

Capítulo 3

ESTIMACIÓN ESPECTRAL POR EL MÉTODO DE THOMSON

3.1. INTRODUCCIÓN.

En este capítulo introducimos ya el objeto fundamental de estudio y análisis del presente Proyecto Fin de Carrera: el Método de Thomson, un nuevo método no paramétrico de estimación espectral.

Este método ha sido ampliamente utilizado en problemas de análisis de señales geofísicas, incluyendo análisis de datos atmosféricos y oceánicos (Ghil y Vautard, 1991; Mann and Park, 1993;...), datos paleoclimáticos (Berger, 1991; Mann y Lees, 1996;...), datos geoquímicos y sismológicos,...

Dicho método equivale a usar la media ponderada de un conjunto de estimaciones de “espectros-directos” basados en unas ventanas ortogonales de datos (‘discrete prolate spheroidal sequences’ o secuencias de Slepian).

Algunas de las principales características de esta estimación, como veremos, son las siguientes: no hay que ‘enfrentarse’ a ventanas arbitrarias, se trata de una sencilla teoría muestral, es consistente, proporciona un test de análisis de varianza de las componentes lineales, y además tiene alta resolución.

Como ya se ha indicado anteriormente, se trata de una estimación no paramétrica que, al contrario de las paramétricas, no asume una forma funcional específica. Tradicionalmente las estimaciones de espectro no paramétricas han sido divididas en dos clases: directas e indirectas. Las primeras son más antiguas. Debido a la carga computacional de las estimaciones directas antes del descubrimiento de la transformada rápida de Fourier, las indirectas (que son aquéllas basadas en estimaciones de las autocovarianzas de las señales) fueron comúnmente empleadas en los trabajos de Bartlett, Blackman y Tukey, ... En nuestro caso, podríamos proceder siguiendo los pasos que se indican a continuación:

- En primer lugar, se formaría una estimación de espectro *directa*:

$$\hat{S}_D(\omega) = \left| \sum_{n=0}^{N-1} x(n) D_n e^{-i\omega n} \right|^2$$

donde D_n es una ventana de datos.

- A continuación, como $\hat{S}_D(\omega)$ es inconsistente en el sentido de que su varianza no se decrementa con las dimensiones de la secuencia de muestras, lo ‘suavizaríamos’ mediante una convolución con una segunda ventana $G(\omega)$:

$$\tilde{S}(\omega) = \tilde{S}_D(\omega) * G(\omega)$$

En esta ecuación se halla implícita la conexión con la estimación *indirecta*.

Sin embargo, este proceso tiene problemas importantes: excesiva dependencia de la ventana de datos D_n en la estimación directa, la operación de suavizado es normalmente poco satisfactoria ya que no se tiene en cuenta la información de fase presente en los datos originales,...

Por todo ello, como sustitución a esas dos fases independientes en la estimación del espectro, planteamos un algoritmo unificado que tiene características muy interesantes:

- Primero, es una pequeña y sencilla teoría que mantiene explícitamente las dimensiones de las muestras dentro de los límites de representación y ejecución.
- Segundo, el uso de las ventanas de datos queda justificado.
- Tercero, la estimación es consistente.
- Cuarto, el procedimiento es adaptivo y proporcionará estimaciones estables incluso en situaciones y zonas del espectro difíciles donde el rango del espectro es amplio.
- Quinto, proporciona un test de análisis de varianza para las componentes lineales.
- Y sexto, crea unas clases nuevas de estimaciones para determinados tipos de datos.

El desarrollo de este nuevo método se debe a la intención de superar los inconvenientes de los métodos empleados anteriormente. Así, hasta la aparición del método de Thomson, la mejor solución existente al problema de estimación del espectro no era completamente satisfactoria. En concreto, cuando se tenía una señal corta, o el espectro estaba mezclado, o bien el rango de los espectros era amplio, los problemas eran más que probables.

Dado que lo que se pretende es estudiar y analizar la solución al problema de estimación del espectro, asumiremos, para mayor simplicidad, datos gaussianos (o casi) constituidos por una secuencia finita de muestras igualmente espaciadas en el tiempo. Supondremos entonces que los datos estarán constituidos por N muestras contiguas, $x(0), x(1), x(2), \dots, x(N-1)$, de una señal temporal gaussiana, estacionaria, real, ergódica y de media cero.

3.2. LA ECUACIÓN INTEGRAL BÁSICA.

El motivo principal para estudiar el espectro de potencia de un proceso es que, para cualquiera que sea estacionario, existe una representación espectral de Crámer tal y como sigue a continuación:

$$x(t) = \int_{-1/2}^{1/2} e^{i2\pi ft} dZ(f)$$

y para todo instante t . Para procesos de media cero, se verifica:

$$E\{dZ(f)\} = 0$$

Su segundo momento, es decir, la densidad espectral de potencia o, simplemente, espectro del proceso $S(f)$ queda definido por:

$$S(f)df = E\{|dZ(f)|^2\}$$

Esta última expresión muestra los problemas de estimación de las propiedades estadísticas de $dZ(f)$, en particular y sobre todo, sus momentos.

El proceso $dZ(f)$ es un proceso de incremento ortogonal. Esto implica que, para frecuencias diferentes f y v , $dZ(f)$ y $dZ^*(v)$ son estadísticamente incorreladas, lo cual no significa independientes. Por simplicidad en la notación, es conveniente trasladar el origen de tiempo al centro del intervalo de observación. Si además cambiamos la definición de $dZ(f)$ por un factor de fase, podremos escribir lo siguiente:

$$x(t) = \int_{-1/2}^{1/2} e^{i2\pi v[t-(N-1)/2]} dZ(v)$$

Como deseamos estimar los estadísticos de $dZ(f)$ a partir de las N muestras contiguas, pasamos al dominio de la frecuencia usando la transformada discreta de Fourier:

$$y(f) = \sum_{t=0}^{N-1} e^{-i2\pi f[t-(N-1)/2]} x(t)$$

Lógicamente, podríamos obtener los datos de $y(f)$ mediante la transformada inversa:

$$x(t) = \int_{-1/2}^{1/2} e^{i2\pi f[n-(N-1)/2]} y(f) df$$

Debido a esta ‘equivalencia’, podemos usar indistintamente como datos $\{x(t)\}$ o $y(f)$.

A partir de las dos ecuaciones anteriores podemos obtener la siguiente expresión:

$$y(f) = \int_{-1/2}^{1/2} \sum_{t=0}^{N-1} e^{i2\pi(v-f)[t-(N-1)/2]} dZ(v)$$

donde es posible reconocer el sumatorio como el núcleo de Dirichlet. Así, deducimos ya la que será ecuación básica en la estimación del espectro:

$$y(f) = \int_{-1/2}^{1/2} \frac{\text{sen} N\pi(f-v)}{\text{sen} \pi(f-v)} dZ(v)$$

Puede interpretarse como una convolución que describe las “pérdidas de enventanado”, la “mezcla de frecuencias”...(consecuencias de usar la transformada finita de Fourier). Por todo ello, no hay razón obvia para esperar que los estadísticos de $y(f)$ se asemejen a los de $dZ(f)$.

Es importante destacar que el periodograma básico, $P_x(f)=|y(f)|^2$, no es un estadístico suficiente para los datos ya que implica un abandono de la información de fase (que es esencial). Consecuentemente, el periodograma es una mala elección como punto de partida para cualquier técnica de análisis de datos seria.

La importancia de esta ecuación es tal que merece que se le preste mayor atención a través de las siguientes observaciones:

- i) La insuficiencia del periodograma es heredada por toda estimación que se base en él. Obviamente, esto incluye tanto a los periodogramas suavizados (siendo irrelevante si el suavizado es realizado directamente sobre el periodograma o no) como a todas las descomposiciones basadas en muestras de la autocovarianza (ya que ésta es la transformada del periodograma).
- ii) El problema tiene mucho en común con el modelo lineal clásico

$$y = X'\beta + e$$

donde y representa las observaciones, X el modelo, β los coeficientes que se quieren estimar, y e el error entre el modelo hipotético y las observaciones. En el caso de la estimación espectral, la transformada de Fourier corresponde a las observaciones, el modelo es especificado por Dirichlet, y los coeficientes generan la estimación. En las regresiones clásicas y los análisis de problemas de varianzas, lo primero normalmente es llevar a cabo la

aproximación adecuada para resolver las ecuaciones (bien por mínimos cuadrados bien por cualquier otra técnica de aproximación) y luego, examinamos los estadísticos de los coeficientes estimados. Sin embargo, en los casos de estimación espectral se procede habitualmente como se indica a continuación: primero, se elevan al cuadrado las observaciones (es decir, se calcula el periodograma); segundo, se ignora el modelo; y tercero, usa las observaciones al cuadrado como posible solución. Si convertimos el modelo lineal analógico en un simple problema de regresión, estamos utilizando los propios datos para crear una recta de regresión. Esto puede ser una buena aproximación a la línea si la relación señal-ruido es bastante alta. Sin embargo, no da información sobre los coeficientes. Además, en la escala logarítmica necesaria para la estimación espectral, la muestra finita de Dirichlet es muy diferente de una delta de Dirac y, consecuentemente, la aproximación basada en usar el modelo y resolver las ecuaciones resultantes presenta algún inconveniente.

Para procesos con espectros típicos de ingeniería, el tamaño de las muestras debe ser extraordinariamente amplio para que el periodograma esté razonablemente insesgado. Así se ha podido comprobar con pruebas realizadas por Thomson utilizando 1.2 millones de puntos: los resultados fueron tan obviamente incorrectos que, al menos, no resultaban engañosos. En otras aplicaciones donde el proceso sea menos conocido, los errores pueden no ser tan obvios. Por ello, aunque las estimaciones que veremos son más difíciles y costosas en la computación que el periodograma, esto queda compensado porque evitamos las respuestas erróneas.

A lo largo de este procedimiento, debemos tener en cuenta que el problema esencial del análisis espectral es estimar las propiedades estadísticas de $dZ(v)$ como oposición a aquellas de $y(f)$. A pesar del esfuerzo empleado en las propiedades estadísticas de $y(f)$, no está claro que las estimaciones basadas en el periodograma sean de mucho interés.

La elección de una técnica para computar soluciones aproximadas depende básicamente de las características deseadas en dicha solución. Algunos de estos deseados aspectos son los siguientes:

- 1) La solución debería ser ‘local’, es decir, el espectro estimado en una frecuencia no debería depender fuertemente de detalles del espectro en frecuencias distantes.

- 2) Debe ser fácil de computar. Esto implica estabilidad numérica y procedimientos iterativos convergentes.
- 3) Los estadísticos de la solución deben poder ser caracterizados.
- 4) Dado que la estimación espectral suele ser un paso intermedio dentro de un análisis más completo, tanto la propia estimación como su logaritmo deben ser ‘buenas’ estimaciones.

El método más exitoso para estimar soluciones aproximadas es la expansión en autofunciones combinada con un criterio de error de mínimos cuadrados. Recordemos que las soluciones de la ecuación integral del tipo

$$y = K * z$$

donde K es un núcleo con funciones propias

$$\lambda_m \psi_m = K * \psi_m$$

estandarizadas porque cumplen

$$\int \psi_m \psi_n = \delta_{n,m}$$

es dada por

$$\hat{z} = \sum_m \frac{1}{\lambda_m} y_m \psi_m$$

donde los coeficientes y_m vienen dados por la fórmula de Fourier-Bessel

$$y_m = \int y \psi_m$$

Como los autovalores decaen generalmente de forma exponencial, tales soluciones no tienen uso en la práctica. Consecuentemente, la clase de soluciones realizables es la correspondiente a autovalores grandes. Estas soluciones pueden ser obtenidas simplemente truncando el sumatorio o bien, como veremos, por ponderación de los coeficientes de forma que podamos escribirlas como sigue

$$\hat{z} = \sum_{m=0} D(\lambda_m, y_m) \cdot \psi_m$$

donde $D(\lambda_m, y_m)$ es una función de ponderación o peso.

En un conjunto considerable de documentos, en el que el más reciente es el de Slepian, se ha demostrado que las funciones propias del núcleo de Dirichlet, conocidas como ‘prolate spheroidal wave functions’, son fundamentales en el estudio de los sistemas limitados en el tiempo y en la frecuencia. Usaremos dichas funciones como una base en la resolución de la ecuación integral en algunos intervalos de frecuencia (f-

$W, f+W)$. A estos intervalos se les llama *locales o dominio interior* mientras que al resto del dominio frecuencial se denomina *exterior*.

3.3.DISCRETE PROLATE SPHEROIDAL WAVE FUNCTIONS O SEQUENCES.

En este apartado estudiamos una serie de propiedades y características de estas funciones propias extraídas del documento de Slepian.

Las autofunciones son conocidas como “*discrete prolate spheroidal wave functions*” (DPSWF, Slepian taper o sequence). Se denotan por $U_k(N, W; f)$, $k = 0, 1, \dots, N-1$ y son las soluciones de la ecuación:

$$\int_{-W}^W \frac{\sin N\pi(f - f')}{\sin \pi(f - f')} \cdot U_k(N, W; f') df' = \lambda_k(N, W) \cdot U_k(N, W; f)$$

donde W cumple que $0 < W < 1/2$ (ancho de banda ‘local’).

Las funciones son ordenadas por sus autovalores:

$$1 > \lambda_0(N, W) > \lambda_1(N, W) > \dots > \lambda_{N-1}(N, W)$$

Los primeros $2NW$ autovalores son muy próximos a 1.

Es importante tener claro que, de todas las funciones que son transformadas de Fourier de una secuencia de índice limitado DPSWF, $U_0(N, W; f)$ tiene la mayor concentración de energía en $(-W, W)$. El grado de esta concentración viene dado por la siguiente expresión asintótica de Slepian:

$$1 - \lambda_k(N, W) \approx \frac{\sqrt{2\pi}}{k!} \left[\frac{8N \sin \pi W}{\cos^2 \pi W} \right]^{k+1/2} \left[\frac{1 - \sin \pi W}{1 + \sin \pi W} \right]^N, \text{ para } k \text{ pequeño y } N \text{ elevado}$$

$$1 - \lambda_k(N, W) \approx \frac{\sqrt{2\pi}}{k!} (8c)^{k+1/2} e^{-2c}, \text{ para } k < 2NW \text{ y } N \text{ aún mayor que antes con } N\pi W = c$$

Estas expresiones proporcionan la fracción de energía de la ventana espectral que queda fuera del lóbulo principal, es decir, del intervalo $(-W, W)$.

Las autofunciones $U_k(N, W; f)$ son *doblemente ortogonales*, lo que quiere decir que son ortogonales en $(-W, W)$

$$\frac{1}{\lambda_k(N, W)} \int_{-W}^W U_j(N, W; f) \cdot U_k(N, W; f) df = \delta_{j,k}$$

y ortonormales en $(-1/2, 1/2)$

$$\int_{-1/2}^{1/2} U_j(N, W; f) \cdot U_k(N, W; f) df = \delta_{j,k}$$

Las transformadas de Fourier de las DPSWF son conocidas como las “*discrete prolate spheroidal sequences*” (DPSS):

$$v_n^{(k)}(N, W) = \frac{1}{\varepsilon_k \lambda_k(N, W)} \int_{-W}^W U_k(N, W; f) e^{-i2\pi f \left[n - \frac{(N-1)}{2} \right]} df$$

válida para $k = 0, 1, \dots, N-1$ y todo n , y con $\varepsilon_k = \begin{cases} i, & k \text{ impar} \\ 1, & k \text{ par} \end{cases}$

Debido a la doble ortogonalidad, hay una segunda expresión para la transformada de Fourier de estas funciones:

$$v_n^{(k)}(N, W) = \frac{1}{\varepsilon_k} \int_{-1/2}^{1/2} U_k(N, W; f) e^{-i2\pi f \left[n - \frac{(N-1)}{2} \right]} df$$

válida para $n, k = 0, 1, \dots, N-1$.

Lógicamente, la transformada discreta de Fourier de la DPSS da como resultado la correspondiente DPSWF

$$U_k(N, W; f) = \varepsilon_k \sum_{n=0}^{N-1} v_n^{(k)}(N, W) e^{i2\pi f \left[n - \frac{(N-1)}{2} \right]}$$

Es interesante saber que las DPSS satisfacen una ecuación como la siguiente

$$\sum_{m=0}^{N-1} \frac{\sin 2\pi W(n-m)}{\pi(n-m)} \cdot v_m^{(k)}(N, W) = \lambda_k(N, W) \cdot v_n^{(k)}(N, W)$$

donde $\lambda_k(N, W)$ son los autovalores de una matriz de Toeplitz. Estas ecuaciones son fácilmente computadas para valores moderados de N .

Debemos tener en cuenta que, al igual que las DPSWF, las $\{v_n^{(k)}\}$ también son doblemente ortogonales, en concreto, son ortogonales en $(-\infty, \infty)$ y ortonormales en $[0, N-1]$.

3.4. ESTIMACIONES PROPIAS: COEFICIENTES.

Calcularemos estimaciones del espectro a partir de las soluciones aproximadas de la ecuación integral, utilizando para ello todo lo comentado anteriormente.

Empezamos considerando los coeficientes de expansión de dZ en $(f-W, f+W)$

$$Z_k(f) = \frac{1}{\sqrt{\lambda_k(N, W)}} \int_{-W}^W U_k(N, W; v) dZ(f + v)$$

Estos coeficientes $Z_k(f)$ son de considerable interés analítico en el sentido de que serían los que se obtendrían al pasar la secuencia completa a través de un filtro ideal paso de banda desde $f-W$ hasta $f+W$ (y antes de truncar al tamaño finito de la muestra). Se les llama *coeficientes idealizados*. En esta definición está implícito el resultado $E\{|Z_k(f)|^2\} = S$, cuando el espectro S es blanco.

A continuación realizamos una estimación de los coeficientes anteriores. Para ello, expandimos la transformada de Fourier de la muestra $y(f)$ sobre el intervalo $(f-W, f+W)$ en el mismo conjunto de funciones base

$$y_k(f) = \frac{1}{\lambda_k(N, W)} \int_{-W}^W U_k(N, W; v) y(f+v) dv$$

Éstos últimos se denominan *coeficientes estimados* o también *autocoficientes* de la muestra. Si usamos la ecuación integral para $y(f)$, que representa la proyección de dZ sobre y , junto con la ecuación integral que define las DPSWF, $y_k(f)$ puede ser expresado en términos de dZ como

$$y_k(f) = \int_{-1/2}^{1/2} U_k(N, W; \xi) dZ(f + \xi)$$

Esta expresión es útil para evaluar diversos momentos de las estimaciones, gracias a las propiedades de $dZ(f+\xi)$.

En resumen, las diferencias entre los coeficientes idealizados y los estimados son importantes. Notemos particularmente que el dominio de integración está limitado en banda en los coeficientes idealizados pero incluye el dominio principal de la estimación.

Podemos obtener otra forma alternativa para los coeficientes estimados si utilizamos la definición de las DPSS y escribimos la transformada de Fourier de $y(f)$ directamente en función de los datos. En este caso pueden expresarse como sigue:

$$y_k(f) = \sum_{n=0}^{N-1} x(n) \frac{v_n^{(k)}(N, W)}{\mathcal{E}_k} \cdot e^{i2\pi f \left[n - \frac{(N-1)}{2} \right]}$$

Vemos que el k -ésimo autocoficiente es la transformada discreta de Fourier de los datos multiplicado por la DPSS actuando como una ventana de datos. En la práctica, es más sencillo considerar los $v_n^{(k)}(N, W)$ reales y los $U_k(N, W; f)$ complejos para que así desaparezcan los factores de compensación de fase y podamos utilizar los algoritmos FFT para los cálculos.

NOTA.- en las ventanas de datos DPSS, al aumentar el orden, más se concentra la energía en los extremos del intervalo.

3.5. ESTIMACIÓN ESPECTRAL DE ALTA RESOLUCIÓN.

Para continuar el desarrollo de la solución de la ecuación integral, damos el siguiente paso al formar la estimación de alta resolución $\hat{z}_h(f; f_0)$, que es válida en $f_0 - W < f < f_0 + W$.

$$\hat{z}_h(f; f_0) = \sum_{k=0}^{K-1} U_k(N, W; f - f_0) \frac{y_k(f_0)}{\lambda_k(N, W)}$$

En esta expresión hemos usado coeficientes de peso unidad para los primeros $k = 2NW$ términos (normalmente se trabaja con estos términos).

Existe una *estimación espectral de alta resolución* correspondiente a la estimación anterior de dZ

$$\hat{S}_h(f; f_0) = \frac{1}{N} \cdot \left| \sum_{k=0}^{K-1} U_k(N, W; f - f_0) \frac{y_k(f_0)}{\lambda_k(N, W)} \right|^2$$

Dado que se trata de un cuadrado absoluto de una variable compleja gaussiana, está distribuida como un múltiplo de χ_2^2 y eso es inconsistente. Por ello, tomamos la media sobre el intervalo de frecuencia $(f_0 - W, f_0 + W)$:

$$\begin{aligned} \bar{S}(f_0) &= \frac{1}{2W} \int_{f_0 - W}^{f_0 + W} \hat{S}_h(f; f_0) df, \quad \bar{S}(f_0) = \sum_{k=0}^{K-1} |d_k(f_0) \cdot y_k(f_0)|^2 \Rightarrow \\ \Rightarrow \bar{S}(f_0) &= \frac{1}{2NW} \sum_{k=0}^{K-1} \frac{1}{\lambda_k(N, W)} \cdot |y_k(f_0)|^2 \end{aligned}$$

Como los autocorrelaciones $y_k(f)$ se calculan transformando los datos y multiplicándolos por la k -ésima ventana $v_n^{(k)}(N, W)$, sus cuadrados absolutos

$$\hat{S}_k(f) = |y_k(f)|^2$$

son *individualmente* estimaciones directas del espectro. Nos referimos a ellas como ‘prolate eigenspectrum estimate of order k’ o, simplemente, el *k-ésimo autoespectro*.

Estos autoespectros están normalizados como sigue, ya que las DPSS son ortonormales en $[0, N-1]$: $E\{\hat{S}_k(f)\} = S$ cuando el espectro S es blanco.

Esta $\hat{S}_0(f)$ es la mejor estimación directa conocida de un espectro; sin embargo, cuando se usaba sola por sí misma, como siempre ocurría en el pasado, esta estimación era suavizada para presentar resultados consistentes. La operación de suavizado tiene la indeseable característica de aumentar el ancho de banda efectivo a varias veces W , con

el consiguiente aumento de sesgo ($\rightarrow$ pérdida de resolución). Este efecto es conocido desde Bartlett y ha sido tratado por muchos desde entonces.

En resumen, debido al incremento de sesgo y pérdida de resolución intrínseca a tales procesos de suavizado por convolución, no son muy recomendables.

Sin embargo, en el procedimiento de estimación presentado aquí hay un elemento importante que nunca se empleó en el pasado: los autocorrelaciones $y_1(f), \dots, y_K(f)$. La presencia de estos términos tiene cuatro efectos fundamentales en la estimación:

- i) Como veremos, los diferentes autocorrelaciones son casi incorrelados para espectros localmente blancos en grado razonable. Por ello, y como individualmente son complejos gaussianos, sus cuadrados absolutos están distribuidos como χ^2_2 y, consecuentemente, la estimación tiene unos $2K$ grados de libertad que, para W fijado, equivale a $4NW$ grados de libertad.
- ii) Hacen que la estabilidad sea alcanzada sin la pérdida de resolución asociada al suavizado por convolución.
- iii) La eficiencia de la estimación es buena, normalmente, mejor que el 80%. Esto es debido a que la información que antes se perdía por utilizar $\hat{S}_0(f)$ sola, ahora no se pierde (gracias al uso de los autocorrelaciones).
- iv) Hacen que las ventanas espectrales efectivas asociadas a la estimación estabilizada $\bar{S}(f)$ tengan una forma casi ideal.

3.6. PROPIEDADES DE MUESTREO. LOS MOMENTOS.

En este apartado analizaremos los momentos de menor orden de las estimaciones que estamos considerando en nuestro proyecto.

Como $E\{dZ(f)\} = 0$, resulta inmediato que $E\{y_k(f)\} = 0$. El segundo momento o momento de segundo orden será

$$E\{y_j(f)y_k^*(f')\} = E\left\{\int_{-1/2}^{1/2} \int_{-1/2}^{1/2} U_j(N, W; \zeta) \cdot U_k(N, W; \xi) dZ(f + \zeta) dZ^*(f' + \xi)\right\}$$

que, tras aplicar las propiedades del proceso ortogonal dZ , se convierte en

$$E\{y_j(f)y_k^*(f')\} = E\left\{\int_{-1/2}^{1/2} U_j(N, W; f + \zeta) \cdot U_k(N, W; f' + \xi) d\zeta\right\}$$

Para el caso especial $j=k$, tendremos:

$$E\{\hat{S}_k(f)\} = |U_k(N, W; f)|^2 * S(f)$$

De esta manera, el valor esperado del k-ésimo autoespectro $\hat{S}_k(f)$ es la convolución del espectro real $S(f)$ con la k-ésima ventana espectral $|U_k(N, W; f)|^2$.

Por otro lado, como las DPSWF tienen casi toda su energía en $(-W, W)$, salvo una fracción $1-\lambda_k$, los autoespectros de bajo orden son buenas estimaciones desde el punto de vista del sesgo.

Las estimaciones individuales de los autoespectros, $\hat{S}_k(f) = |y_k(f)|^2$, son inconsistentes y están distribuidas como múltiplo de χ^2_2 . Sin embargo, su media está distribuida aproximadamente como un múltiplo de χ^2_{4NW} y así, sí es consistente.

Resulta conveniente dividir la integral de la estimación en dos partes:

- Sesgo local:
$$\tilde{S}_k(f) = \int_{-W}^W |U_k(N, W; \xi)|^2 S(f + \xi) d\xi$$
- Sesgo de banda-ancha:

$$B_k(f) = \left| \int_{-1/2}^{1/2} U_k(N, W; \xi) dZ(f + \xi) - \int_{-W}^W U_k(N, W; \xi) dZ(f + \xi) \right|^2$$

El valor esperado del sesgo de banda-ancha es

$$E\{B_k(f)\} = \int_{-1/2}^{1/2} U_k^2(N, W; \xi) S(f + \xi) d\xi - \int_{-W}^W U_k^2(N, W; \xi) S(f + \xi) d\xi$$

Es posible acotar este valor a través de la desigualdad de Cauchy, obteniendo el siguiente resultado:

$$E\{B_k(f)\} \leq (1 - \lambda_k(N, W))^2 \sigma^2$$

3.7. SESGO LOCAL Y CUADRÁTICO.

En muchos casos, particularmente cuando se trabaja con series muy cortas, no es realista asumir que el espectro varíe lentamente en $f-W$, $f+W$ y, consecuentemente, los términos de sesgo local pueden ser significativos. Para evaluar este efecto suponemos que el espectro puede extenderse como una serie de Taylor alrededor de f_0 . De este modo, el valor esperado del k-ésimo autoespectro puede escribirse como:

$$E\{\hat{S}_k(f_0)\} = S(f_0) + \frac{S'(f_0)}{1!} \int_{-1/2}^{1/2} \xi |U_k(\xi)|^2 d\xi + \frac{S''(f_0)}{2!} \int_{-1/2}^{1/2} \xi^2 |U_k(\xi)|^2 d\xi + \dots$$

Por simetría, los coeficientes de S' , S''' , ... se anulan entre sí, mientras que los de S'' y los términos de mayor orden par no lo hacen. Como domina el sesgo asintótico, sólo

consideraremos el término cuadrático, es decir, el coeficiente de $S''(f_0)$, que escribimos en forma normalizada como sigue:

$$G_k(N, W) = N^2 \int_{-1/2}^{1/2} \xi^2 |U_k(\xi)|^2 d\xi$$

Con esta definición, el término asintóticamente dominante del sesgo local del k -ésimo autoespectro se convierte en

$$\frac{1}{N^2} \frac{S''}{2!} G_k(N, W)$$

A causa de la gran concentración en las DPSWF, el valor de esta integral surge realmente como resultado de integrar en $(-W, W)$. Por otro lado, debido a que las ventanas espectrales reales $|U_k(N, W, \xi)|^2$ se ensanchan conforme aumenta k , está claro que $G_k(N, W)$ debería incrementarse también con el aumento de k . El hecho de que la energía sea prácticamente 1 en $(-W, W)$ puede ser utilizado para obtener una aproximación a esta integral como la que se indica a continuación:

$$G_k(N, W) \approx (NW)^2 / 3$$

Aunque esta aproximación es pobre para los términos G_k individuales, resulta bastante buena para el valor medio de éstos cuando la ventana espectral ponderada (media) se aproxima a una ventana rectangular. Ya que $NW=c/\pi$, podemos afirmar con claridad que, asintóticamente, el término de sesgo o desviación local cuadrático decrece con N^{-2} .

Como $G_k(N, W)$ aumenta con k , es razonable plantearse si la desviación cuadrática asociada con la aproximación de los autocorrelaciones es mayor o no que la obtenida por $\hat{S}(f_0)$ y los procesos de convolución convencionales. Para hacer tal comparación, se fijan los grados de libertad de la estimación y usamos luego la ventana de Papoulis (que fue optimizada con respecto al sesgo cuadrático) para suavizar el autoespectro de orden cero. Esto resultaría en coeficientes de desviación cuadrática de 2.64 y 18.5 para procesos de convolución con 4 y 10 grados de libertad respectivamente. Sin embargo, con la estimación de los autocorrelaciones, se obtendría 1.287 y 3.282 respectivamente. De esta forma, podemos apreciar que la desviación cuadrática de la aproximación de autocorrelaciones es mucho menor que la de los procesos convolucionales. También veremos más adelante que la eficiencia de la varianza y la capacidad de estimación de estos últimos procesos son considerablemente peores que los que se obtienen en la aproximación del autoespectro.

3.8. DISTRIBUCIONES DE LAS ESTIMACIONES DE AUTOESPECTROS.

Además de su valor esperado, también es de interés la distribución de $\hat{S}_k(f)$. En primer lugar, si el proceso es gaussiano, $\hat{S}_k(f)$ estará distribuida según la función chi-cuadrado con dos grados de libertad, es decir, χ_2^2 . Es más, incluso cuando los datos originales no son del todo gaussianos, el filtrado adicional implícito en el procedimiento de estimación hará que los coeficientes $y_k(f)$ sean de una forma compleja tal que sus cuadrados están distribuidos también como χ_2^2 .

Un problema más interesante surge de la división en componentes de sesgo o desviación local y de banda-ancha cuando ésta última es la dominante. En tales casos, la distribución local aparece significativamente diferente de aquellas ocasiones en las que domina el sesgo local. Así, aunque en conjunto la distribución es proporcional a χ_2^2 , para una determinada muestra es más apropiado pasar a un modelado en términos de una función de distribución de chi-cuadrado no centrada, con un parámetro libre. Si identificamos este parámetro libre no centralizado con la componente de desviación de banda-ancha, está claro que, en las regiones donde este término domina, la variabilidad de la estimación será mucho menor que lo normal.

Otro caso especial se produce cuando $f=f'$, siendo $j \neq k$:

$$E\{y_j(f)y_k^*(f)\} = \int_{-1/2}^{1/2} U_j(N,W;f+\xi)S(\xi)U_k(N,W;f+\xi)d\xi$$

Utilizando de nuevo las propiedades de los autovalores podemos acotar la contribución del dominio exterior a cantidades exponencialmente pequeñas en NW de forma que, en la mayoría de los casos, pueda ser ignorada dicha contribución. Luego, si el espectro dentro de $(f-W, f+W)$ es razonablemente llano, los coeficientes serán incorrelados debido a las propiedades de ortogonalidad de las DPSWF. Sin embargo, si el espectro es muy puntiagudo o cambia rápidamente en esta zona, la correlación puede ser significativa.

El último caso que consideramos por su interés es aquél en el que $f'=f+\Delta f$ de modo que las DPSWF están desplazadas y no son tan ortogonales. Si asumimos que el espectro es blanco, tenemos el siguiente coeficiente de correlación:

$$\Lambda_{jk}(\Delta f) = \int_{-1/2}^{1/2} U_j(N,W;f)U_k(N,W;f+\Delta f)df$$

3.9. ESTIMACIÓN ESPECTRAL POR PROMEDIO O PONDERACIÓN ADAPTIVA.

Aún cuando las propiedades de desviación de los autoespectros de bajo orden son generalmente excelentes debido al decremento de los autovalores conforme k se acerca a $2NW$, las características de polarización de las sucesivas estimaciones pueden degradarse. Consecuentemente, en las regiones donde el espectro es pequeño, las estimaciones de mayor orden serán menos fiables que los autoespectros de bajo orden por lo que tendrán que ser ponderados ‘a la baja’.

Si vemos la contribución a la estimación del k -ésimo espectro de la región $(f-W, f+W)$ como “señal”, las contribuciones del resto de la frecuencia como “ruido”, y el orden k como “frecuencia”, el proceso de peso o ponderación adaptivo se asemeja al filtrado de Wiener, teniendo como única diferencia las funciones base.

En este proceso se introducen una serie de funciones de ponderación $d_k(f)$ que, tal como los coeficientes que modifica, dependen de la frecuencia. Estas funciones minimizan el error de mínimos cuadrados entre los coeficientes ideales y los reales:

$$Z_k(f) - y_k(f)d_k(f) = \frac{1}{\sqrt{\lambda_k(N, W)}} \cdot \int_{-W}^W U_k(N, W; \xi) dZ(f + \xi) - d_k(f) \int_{-1/2}^{1/2} U_k(N, W; \xi) dZ(f + \xi)$$

Agrupando las regiones de integración, podemos deducir de esta expresión que el error está constituido por la suma de dos términos: uno, definido en $(-W, W)$ y el otro sobre el resto del dominio principal. Como ambas integrales se expresan en función de dZ , ambas son independientes y consecuentemente el error de mínimos cuadrados es simplemente la suma de los cuadrados de los dos términos. Usando entonces las propiedades de ortogonalidad de dZ y asumiendo que el espectro varía lentamente en $(-W, W)$, podemos aproximar el error de mínimos cuadrados de la primera integral por la siguiente expresión:

$$E \left\{ \left| \int_{-W}^W U_k(N, W; \xi) dZ(f + \xi) \right|^2 \right\} \approx \lambda_k(N, W) S(f)$$

La segunda integral es el sesgo o desviación de banda-ancha $B_k(f)$ de la estimación del k -ésimo autoespectro. Esta función depende del conjunto de características del espectro en el dominio exterior. Ya que el problema es la estimación del espectro, esta información debe ser aproximada por la muestra y usaremos un proceso de dos pasos. Primero, considerando su valor medio sobre todas las frecuencias, obtenemos una buena estimación inicial:

$$\int_{-1/2}^{1/2} E\{B_k(f)\} = (1 - \lambda_k(N, W))^2 \sigma^2$$

donde σ^2 es el proceso de varianza:

$$\sigma^2 = \int_{-1/2}^{1/2} S(f) df$$

Combinando estas dos expresiones y minimizando el error de mínimos cuadrados respecto a $d_k(f)$, encontramos la aproximación óptima para la función de ponderación o promediado:

$$d_k(f) \approx \frac{\sqrt{\lambda_k(N, W)} S(f)}{\lambda_k(N, W) S(f) + E\{B_k(f)\}}$$

La media correspondiente a la función de densidad espectral:

$$\hat{S}(f) = \frac{\sum_{k=0}^{K-1} |d_k(f)|^2 \hat{S}_k(f)}{\sum_{k=0}^{K-1} |d_k(f)|^2}$$

Como en esta expresión aparecen valores del espectro y del sesgo de banda-ancha (y éstos son, obviamente, desconocidos), los sustituimos por estimaciones.

La definición es recursiva y la estimación del espectro resultante es una solución de la ecuación:

$$\sum_{k=0}^{K-1} \frac{\lambda_k(\hat{S}(f) - \hat{S}_k(f))}{[\lambda_k \hat{S}(f) + \hat{B}_k(f)]^2}$$

y así se debe proceder en el intervalo acotado por el menor y máximo valor de $\hat{S}_k(f)$.

En la práctica, la ecuación es resuelta iterativamente usando la media de las dos estimaciones de menor orden como valor de partida. Se demuestra que la convergencia es bastante rápida (pueden usarse estimaciones distantes en menos de un 5%).

Un parámetro útil en el procedimiento de estimación que determina la estabilidad de estas estimaciones es:

$$v(f) = 2 \sum_{k=0}^{K-1} |d_k(f)|^2$$

Podemos apreciar que $v(f)$ es una función de la desviación.

Si $v(f)/2k$ es significativamente menor que 1, entonces o el valor W es demasiado pequeño o debería ser usado un prefiltrado adicional. Como veremos, combinando $v(f)$ con el coeficiente de eficiencia de la varianza, obtendremos una ‘regla de parada’ en la iteración.

Una vez que ha sido hecha esta estimación inicial, el espectro estimado puede utilizarse para mejorar la estimación del sesgo de banda-ancha $B_k(f)$. Para este propósito definimos una *ventana de retraso*:

$$L_k^{(o)}(\tau) = \int_{-1/2}^{1/2} e^{i2\pi\tau\xi} U_k(N, W; \xi)^2 d\xi - \int_{-W}^W e^{i2\pi\tau\xi} U_k(N, W; \xi)^2 d\xi$$

de forma que

$$\hat{B}_k(f) = \sum_{\tau} e^{-i2\pi f\tau} L_k^{(o)}(\tau) R^{(c)}(\tau)$$

donde $R^{(c)}(\tau)$ es la función de autocovarianza correspondiente a la estimación del espectro al comienzo de la iteración actual.

Para realizar eficientemente estos cálculos se emplean algoritmos basados en la transformada rápida de Fourier (FFT). Debemos tener en cuenta los dos hechos siguientes: en primer lugar, el resultado del procedimiento de estimación es una secuencia de autocovarianza extrapolada, que es distinta de cero para retrasos $|\tau| > N$; segundo, al igual que en todas las operaciones numéricas que usan la función de autocovarianza, es aconsejable el uso de aritmética de doble precisión para muchos espectros de interés.

Por último, podemos reseñar que la estimación de $\hat{S}(f)$ por el método de máxima-probabilidad culmina en resultados muy parecidos a los mostrados anteriormente. Para ser más precisos, podemos afirmar que los pesos efectivos procedentes de la técnica de máxima-probabilidad son un poco más altos que aquellos obtenidos por el método de mínimos cuadrados. Sin embargo, como los resultados en las estimaciones espectrales raramente difieren en más de un 1 ó 2 por ciento y dado que la técnica de mínimos cuadrados es mucho más simple, omitimos más detalles sobre el otro método.

3.10.RELACIONES ENTRE ESTIMACIONES PROPIAS Y EL PERIODOGRAMA.

Hay una interesante relación entre las estimaciones de los espectros propios y el periodograma, que muestra los problemas de desviación de las estimaciones basadas en el periodograma en términos de los coeficientes de ponderación.

Empezamos por extender el núcleo de Dirichlet utilizando una fórmula bilineal:

$$\frac{\text{sen}N\pi(f - f')}{\text{sen}\pi(f - f')} = \sum_{k=0}^{N-1} U_k(N, W; f) \cdot U_k(N, W; f')$$

y, sustituyendo f y f' por $f-f_0$ y $f'-f_0$ respectivamente:

$$\frac{\sin N\pi(f-f')}{\sin \pi(f-f')} = \sum_{k=0}^{N-1} U_k(N, W; f-f_0) \cdot U_k(N, W; f'-f_0)$$

Multiplicando ambos lados de la ecuación anterior por $dZ(f')$ e integrando (se emplea la ecuación integral básica) podemos obtener

$$y(f) = \sum_{k=0}^{N-1} U_k(N, W; f-f_0) y_k(f)$$

Elevando al cuadrado y multiplicando por $1/N$ ambos lados obtenemos una expresión para el periodograma válida en $|f-f_0| < W$

$$I(f) = \frac{1}{N} \cdot \left| \sum_{k=0}^{N-1} U_k(N, W; f-f_0) y_k(f) \right|^2$$

Observamos que todos los autocorrelaciones aparecen en esta expresión con peso unidad. Si ahora consideramos un caso semejante pero con anchura $2W$ tendremos

$$I_W(f_0) = \frac{1}{2NW} \cdot \sum_{k=0}^{N-1} \lambda_k(N, W) \hat{S}_k(f_0)$$

Esto muestra que el periodograma suavizado es equivalente a una media ponderada de los espectros propios con pesos independientes de la frecuencia. En resumen, la media o promedio es de todos los autoespectros, no sólo de aquellos $2NW$ cuyos grandes autovalores implican que la información contenida en los coeficientes es de origen local.

Si se comparan los pesos aplicados a cada autoespectro con los obtenidos por el criterio de mínimos cuadrados, sólo son iguales cuando $S(f)$ es blanco: $S(f) = \sigma^2$

La relación entre los pesos efectivos y los obtenidos por el criterio de mínimos cuadrados es la siguiente:

$$w_k(f) = |d_k(f)|^2 = \frac{\lambda_k(N, W) S^2(f)}{[\lambda_k(N, W) S(f) + \sigma^2 (1 - \lambda_k(N, W))]^2}$$

3.11. VARIANZA Y EFICIENCIA.

La caracterización de las estimaciones espectrales es un asunto complejo, sobre el que se han escrito muchos informes y documentos y sobre el que, a pesar de lo dicho, hay poco convenio.

Para estudiar las propiedades de muestreo de estas estimaciones, asumimos que los datos son gaussianos. Empleando la fórmula del cuarto momento de un proceso gaussiano, se obtiene

$$\text{cov}\{\hat{S}_j(f+g), \hat{S}_k(f-g)\} = \left| \int_{-1/2}^{1/2} S(\xi) U_j(f+g-\xi) U_k(f-g-\xi) d\xi \right|^2 + \left| \int_{-1/2}^{1/2} S(\xi) U_j(f+g-\xi) U_k(f-g+\xi) d\xi \right|^2$$

En esta expresión, el último término es grande únicamente cerca del origen y de la frecuencia de Nyquist. Sin embargo, el primer término suele ser dominante y presenta una serie de características de interés:

- Primero, si la frecuencia de separación $2g$ es mayor que $2W$ de manera que los lóbulos centrales de ambas estimaciones no se superpongan, el intervalo de integración puede dividirse y la covarianza acotada en función a los autovalores. Tales acotaciones son pequeñas típicamente.
- Segundo, si la frecuencia de separación es menor que $2W$ tal que los lóbulos centrales ahora se superponen, la covarianza depende del espectro principalmente en $(f-g-W, f+g+W)$ y la contribución de las componentes espectrales fuera de este entorno puede ser acotada por las propiedades de los autovalores en una manera similar a la considerada anteriormente.

Si el espectro alrededor de f es constante o lineal, la correlación en la frecuencia de separación $\Delta=2g$ es

$$\Lambda_{jk}^2(\Delta) = |U_j(\Delta) * U_k(\Delta)|^2 = \left| \sum v_n^{(j)} v_n^{(k)} e^{in\Delta} \right|^2$$

En particular, cuando Δ es cero, la correlación entre autoespectros es

$$\text{corr}\{\hat{S}_j(f), \hat{S}_k(f)\} = \left| \int_{-1/2}^{1/2} U_j(\xi) U_k(\xi) d\xi \right|^2$$

que es cero para $j \neq k$. Incluso para casos en los que Δ no es cero, las correlaciones cruzadas en la frecuencia de off-set son bastantes pequeñas. Así, por ejemplo, con $c=4\pi$ tenemos $\Lambda_{01}^2(1.59/T) = 0.366$ y $\Lambda_{02}^2(2.2/T) = 0.264$ como valores máximos.

El concepto de eficiencia se halla íntimamente relacionado al de varianza de un estimador. La eficiencia de una estimación espectral engloba no sólo a la varianza en una frecuencia determinada sino también a la covariabilidad de las estimaciones en diferentes frecuencias, de lo contrario la eficiencia podría ser incrementada por un

suavizado adicional. Llevando esta idea al límite y usando el espectro integrado como una estimación de varianza para los datos incorrelados, obtenemos una medida de la eficiencia total de la estimación. Para una estimación de un autoespectro, la eficiencia de la varianza es

$$\xi_k = \frac{1}{N \sum_{n=0}^{N-1} [v_n^{(k)}(N, W)]^4}$$

Mientras que para la media de las primeras k estimaciones se tendrá:

$$\Xi_k = \frac{1}{N \sum_{n=0}^{N-1} \left[\frac{1}{K} \sum_{k=0}^{K-1} [v_n^{(k)}(N, W)]^2 \right]^2}$$

Dichas eficiencias de las estimaciones pueden ser altas; por ejemplo, con $c=3\pi$, $K=4$ y $c=4\pi$, $K=5$, tienen una eficiencia superior al 80% y todavía proporcionan una excelente protección frente al sesgo. Cuando se utiliza la ponderación adaptiva, la desviación de estas estimaciones es mucho menor debido a la baja pérdida del autoespectro inicial. Puede demostrarse que, si se usan más de las primeras $2c/\pi$ estimaciones, no sólo cae rápidamente la protección frente a la desviación sino también la eficiencia.

Es muy importante recalcar que la eficiencia de la varianza no debería ser considerada como un criterio de fuerza en la comparación de estimas espectrales por las siguientes razones:

- Es válido estrictamente para procesos de ruido blanco.
- Ignora el sesgo y sus consecuencias.

De este modo, si uno juzgara las estimaciones exclusivamente según el criterio de la eficiencia, tendríamos que quedarnos como única forma admisible con el periodograma y sus variantes.

Lo más importante a resaltar es que si utilizamos los procedimientos convencionales, es decir, si la estimación se realiza siguiendo los pasos siguientes:

1. Se multiplican los datos por una buena ventana como $v_n^{(0)}(N, W)$ o una aproximación de Káiser,
2. Se transforman los datos enventanados y se eleva al cuadrado,
3. Se utiliza un proceso de convolución para suavizar y estabilizar el resultado,

la eficiencia no puede exceder $\Xi_1(c)$ para ninguna técnica de suavizado. Así para $c=3\pi$ y 4π la aproximación del autoespectro es dos veces más eficiente que estimaciones enventanadas convencionales.

Con respecto a la eficiencia, recordemos que varias estimaciones espectrales óptimas basadas en formas cuadráticas $\hat{S}(f) = X^T A(f) X$, han sido ya propuestas donde se asumía como conocido el espectro para poder calcular la función óptima de ponderación, $A(f)$. Dado que las estimaciones no ponderadas son también formas cuadráticas, resulta razonable preguntarse cómo las mencionadas inicialmente pueden ser más eficientes que estas más antiguas. Hay varias razones:

1. $A(f)$ dependía solamente de las muestras individuales de autocovarianzas y no se permitía que fuese más general.
2. Óptimamente, fue establecida sólo asintóticamente.
3. Se asumía que el periodograma era insesgado.

Como las estimaciones de $A(f)$ resultan ser desconocidas, al no estar restringidas a los periodogramas, sus buenos resultados no deben tomarse demasiado en serio.

Por otro lado, la estimación ponderada no es una simple forma cuadrática sino una combinación racional de ellas.

Podemos anticipar que, aunque el periodograma corresponde a un máximo de la función de probabilidad, éste no es el máximo global.

Combinando los resultados anteriores con la estimación de estabilidad $v(f)$, se tiene una medida más completa de la eficiencia

$$eff \approx \overline{v(f)} \Xi_K(c) N$$

que puede usarse para comparar la efectividad del prefiltro (que reduce N), de variar W y K , ...