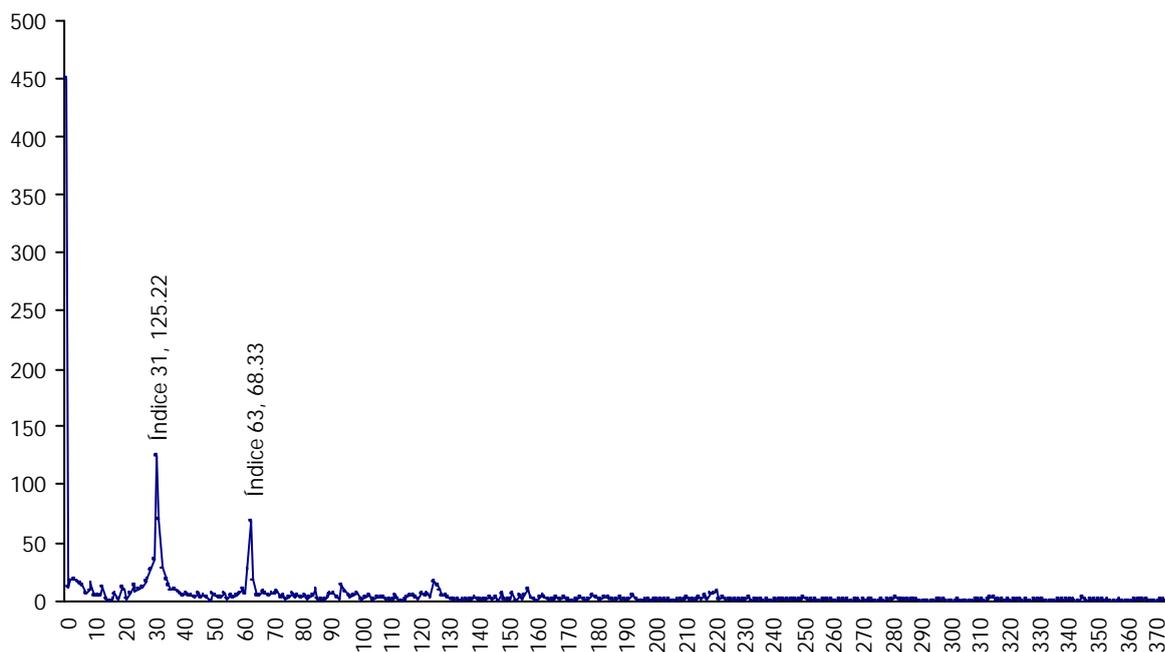


Gráfico 4.6. Transformada Discreta de Fourier de la serie del depósito Polígono Ofra.

Depósito Salamanca

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	1568.27	80.02%	80.02%		
☒24.32	31	342.47	7.63%	87.65%	38.20%	38.20%
11.97	63	213.22	2.96%	90.61%	14.81%	53.01%
☒23.56	32	182.10	2.16%	92.77%	10.80%	63.81%
⊙188.50	4	96.48	0.61%	93.37%	3.03%	66.84%
☒22.85	33	92.71	0.56%	93.93%	2.80%	69.64%
☒25.13	30	84.11	0.46%	94.39%	2.30%	71.94%
12.16	62	75.98	0.38%	94.77%	1.88%	73.82%
11.78	64	71.31	0.33%	95.10%	1.66%	75.48%
⊙150.80	5	65.72	0.28%	95.38%	1.41%	76.88%
☒26.00	29	60.29	0.24%	95.62%	1.18%	78.07%
21.54	35	56.13	0.21%	95.82%	1.03%	79.09%
6.03	125	51.09	0.17%	95.99%	0.85%	79.94%
8.02	94	49.03	0.16%	96.15%	0.78%	80.73%
377.00	2	47.76	0.15%	96.30%	0.74%	81.47%

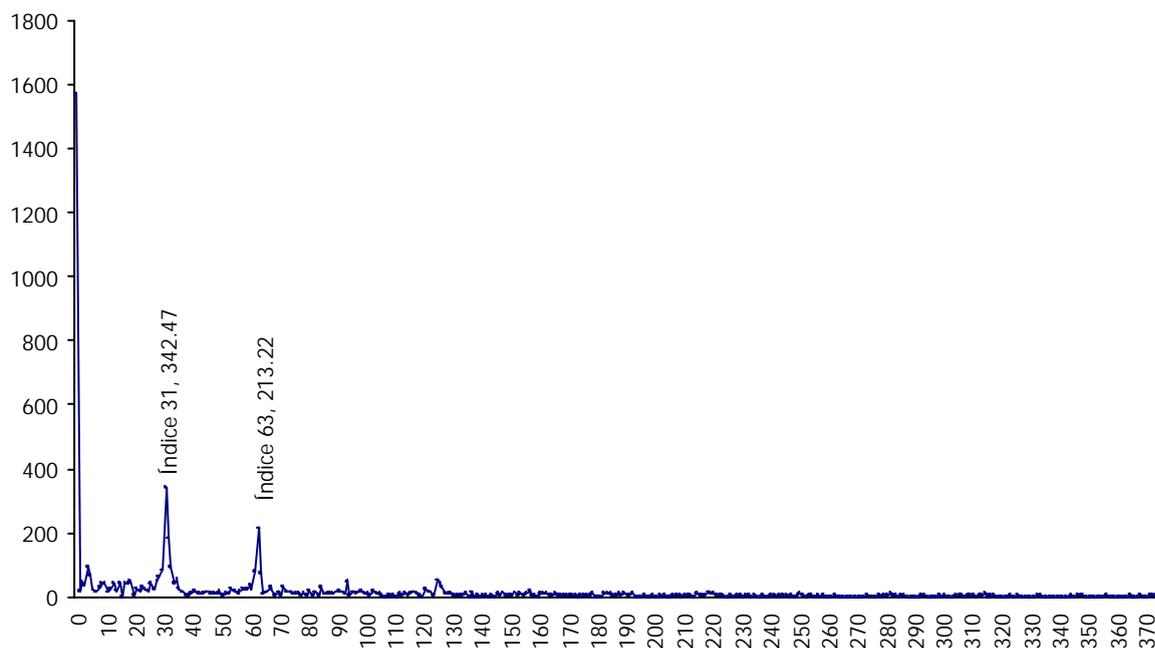


Gráfico 4.7. Transformada Discreta de Fourier de la serie del depósito Salamanca.

Depósito Salud

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	1137.46	80.30%	80.30%		
☑24.32	31	264.52	8.68%	88.98%	44.08%	44.08%
11.97	63	155.54	3.00%	91.99%	15.24%	59.33%
☑23.56	32	153.37	2.92%	94.91%	14.82%	74.15%
☑25.13	30	66.15	0.54%	95.45%	2.76%	76.90%
☑22.85	33	59.72	0.44%	95.89%	2.25%	79.15%
12.16	62	58.42	0.42%	96.32%	2.15%	81.30%
☑26.00	29	47.29	0.28%	96.59%	1.41%	82.71%
27.93	27	38.89	0.19%	96.78%	0.95%	83.66%
11.78	64	38.80	0.19%	96.97%	0.95%	84.61%
6.03	125	38.61	0.19%	97.15%	0.94%	85.55%
5.98	126	33.65	0.14%	97.29%	0.71%	86.26%
☑22.18	34	33.11	0.14%	97.43%	0.69%	86.96%
26.93	28	27.82	0.10%	97.53%	0.49%	87.44%
21.54	35	24.65	0.08%	97.60%	0.38%	87.83%

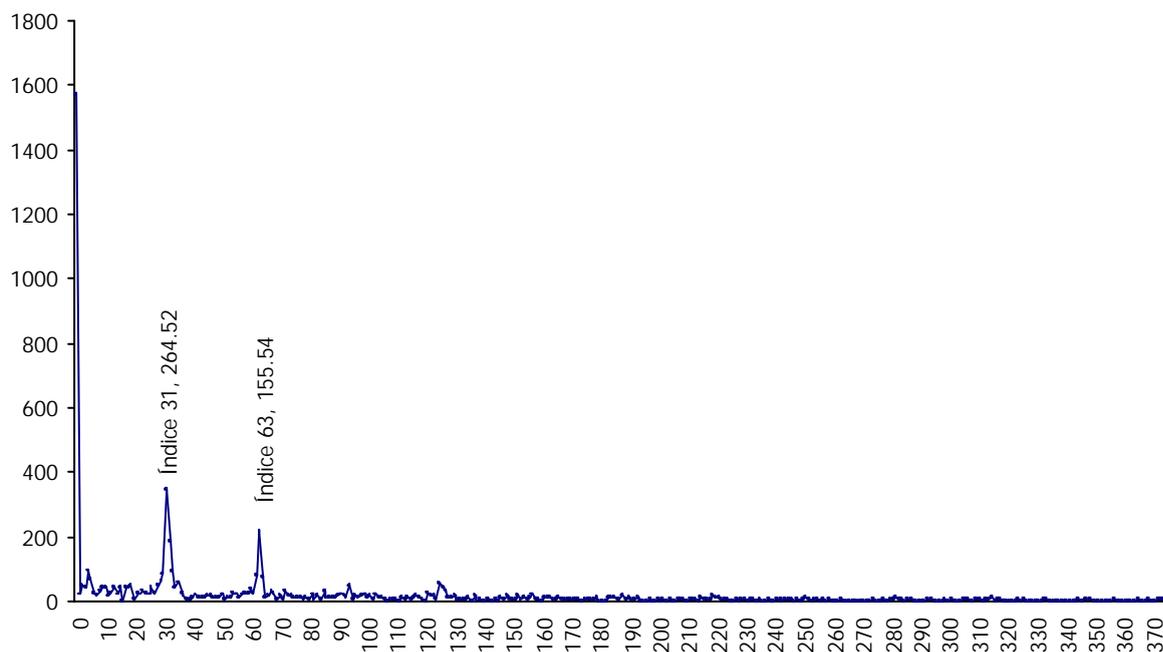


Gráfico 4.8. Transformada Discreta de Fourier de la serie del depósito Salud.

Depósito San Andrés

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	246.24	86.36%	86.36%		
☑24.32	31	45.21	5.82%	92.18%	42.66%	42.66%
☑23.56	32	28.28	2.28%	94.46%	16.69%	59.36%
11.97	63	17.18	0.84%	95.30%	6.16%	65.52%
☑25.13	30	12.07	0.42%	95.71%	3.04%	68.56%
22.85	33	10.98	0.34%	96.05%	2.52%	71.08%
83.78	9	8.28	0.20%	96.25%	1.43%	72.51%
12.16	62	7.88	0.18%	96.43%	1.30%	73.81%
27.93	27	7.74	0.17%	96.60%	1.25%	75.06%
377.00	2	7.47	0.16%	96.76%	1.17%	76.22%
☑26.00	29	7.41	0.16%	96.91%	1.15%	77.37%
⊙251.33	3	6.82	0.13%	97.04%	0.97%	78.34%
125.67	6	6.80	0.13%	97.18%	0.97%	79.30%
68.55	11	6.77	0.13%	97.31%	0.96%	80.26%
39.68	19	6.61	0.12%	97.43%	0.91%	81.17%

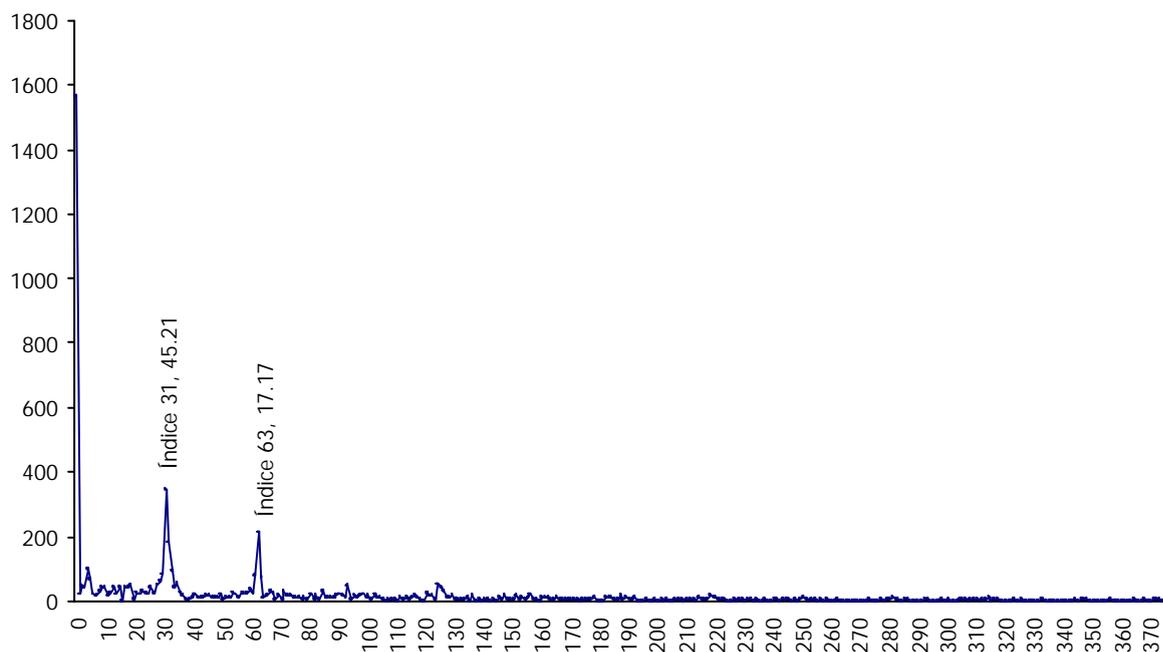


Gráfico 4.9. Transformada Discreta de Fourier de la serie del depósito San Andrés.

Depósito Tío Pino

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0	228.05	81.03%	81.03%		
☑24.32	31	52.04	8.44%	89.47%	44.50%	44.50%
☑23.56	32	30.14	2.83%	92.30%	14.92%	59.42%
11.97	63	29.06	2.63%	94.94%	13.88%	73.30%
☑25.13	30	13.40	0.56%	95.49%	2.95%	76.25%
☑22.85	33	13.22	0.54%	96.04%	2.87%	79.12%
12.16	62	10.60	0.35%	96.39%	1.85%	80.96%
☑26.00	29	10.47	0.34%	96.73%	1.80%	82.76%
☑22.18	34	8.89	0.25%	96.98%	1.30%	84.06%
6.03	125	6.86	0.15%	97.12%	0.77%	84.84%
11.78	64	6.84	0.15%	97.27%	0.77%	85.60%
27.93	27	6.56	0.13%	97.40%	0.71%	86.31%
5.98	126	6.20	0.12%	97.52%	0.63%	86.94%
4.80	157	5.91	0.11%	97.63%	0.57%	87.52%
26.93	28	5.72	0.10%	97.73%	0.54%	88.05%

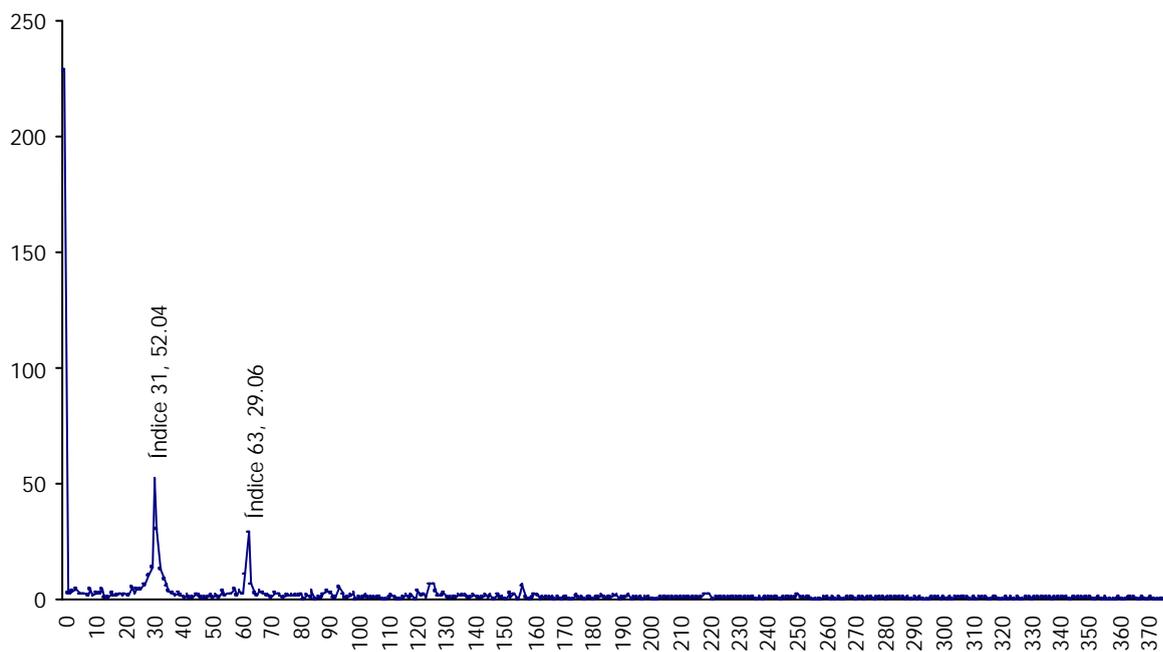


Gráfico 4.10. Transformada Discreta de Fourier de la serie del depósito Tío Pino.

Depósito Vuelta Grande

p_r	r	w_r	$e_r \%$	$\sum e_r \%$	$e_r \%$	$\sum e_r \%$
	0					
	0	19518.93	86.52%	86.52%		
☒24.32	31	4527.23	9.31%	95.83%	69.06%	69.06%
12.16	62	1811.37	1.49%	97.32%	11.06%	80.12%
⊙188.50	4	706.11	0.23%	97.55%	1.68%	81.80%
754.00	1	647.08	0.19%	97.74%	1.41%	83.21%
⊙150.80	5	608.76	0.17%	97.91%	1.25%	84.46%
☒25.13	30	550.35	0.14%	98.04%	1.02%	85.48%
6.08	124	523.02	0.12%	98.17%	0.92%	86.40%
⊙251.33	3	510.78	0.12%	98.29%	0.88%	87.28%
83.78	9	481.20	0.11%	98.39%	0.78%	88.06%
58.00	13	466.90	0.10%	98.49%	0.73%	88.79%
34.27	22	422.44	0.08%	98.57%	0.60%	89.40%
21.54	35	378.96	0.07%	98.64%	0.48%	89.88%
8.11	93	366.40	0.06%	98.70%	0.45%	90.33%
377.00	2	362.15	0.06%	98.76%	0.44%	90.77%

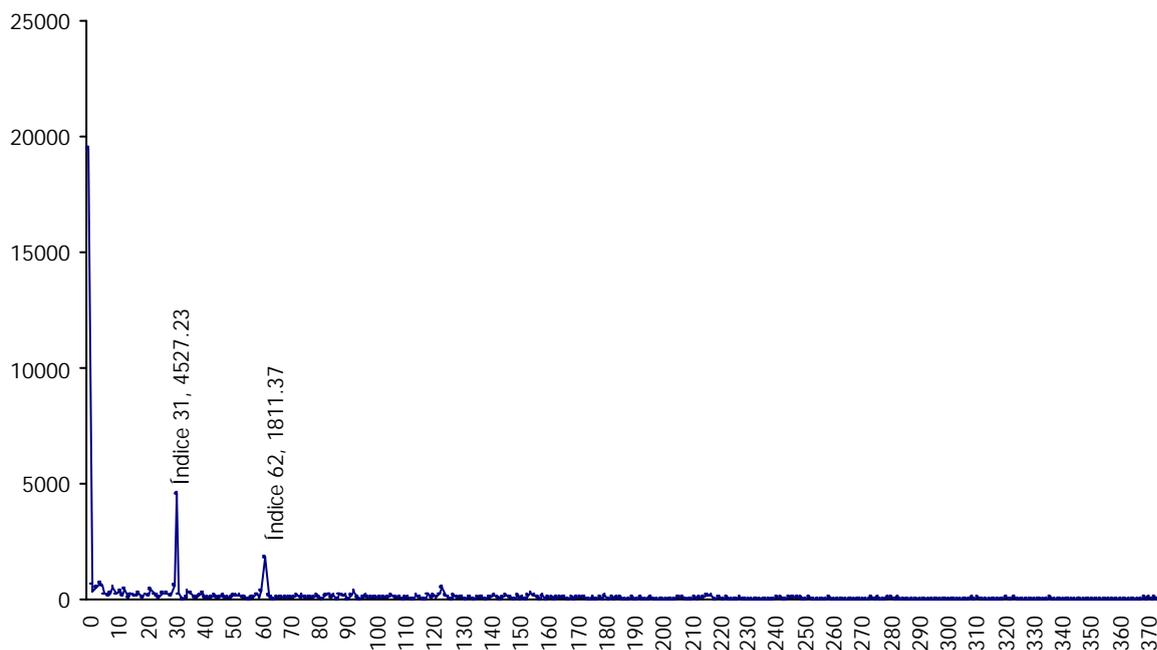


Gráfico 4.11. Transformada Discreta de Fourier de la serie del depósito Vuelta Grande.

4.3. La periodicidad semanal

Por lo expuesto anteriormente, la periodicidad que más influye en el desarrollo de las series de consumo de agua es la diaria. Aparece en los primeros puestos en todas las tablas de pesos y lidera el índice más alto de frecuencia relevante excepto en el Depósito Barranco Grande, donde el índice *94* que corresponde a un periodo de *8* horas, es el más influyente.

Ésta es la razón que nos permitiría simplificar la serie cronológica de datos, agrupando y obteniendo las medias diarias, pasando de un tamaño $N=754$ a otro menor $N=24$. Esta agrupación es suficiente para recoger no sólo la influencia diaria, sino cualquier de índice superior a ésta: *12* horas, *8* horas, *6* horas etc.

La razón de esta simplificación surge de la necesidad de reducir el índice de la mayor frecuencia con relevancia. Si en la mayoría de los depósitos éste es *62-63*, el grado del polinomio predictor usado en los modelos sería de esa magnitud, con lo que serían necesarios, al menos, tantos coeficientes, con lo que complicamos el modelo. Al ser efectuada la agrupación, con $N=24$, el mayor índice relevante pasa a ser sólo *2*, puesto que $24/2=12$, que es el periodo cuyo peso se ha de conservar. Aumentando el grado del polinomio hasta cuatro se cubren también las periodicidades de *8* horas y *6* horas. Por lo tanto, se ha reducido considerablemente el número de coeficientes que intervienen en el modelo, simplificándolo.

Sin embargo, esta drástica simplificación no es aplicable a todas las series. Agrupar y calcular medias supone pérdida de información por uniformidad con lo que tenemos que ser cuidadosos, comprobando si dicha pérdida es sensible. Aplicar un modelo de comportamiento diario implica que la serie modelizada no se ve afectada por el día de la semana que es, esto es, se comporta igual

un lunes que un viernes. Esto ha de ser verificado, para lo cual debemos estudiar la periodicidad semanal: en aquellas series en la que tenga peso esta periodicidad la simplificación se realizará pero a $N=168$, esto es, las 24 horas de los 7 días de la semana, o lo que es lo mismo, un modelo de patrón de comportamiento semanal; para aquellas series en las que esta periodicidad no tenga influencia contruiremos el modelo de patrón de comportamiento diario.

En el *Gráfico 4.12* se suman los pesos de los índices 3, 4 y 5 que corresponderían a la influencia semanal. Se muestran por orden decreciente y divididos en 2 grupos según lo que hemos considerado que pueda ser una separación entre series de consumo con o sin influencia de la periodicidad semanal. El primer grupo aparece en un color más oscuro que el segundo, indicando en porcentajes el peso de la influencia semanal, y siendo el valor de corte *0.30%*.

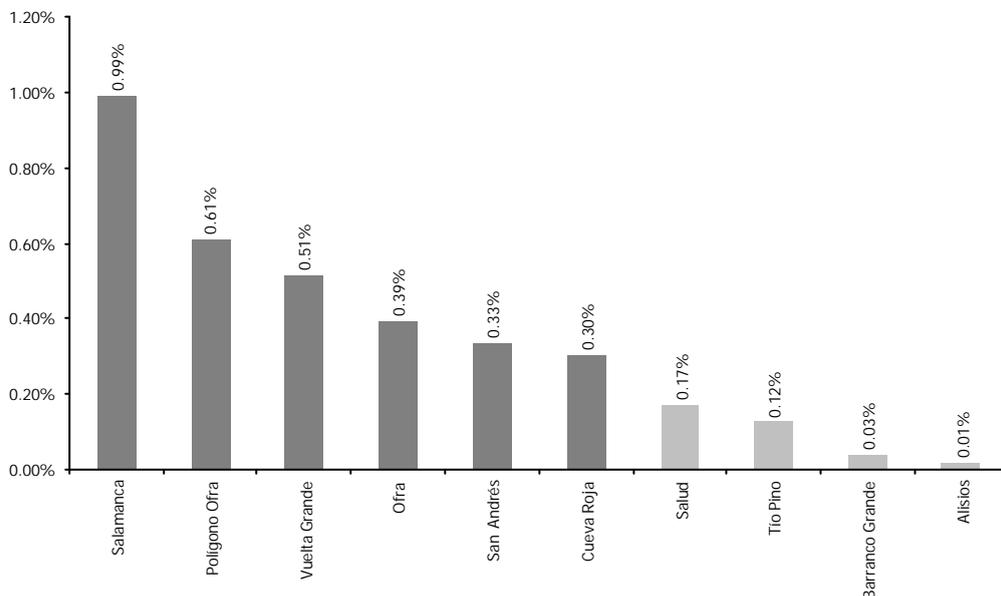


Gráfico 4.12. Peso de la influencia semanal en las series de cada uno de los depósitos.

Por lo tanto, se deberá construir un modelo semanal para los depósitos de Cueva Roja, Ofra, Polígono Ofra, Salamanca, San Andrés y Vuelta grande. Y, en cambio, se deberá calcular un modelo diario para Alisios, Barranco Grande, Salud y Tío Pino.

Procede ahora un estudio del comportamiento de los modelos y sus errores para las secuencias de datos agrupados correspondientes a los distintos depósitos.

4.4. Estudio de los modelos

Dado que la modelización que estudiamos se presenta desde el dominio de la frecuencia, y puede ser considerado su posterior traslado al dominio del tiempo, son dos los errores que deben ser estudiados. Por un lado, en el dominio de la frecuencia, el que se establece entre la secuencia de la *DFT* y su modelo y por otro, en el dominio del tiempo, el que se establece entre la secuencia de medias y los modelos de la *DFT* trasladados al tiempo. Sin embargo, como la modelización mediante el *aproximante de Padé-Laurent* no construye la secuencia en el dominio de la frecuencia, sino una forma cerrada polinomial de la transformada inversa de Fourier discreta, el estudio para este último caso lo realizamos sólo en el dominio del tiempo.

Además, para la comparativa entre el modelo y la serie original, y el estudio de los errores, vamos a hacer intervenir una modelización estándar mediante secuencias truncadas de Fourier que usen el mismo número de parámetros. Para cualquiera de los modelos: *lineal*, *PL (aproximante de Padé-Laurent)* y *TF (truncada de Fourier)*, evaluaremos el error cuadrático de la aproximación.

Hemos escogido dos depósitos para el desarrollo de los comentarios: Depósito Alisios para el agrupamiento diario, y Depósito Cueva Roja para el agrupamiento semanal.

Modelos para el depósito Alisios

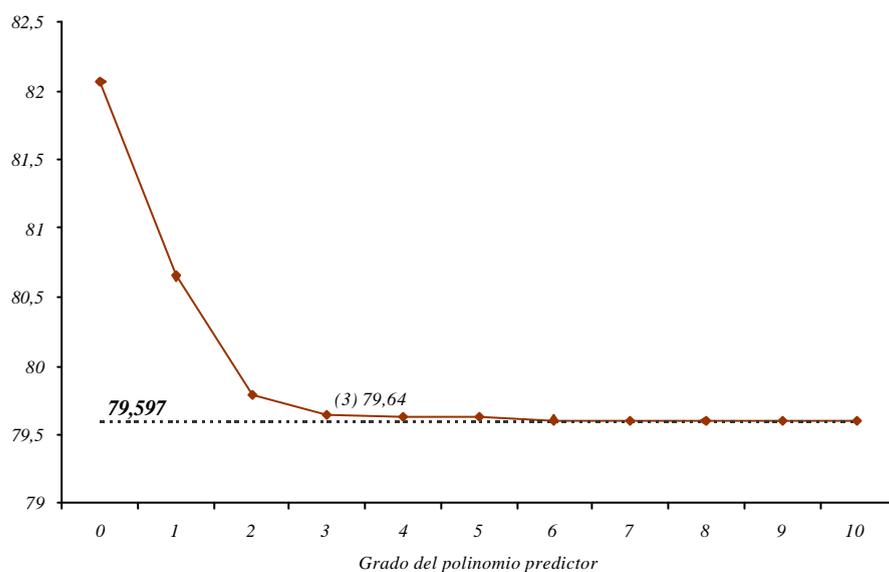


Gráfico 4.13. Ganancia del filtro para el modelo lineal

El *Gráfico 4.13* muestra la convergencia de la ganancia del filtro al peso de la componente continua de la *DFT*, el valor *79,597*. Como comprobamos no hace falta un grado muy elevado para lograr una buena aproximación: el grado *3* es suficiente. Usaremos este grado como referencia para la construcción de los modelos. Este orden también es suficiente para el modelo *PL* pues asegura la inclusión en el mismo hasta el peso de la frecuencia de índice *3*.

Esta decisión también se comprueba en el *Gráfico 4.14* en el que se muestran los errores cuadráticos del modelo lineal y una truncada para el mismo número de parámetros. Nótese en este caso que, si el

grado del polinomio predictor es p , el modelo lineal usa $p+1$ parámetros, con lo que usaremos una truncada de tal número de coeficientes. Si se elige por la convergencia de la ganancia $p=3$, el modelo lineal muestra un error sensiblemente menor que el de la truncada con 4 coeficientes.

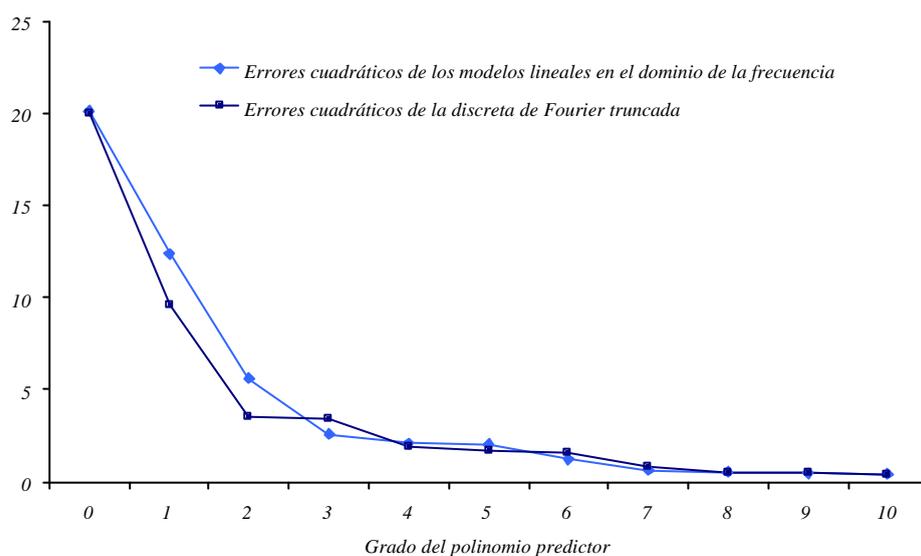


Gráfico 4.14. Errores cuadráticos en el dominio de la frecuencia

El *Gráfico 4.15* apoya de nuevo la decisión tomada respecto al orden del modelo, pues es en este grado donde la truncada de Fourier muestra, aunque sensiblemente, peores resultados. En este gráfico se comparan los modelos *lineal*, *PL* y *TF* en el dominio del tiempo.

La coincidencia entre la *DFT* y el modelo lineal para $p=3$ se comprueba en el *Gráfico 4.16*, y, finalmente, el *Gráfico 4.17* muestra los 3 modelos junto a la serie de datos original agrupada de forma diaria. Se pone de manifiesto en este gráfico que los modelos

suavizan el comportamiento de la serie original y que, entre todos, el de mejor comportamiento es el *lineal*.

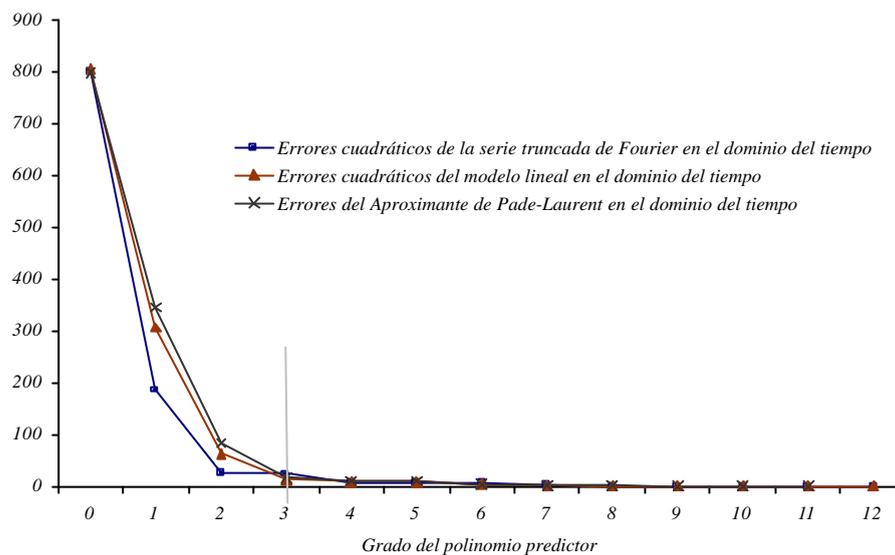


Gráfico 4.15. Errores cuadráticos en el dominio del tiempo

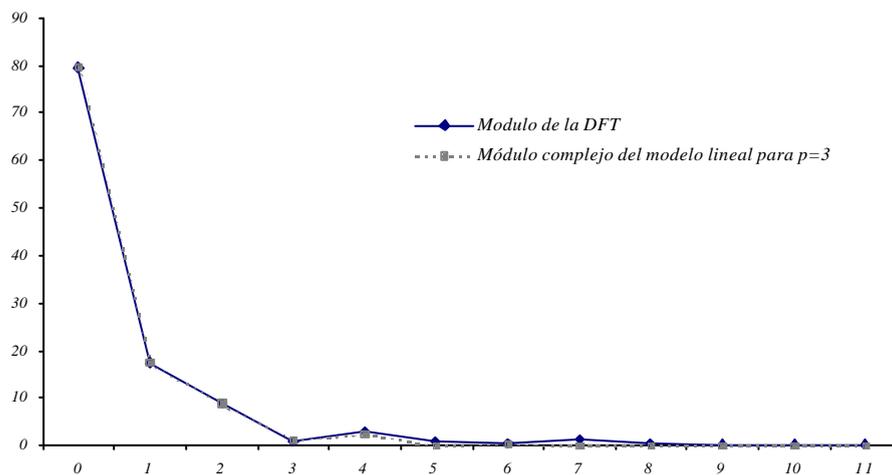


Gráfico 4.16. Módulos complejos de la DFT y su modelo lineal de orden 3

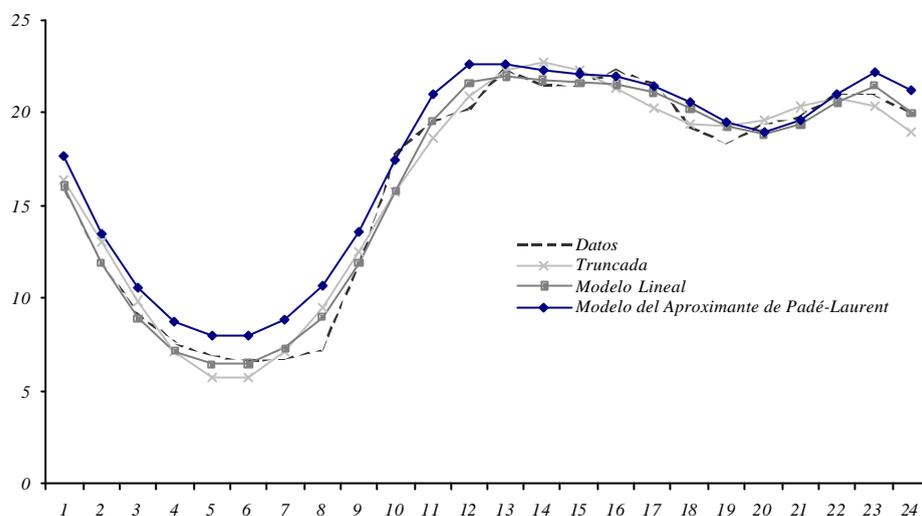


Gráfico 4.17. Comparativa en el dominio del tiempo

La siguiente sección repite la aplicación de los modelos para el depósito de Cueva Roja. Sobre la secuencias agrupadas semanalmente se comprueba más fácilmente las conclusiones de convergencia y ajuste al ser las secuencias de datos más largas.

Modelos para el depósito de Cueva Roja

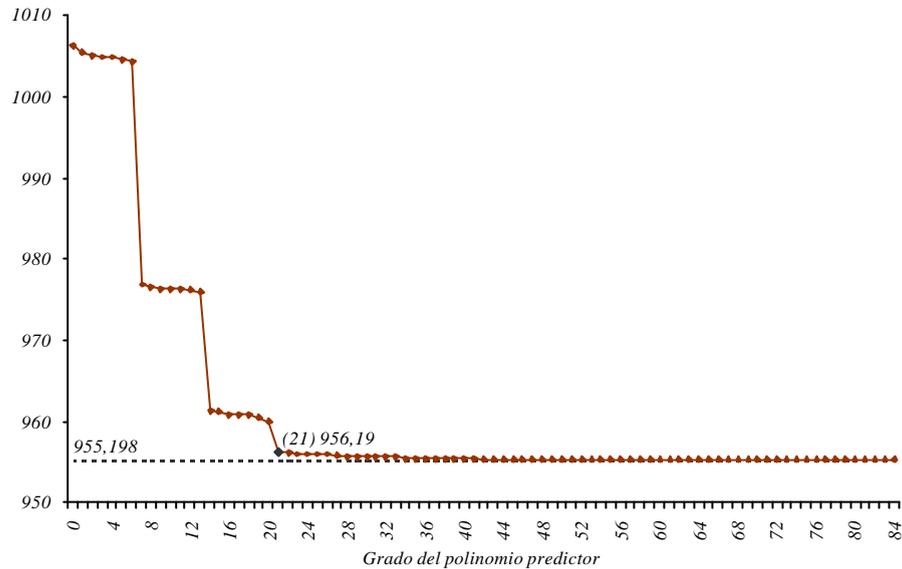


Gráfico 4.18. Ganancia del filtro para el modelo lineal

En *Gráfico 4.18*, se muestra la convergencia en saltos de la ganancia del filtro. Las zonas horizontales del desarrollo de la convergencia indican que no se produce variación estimable del coeficiente de reflexión, lo que nos va a permitir usar la versión acelerada del algoritmo.

Es $p=21$ el valor escogido para asegurar una convergencia suficiente. Este valor se justifica en los gráficos siguientes, *4.19* y *4.20* de errores cuadráticos en ambos dominios. Es el modelo lineal el que mejor resultado ofrece, mientras que el *PL* tiene un comportamiento similar al calculado con la trunca de Fourier. Nótese que para el orden elegido, el error del modelo lineal avanza al menos en 7 parámetros el comportamiento de los restantes modelos en ambos dominios.

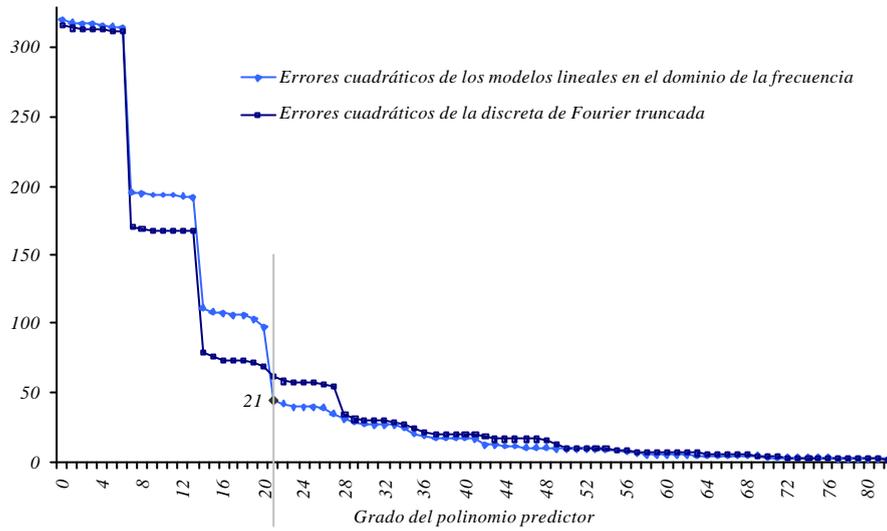


Gráfico 4.19. Errores cuadráticos en el dominio de la frecuencia

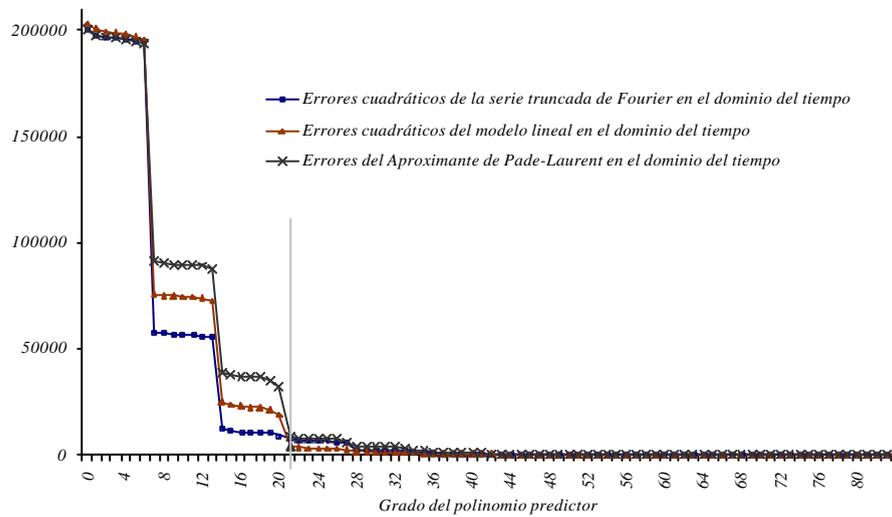


Gráfico 4.20. Errores cuadráticos en el dominio del tiempo

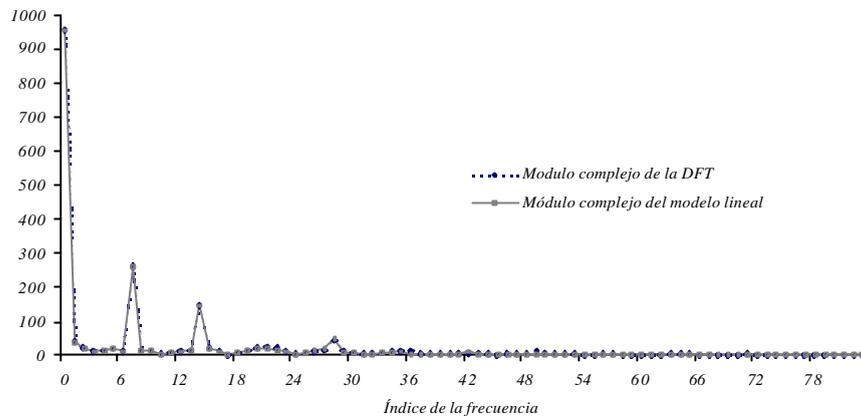


Gráfico 4.21. Comparativa entre el módulo complejo de la DFT y el modelo lineal para $p=21$

Hemos ampliado la comparativa de los modelos lineales al usar la versión acelerada del algoritmo para este modelo, con un corte de 0.05 . Es decir, son anulados los coeficientes de reflexión de módulo inferior a dicha cantidad. En el *Gráfico 4.21* se muestra el buen ajuste de la secuencia *DFT* de la serie de datos agrupados semanalmente, por su modelo lineal para $p=21$. El *Gráfico 4.22* muestra un ajuste equivalente para el modelo contruido a través el algoritmo acelerado.

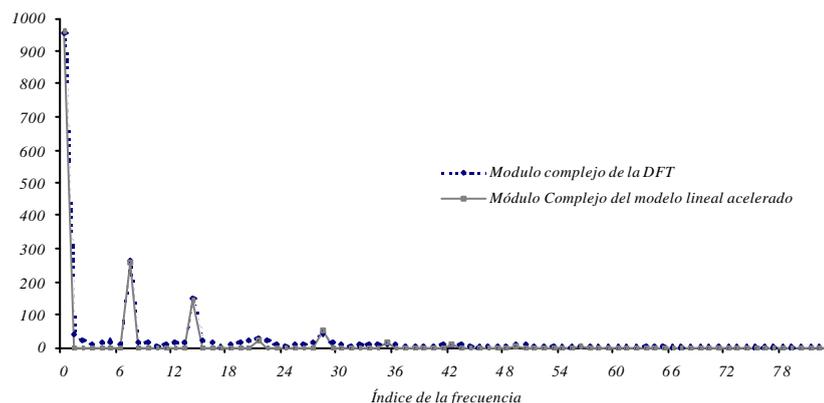


Gráfico 4.22. Comparativa entre el módulo complejo de la DFT y el modelo lineal acelerado para $p=21$

Además quedan igualmente anulados muchos de los coeficientes del polinomio predictor sin modificarse sustancialmente el ajuste. En particular el filtro lineal queda simplificado a,

$$\frac{959.102}{1 + (0.2056 - 0.1787i)z^7 - (0.1140 + 0.1598i)z^{14} - (-0.0931 + 0.0150i)z^{21}},$$

esto es, sólo 3 coeficientes de los 21 iniciales.

Para finalizar el presente capítulo, mostramos la serie de datos semanal junto a los modelos que se presentan en la memoria. De nuevo es el modelo lineal el que, visualmente se ajusta mejor a la secuencia original.

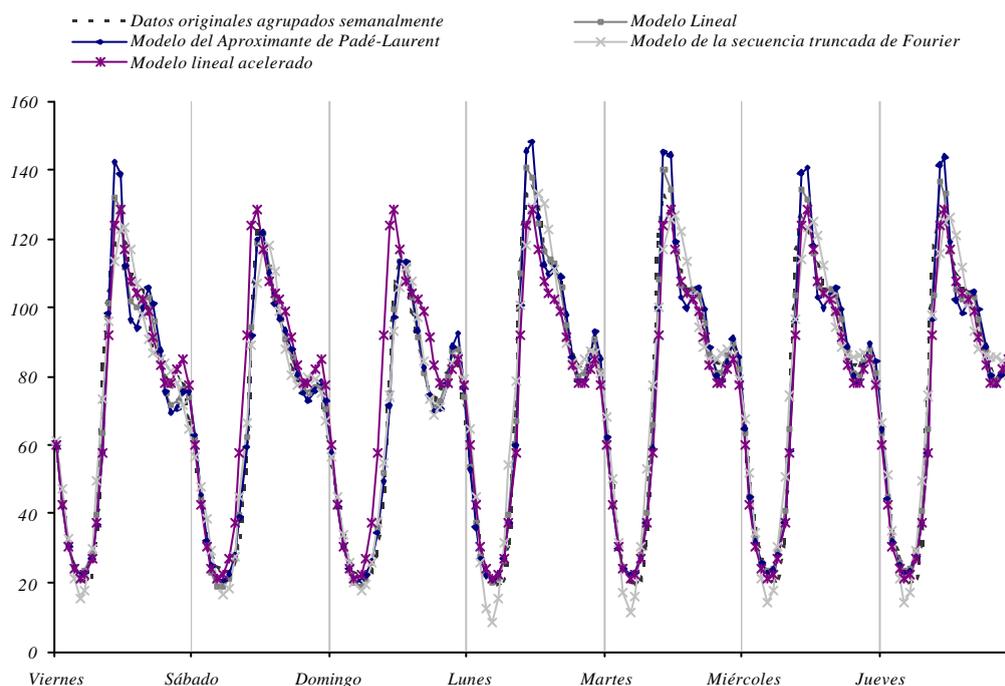


Gráfico 4.23. Comparativa de los modelos en el dominio del tiempo.

Conclusiones y Cuestiones abiertas

En la presente memoria proponemos el uso de dos modelos para aproximar el dominio de la frecuencia. Mediante el primer modelo, se ha adaptado una estructura lineal por retardos ampliamente usada para secuencias estacionarias a una serie de datos que ni siquiera es cronológica. En el segundo caso, se ha contruido un modelo racional usando aproximantes de Padé-Laurent. Para ambos casos, la computación de los parámetros del modelo se ha resuelto mediante el algoritmo de Levinson.

La justificación del uso de un filtro lineal para la modelización de una transformada discreta de Fourier de datos reales, no viene únicamente determinada por un buen ajuste. Una conveniente interpretación de la modelización a través de los coeficientes de reflexión y no de los coeficientes de los retardos completa el estudio. Precisamente es la interpretación que se hace de estos valores, la que posibilita la escritura de una versión acelerada del algoritmo de Levinson, que elimina la computación de valores no relevantes. Se produce entonces un ahorro en el número de cálculos recursivos que no merma el ajuste del modelo final.

La identificación del orden del modelo se realiza usando propiedades de convergencia del valor de la ganancia del filtro a su

valor verdadero. Este método, aunque se basa en apreciaciones relativas, resuelve la elección del grado de forma práctica.

La otra vertiente que se toma es la referida al modelo racional mediante aproximantes de series biinfinitas de Padé-Laurent. Tomando la relación del algoritmo de Levinson con estos aproximantes se construye para la serie tratada una formulación cerrada de la transformada inversa. Para la elección del orden del aproximante en este segundo modelo planteado, se estima simplemente hasta qué frecuencia se quiere añadir al modelo.

En la aplicación práctica de ambos modelos, contamos con secuencias periódicas de datos reales que permiten, tras un análisis de periodicidades en el dominio de la frecuencia, una primera simplificación por agrupamientos. Este cálculo previo permite, sin pérdida de información importante, que la complejidad de los modelos sea también menor. Sin embargo, es para la segunda serie de datos agrupados, la correspondiente a una serie semanal, donde las propiedades planteadas se comprueban con más claridad: la convergencia de la ganancia del filtro y la elección del orden del modelo. Una comparativa entre los modelos propuestos y la serie truncada de Fourier pone de manifiesto la ventaja del modelo lineal para un mismo número de parámetros. La conclusión que extraemos de este ejercicio práctico es la de considerar como mejor modelo al lineal, y como equivalentes el construido mediante aproximantes y el deducido mediante secuencias truncadas.

Varias cuestiones quedan abiertas por esta investigación. En primer lugar sería interesante comparar los métodos de identificación de órdenes para aproximaciones racionales de la literatura con los aquí expuestos. Sin embargo, dada la particularidad de éstos últimos para los modelos planteados, no es previsible que puedan ser generalizados.

Una segunda cuestión que puede ser planteada es la posibilidad de uso con éxito de los modelos planteados en estrategias

aceleradoras de convergencia en truncadas de Fourier y sumas trigonométricas, véase [JonHar70]. Este estudio es el que centra en la actualidad nuestro trabajo.

La tercera cuestión globaliza la aplicación práctica realizada con las series de consumo de agua en un entorno urbano. Por un lado, hay que extender la serie de datos a un tamaño más razonable que pueda hacer posible el estudio de frecuencias mensuales, e incluso anuales. Construidos los modelos, el siguiente paso sería clasificatorio, es decir, habría que establecer las relaciones oportunas entre las frecuencias con peso y el tipo de población de la zona. Este estudio aseguraría el éxito a la hora de la planificación de nuevos depósitos para zonas de la ciudad en desarrollo.

Respecto a este tema, podría ser de utilidad contemplar la gestión de los depósitos de agua como un sistema de almacén e inventario. Modelizada la demanda, el depósito debe satisfacerla con dos requisitos: el nivel de almacén debe ser suficiente para que no se produzca desabastecimiento, y menor que una cota máxima que produciría desaprovechamiento. El uso de modelos de inventario no redundaría sólo en ahorro de recursos, sino que, por ejemplo, podría determinar el tamaño óptimo del depósito para el mejor abastecimiento.

Bibliografía

-
- [Alo94] S. Alonso
Una introducción práctica a la aproximación y tratamiento de señales.
Memoria de Licenciatura. Universidad de La Laguna. España. 1994.
-
- [AloGon00] S. Alonso; C. González-Concepción.
Linear Approximation in the Discrete Frequency Domain.
Journal of Computational and Applied Mathematics. Febrero 2000. En Prensa.
-
- [AloGon96] S. Alonso; C. González-Concepción.
Una Aplicación del Algoritmo de Levinson en la Modelización Espectral.
Contribución del Libro Conmemorativo de los 25 años de Matemáticas en la Univesidad de La Laguna. 1996.
-
- [AloGon98] S. Alonso; C. González-Concepción
Análisis y modelización de series de consumo de agua en el dominio de la frecuencia. Una aplicación al caso de Santa Cruz de Tenerife.
Actas del Primer Seminario de Investigación Universitaria sobre la Economía Canarias. Junio de 1998. Tomo I. Páginas 319-330.
-
- [AmmGra88] G. S. Ammar; W. B. Gragg.
Superfast solution of Real Positive Definite Toeplitz Systems.
SIAM J. Matrix Anal. Appl. Vol 9. Número 1. Enero 1988.. Páginas 61-76.
-
- [Ata70] B. S. Atal
Speech Analysis and Synthesis by Linear Prediction of the Speech Wave.
J. Acoust Soc. Am. Vol 47. Número 65(A). 1970.
-
- [BakGra96] G. A. Baker, Jr; P. Graves-Morris.
Padé Approximants.
Cambridge University Press. Encyclopedia of Mathematics and Its Applications 59. 1996.
-
- [Blo76] P. Bloomfield
Fourier Analysis of Time Series: An introduction.
John Wiley & Sons. 1976.
-
- [BoxJen76] G. E. P. Box; G. M. Jenkins.
Times series analysis. Forecasting and control. Revised Edition
Holden-Day. 1976.
-

[Bra86]	R. N. Bracewell. <i>The Hartley transform.</i> Oxford University Press. Oxford Engineering Science Series, 19. 1986.
[BroDav91]	P. J. Brokwell; R. A. Davis. <i>Time Series: Theory and Methods.</i> Springer-Verlag. Springer Series in Statistics. Second Edition, 1991.
[Bul87]	A. Bultheel. <i>Laurent Series and their Padé Approximations.</i> Birkhäuser Verlag. 1987
[BulBar93]	A. Bultheel; M. Van Barel. <i>Linear Prediction: mathematics and engineering.</i> Bulletin of the Belgian Mathematical Society. Vol 1. Número 1. 1991.
[Cho92]	B. S. Choi. <i>ARMA model identification.</i> Springer-Verlag. Springer Series in Statistics: Probability and its Applications. 1992.
[CocCoo67]	W. T. Cochran; J. W. Cooley. <i>What is the Fast Fourier Transform?</i> IEEE Trans. Audio Electroacoust. Vol AU-15. 1967. Páginas 45-55.
[CooLewWel69]	J. W. Cooley; P. A. Lewis; P. D. Welch. <i>The Finite Fourier Transform.</i> IEEE Trans. Audio Electroacoust. Vol AU-17. 1969. Páginas 77-85.
[CooLewWel70]	J. W. Cooley; P. A. Lewis; P. D. Welch. <i>The Fast Fourier Transform Algorithm: Programming Considerations in the calculation of Sine, Cosine and Laplace Transform.</i> J. Sound Vib. Vol 12. 1970. Páginas 315-337.
[CooTuc65]	J. W. Cooley; Tuckey J. W. <i>An algorithm for the machine calculation of complex Fourier series.</i> Mathematics of Computation. Vol 19. Número 90. 1965. Páginas 297-301.
[Cyb80]	G. Cybenko. <i>The Numerical Stability of the Levinson-Durbin Algorithm for Toeplitz Systems of Equations.</i> SIAM Journal on Scientific and Statistical Computing. Vol 1. Número 3. Septiembre 1980. Páginas 303-319.
[DelGen86]	Philippe Delsarte; Yves V. Genin. <i>The split Levinson algorithm.</i> IEEE Trans. Acoust. Speech Signal Process. Vol 34. Número 3. 1986. Páginas 470-478.
[DelGen91]	P. Delsarte; Y. V. Genin. <i>An introduction to the class of Split Levinson Algorithms</i> Invited Lecture in <i>Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms.</i> Springer-Verlag. NATO ASI Series. 1991

-
- [Gil95] C. Gil.
Propuesta Metodológica en el Contexto de Modelos de Función de Transferencia con Expectativas.
Tesis Doctoral. Universidad de La Laguna. 1995. España.
-
- [GilGon97] C. Gil; C. González-Concepción.
Modelización racional de series temporales no causales: algunas propuestas de caracterización dinámica
Estudios de Economía Aplicada. Número 8. 1997. Páginas 77-107.
-
- [Gra94] J. W. Gray.
Mastering Mathematica.
A P Professional. 1994.
-
- [Jim91] M. Jiménez
Sobre R-Funciones Y Aproximacion Tipo-Pade Multipuntual.
Tesis Doctoral. Universidad de La Laguna. 1991.
-
- [JonHar70] W. B. Jones; G. Hardy
Accelerating Convergence of Trigonometric Approximations
Mathematics of Computation, Vol 24, Número 111, Julio 1970, Páginas 547-560.
-
- [Jur64] E. I. Jury.
Theory and Application of the z-transform Method.
Wiley. 1964.
-
- [Kai74] T. Kailath.
A View of Three Decades of Linear Filtering Theory.
IEEE Trans. IT-20. Páginas 146-181. 1974.
-
- [LanTis85] P. Lancaster; M. Tismenetsky.
The theory of matrices.
Academic Press. Second edition. 1985.
-
- [Lat92] B. P. Lathi
Linear Systems and Signals.
Berkeley-Cambirdge Press. 1992
-
- [Lev47] N. Levinson.
The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction.
J. Math. Phys. Vol 25. 1947. Páginas 261-278.
-
- [Mae91] R. Maeder.
Programming in Mathematica.
Addison-Wesley. Segunda Edición. 1991.
-
- [Mak75] J. Makhoul.
Linear Prediction: A Tutorial Review.
Proceedings of the IEEE. Vol 63. Número 4. Abril 1975. Páginas 561-580.
-
- [Man89] U. Manber
Introduction to Algorithms. A Creative Approach.
John Wiley & Sons. 1976.
-

[MarGra76]	J. D. Markel; A. H. Gray Jr. <i>Linear Prediction of Speech.</i> Springer-Verlag. 1976.
[OppSch75]	A. V. Oppenheim; R. W. Schaffer. <i>Digital Signal Processing.</i> Prentice/Hall International Inc. 1975.
[OppWilYou83]	A. V. Oppenheim; A. S. Willsky; I. T. Young. <i>Signals and Systems.</i> Prentice/Hall International, Inc. 1983.
[Par61]	E. Parzen <i>An approach to time series analysis.</i> Ann. Math Statist. Número 32. Páginas 951-989.
[PeaTre69]	K. L. Peacock, S. Treitel. <i>Predictive Deconvolution: Theory and Practice.</i> Geophysics. Vol 34. Número 2. Abril 1969. Páginas 155-169.
[Pri81]	M. B. Priestley. <i>Spectral Analysis and Time Series.</i> Academic Press. 1981.
[Pri88]	M. B. Priestley. <i>Nonlinear and nonstationary time series analysis.</i> Academic Press. 1988.
[Rob67]	E. A. Robinson. <i>Predictive Decomposition of Time Series with Application to Seismic Exploration.</i> Geophysics. Vol XXXII. Número 3. Junio 1967. Páginas 418-184.
[RobMul87]	R. A. Mullis; R. A. Roberts. <i>Digital Signal Processing.</i> Addison-Wesley. 1987.
[RobTre67]	E. A. Robinson; S. Treitel. <i>Principles of Digital Wiener Filtering.</i> Geophysical Prospecting. Vol XV. 1967. Páginas 311-333.
[RobTre77]	E. A. Robinson; S. Treitel. <i>The Spectral Function of a Layered System and the Determination of the Waveforms at Depth.</i> Geophysical Prospecting. Vol 25. 1977. Páginas 434-459.
[Sze67]	G. Szegö <i>Orthogonal polynomials.</i> American Mathematical Society Colloquium Publications, Vol. 23. Third edition. American Mathematical Society. 1967.
[Wei90]	W. S. Wei. <i>Time series analysis. Univariate and multivariate methods.</i> Addison-Wesley. 1990.

[Wie66] N. Wiener.
Extrapolation Interpolation and Smoothing of Stationary Time Series.
M. I. T. Press. 1966.

[Wol91] S. Wolfram.
Mathematica, a system for doing Mathematica by computer.
Addison-Wesley. Segunda Edición. 1991.