

Tema 5

Estimación espectral

5.1. Introducción

La mayoría de las señales biomédicas son el resultado de procesos que tienen lugar en el dominio del tiempo. No obstante, en algunos casos es conveniente realizar su estudio en el dominio de la frecuencia, tanto en casos deterministas como estocásticos, por lo cual es necesario obtener la función densidad espectral de potencias (PSD).

En EEG, el análisis espectral se ha utilizado tanto en clínica como en investigación. Los objetivos han sido la clasificación de desórdenes neurológicos, clasificación automática de estados de sueño y grados de anestesia. En EMG, se ha utilizado para estudiar la fatiga muscular. En señales de voz, se ha utilizado para diagnosticar malfunciones de la laringe.

El valor real del PSD no puede, en principio, obtenerse. La señal de entrada está limitada en el tiempo, es no estacionaria y tiene ruido superpuesto, por lo que es necesario estimar el PSD a partir de un registro de cierta cantidad de datos. Los primeros métodos se basaron en la estimación de la transformada de Fourier (FT). Un paso importante en estimación espectral moderna fue el trabajo de Wiener, que establecía las bases teóricas del tratamiento de procesos estocásticos. Wiener y Khinchin, de forma independiente, demostraron que la FT relaciona la función autocorrelación de un proceso estacionario y su PSD. Se denomina generalmente relación de Wiener-Khinchin.

Con anterioridad a la introducción del algoritmo de la transformada rápida de Fourier (FFT) en 1965, el método usual para estimar el PSD era la implementación de la relación de Wiener-Khinchin sugerida por Blackman y Tukey. Según este método, los coeficientes de autocorrelación se estiman utilizando una secuencia de datos eventanados. Se realiza la transformada de Fourier de esta correlación eventanada para obtener el PSD. Este método suele denominarse también método indirecto.

A partir de la introducción de la FFT, que implicaba la disposición de un algoritmo computacionalmente eficiente para el cálculo de la transformada discreta de Fourier (DFT), se utiliza de forma intensiva la aproximación directa (periodograma), que se

obtiene como la magnitud al cuadrado de la DFT obtenida mediante FFT y aplicada directamente sobre los datos (previamente enventanados). Otros métodos como técnicas de análisis ARMA y procedimientos más especializados como la descomposición armónica de Pisarenko, el método de Prony, etc., proporcionan estimadores con mejores características estadísticas, aunque pueden ser computacionalmente menos eficientes que los basados en la DFT. Básicamente podemos distinguir tres grandes clases:

- Métodos no paramétricos: basados en la DFT.
- Métodos paramétricos: basados en modelos ARMA y variantes.
- Métodos tiempo-frecuencia: que asumen la no estacionariedad de las señales con el tiempo.

5.2. Métodos no paramétricos

Periodograma

La DFT de una función muestreada $x(n)$ viene dada por:

$$X_k = T \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}$$

donde T es el periodo de muestreo. Su **densidad espectral de energía** (ESD) vendrá dada por:

$$S_k = |X_k|^2$$

y suele denominarse también periodograma, siendo un estimador espectral de la señal.

Si el proceso $x(n)$ es un proceso estocástico en lugar de uno determinista, lo que implica que no se trata de una onda con energía finita, sino infinita, el parámetro de interés no es la energía sino la potencia (promediado temporal de la energía). Para este caso, se define el periodograma como **densidad espectral de potencia** (PSD) según:

$$P_k = |X_k|^2 / (N \cdot T)$$

P_k es idéntico a la densidad espectral de energía S_k excepto por la división del intervalo de tiempo $N \cdot T$ segundos, lo que lo convierte en una densidad espectral de potencias.

El periodograma no es un estimador consistente, ya que su varianza no tiende a cero cuando la longitud del registro tiende a infinito. Para mejorar esta característica,

Welch propuso un método según el cual se dividía el registro de N puntos original en segmentos de M puntos solapados entre sí L muestras. Si $L=M$, entonces $N=(K+1)M$, donde K sería el número total de segmentos. Posteriormente se aplica una ventana a cada segmento, y finalmente se calcula el periodograma para cada segmento enventanado. El periodograma final se obtiene promediando todos los periodograma parciales.

Cross-spectrum

La relación de Wiener-Khinchin relaciona el espectro de una señal con la TF de su función autocorrelación. Cuando se aplica al caso de dos señales, $x(t)$ e $y(t)$, se obtiene la función densidad espectral cruzada (*cross-spectrum*), que se define como la TF de la función correlación cruzada.

El espectro cruzado permite realizar un estudio equivalente en el dominio de la frecuencia al realizado en el dominio del tiempo mediante la función de correlación cruzada. Se puede demostrar que el espectro cruzado está acotado por:

$$|S_{xy}(w)|^2 \leq S_x(w)S_y(w)$$

Se define a partir de aquí la función de coherencia, γ_{xy}^2 , dada por:

$$\gamma_{xy}^2(w) = \frac{|S_{xy}(w)|^2}{S_x(w)S_y(w)} \leq 1$$

Cuando la función vale 1 para todas las frecuencias, se dice que $x(t)$ e $y(t)$ son totalmente coherentes. Cuando para alguna frecuencia w_0 vale cero, se dice que son incoherentes a esa frecuencia. Si ambas señales son estadísticamente independientes, la función de coherencia vale cero para todas las frecuencias.

Esta función se ha aplicado, por ejemplo, al análisis de EEG, obtenido simultáneamente en posiciones distintas del cráneo, para la investigación de asimetrías cerebrales, localización de focos epilépticos, el estudio de relaciones entre la actividad cortical y talámica, etc.

Conclusiones

Como conclusión, podemos decir que los métodos no paramétricos presentan las siguientes ventajas e inconvenientes:

- Ventajas:
 - Son computacionalmente eficientes.
 - El PSD obtenido es directamente proporcional a la potencia de las componentes espectrales del proceso.
- Inconvenientes:
 - Introducción de distorsión en el espectro debido a la aparición de lóbulos laterales asociados con el enventanado de la señal. Estos lóbulos pueden enmascarar la contribución de armónicos de baja amplitud situados junto a armónicos de gran amplitud.
 - Resolución de frecuencia limitada por la duración de la señal disponible, independientemente de las características de la misma.
 - Necesidad de algún tipo de promediado para obtener estimaciones estadísticamente consistentes.

5.3. Métodos paramétricos

Los métodos paramétricos permiten una estimación espectral de alta resolución, y se han mostrado como una alternativa interesante a la FT en el análisis de bioseñales, como, por ejemplo, en el caso de series temporales RR.

En el caso más genérico (modelo ARMAX), la señal muestreada (la secuencia) $\{S_k\}$ se supone la salida de un sistema lineal con una entrada (no accesible) $\{U_k\}$ y con ruido aditivo superpuesto $\{n_k\}$. La señal $\{S_k\}$ es la solución de la ecuación de diferencias:

$$S_k = -\sum_{i=1}^p a_i S_{k-i} + \sum_{i=0}^q b_i U_{k-i} + \sum_{i=0}^m f_i r_{k-i}$$

donde:

$$\begin{aligned} A(z^{-1}) &= \sum a_i z^{-i} ; & a_0 &= 1 \\ B(z^{-1}) &= \sum b_i z^{-i} \\ F(z^{-1}) &= \sum f_i z^{-i} \end{aligned}$$

El modelo ARMA (*autoregressive moving average*) se obtiene de la ecuación en diferencias suponiendo que no hay ruido externo:

$$S(z) = [B(z^{-1}) / A(z^{-1})] U(z)$$

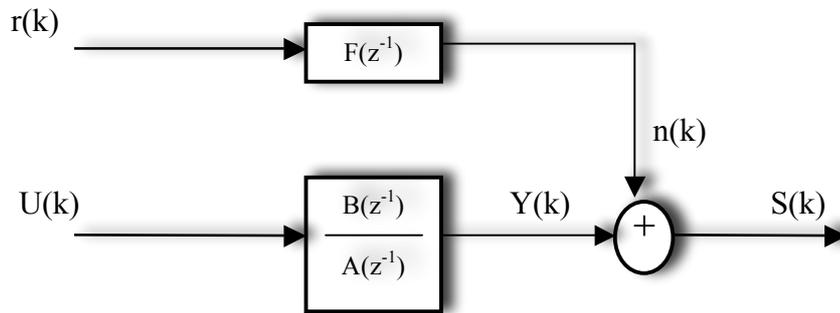


Figura 5.1. Estructura de un proceso ARMAX.

Este modelo se conoce como ARMA de orden (p,q) , denominándose también modelo de polos y ceros. En estos modelos, la muestra k -ésima se expresa como combinación lineal de las p muestras anteriores de la salida y de las $(q+1)$ muestras anteriores de la entrada.

De la ecuación anterior se derivan otros dos modelos frecuentemente usados: el AR (autorregresivo) y el MA (promediado móvil):

El modelo AR sólo contiene polos, derivado del ARMA haciendo $B(z^{-1}) = 1$:

$$S(z) = [1 / A(z^{-1})] U(z)$$

El modelo MA sólo contiene ceros, derivado del ARMA haciendo $A(z^{-1}) = 1$:

$$S(z) = B(z^{-1}) U(z)$$

La idea básica de modelado predictivo lineal es suponer que la secuencia $\{S_n\}$ es la salida de un sistema lineal, por lo que se puede expresar dicha secuencia (de forma paramétrica) por medio de los parámetros del sistema. Dada la secuencia $\{S_n\}$, el problema pues es identificar dichos parámetros. Puesto que la entrada es inaccesible, no puede aplicarse algoritmos para identificación de sistemas directamente. En aplicaciones de modelización de señal y de estimación espectral, se supone que al entrada al modelo es ruido blanco (espectro plano) con media cero y varianza constante.

Cualquier valor de la salida puede expresarse en términos de una suma pesada de valores previos de entrada y salida, representada por la ecuación en diferencias:

$$S_n = -\sum_{k=1}^m a_k S_{n-k} + \sum_{k=0}^q b_k S_{n-k}; b_0 = 1$$

Esta ecuación puede escribirse en términos de z , obteniéndose la función de transferencia del modelo:

$$H(Z) = \frac{S(z)}{X(z)} = \frac{1 + \sum_{k=1}^q b_k Z^{-k}}{1 - \sum_{k=1}^m a_k Z^{-k}}$$

El denominador de esta ecuación contiene los términos autorregresivos, que definen los polos de la función de transferencia, mientras el numerador está compuesto por los términos de promediado móvil, que definen sus ceros. En general, cuando queremos modelar una señal con este tipo de funciones, los modelos ARMA, que vienen definidos por ceros y polos, son siempre los que presentan menor número de coeficientes, por lo que la modelización con ARMA es siempre la mejor que puede obtenerse. No obstante, y debido a la mayor complejidad de obtención de los coeficientes, suele trabajarse con modelos que contienen exclusivamente ceros en $H(Z)$ (modelos MA) o sólo polos (modelos AR). La simplificación a la hora de obtener los coeficientes con estos modelos tiene como contrapartida el requerir un número mayor de coeficientes para modelizar la señal respecto de los necesarios con modelos ARMA.

De las dos variantes comentadas, los modelos AR suelen ser los más utilizados, ya que el control de los polos permiten modelar transiciones más rápidas que con ceros para órdenes menores. El modelo autorregresivo (AR) tiene una función de transferencia dada por:

$$H(Z) = \frac{1}{1 - \sum_{k=1}^p a_k Z^{-k}}; p > (m + q)$$

La transformación al dominio de la frecuencia se realiza evaluando la ecuación anterior para $z = e^{j2\pi fT}$:

$$H(f) = \frac{1}{1 - \sum_{k=1}^p a_k e^{-j2\pi f k T}}$$

donde ΔT es el intervalo de muestreo. La entrada y el espectro de salida del proceso están relacionados por:

$$S(f) = \frac{\sigma^2 T}{\left| 1 - \sum_{k=1}^p a_k e^{-j2\pi kT} \right|^2}$$

produciendo un espectro de potencias de alta resolución. Puede calcularse la raíz cuadrada de cada componente para obtener el espectro de amplitudes.

Los pasos a seguir para modelar una serie temporal dada con este tipo de modelos son:

1. Preprocesar los datos: eliminar tendencias, tramos corrompidos, segmentar datos.
2. Seleccionar la estructura del modelo que mejor represente a la serie (ARMA, AR, MA ,etc.). De los modelos presentados, el AR y el ARMA son los más utilizados. Puede demostrarse que un modelo AR puede sustituir a un ARMA a costa de necesitar más parámetros para modelar la señal. El modelo AR se utiliza frecuentemente debido a su simplicidad y a la existencia de algoritmos efectivos para la estimación de parámetros.
3. Obtener el modelo a partir de los datos y de algún criterio de bondad. Esto implica:

- ORDEN DEL MODELO.

La estimación del orden se realiza por medio de alguna técnica de optimización. Se pretende tener un valor de p que represente correctamente el proceso minimizando a la vez la complejidad del modelo. Se han propuesto diversos estimadores:

- El FPE (*final prediction error* - Akaike) obtiene el orden minimizando la función:

$$FPE(p) = \hat{\sigma}_{wp}^2 \frac{N + p + 1}{N - p - 1}$$

donde N es la longitud de la muestra de datos y σ es la varianza estimada según:

$$\hat{\sigma}_{wp}^2 = \frac{1}{N - p - 1} \sum_{k=p+1}^N e_r(k)^2$$

siendo $e_r(k)$ el error (o residuo) entre el valor de la muestra k-ésima de la serie original y la estimada por el modelo.

- El AIC (*Akaike information error* - Akaike) proporciona el orden que minimiza la función:

$$AIC(p) = \ln \hat{\sigma}_{wp}^2 + 2p/N$$

- El MDL (*minimized description length* - Rissanen) selecciona el orden que minimiza:

$$MDL(p) = N \ln \hat{\sigma}_{wp}^2 + p \ln N$$

- El CAT (*criterion autoregressive transfer* - Parzen) selecciona el orden que minimiza:

$$CAT(p) = \left(\frac{1}{N} \sum_{k=1}^p \frac{1}{\hat{\sigma}_{wk}^2} \right) - \frac{1}{\hat{\sigma}_{wp}^2}$$

$$\hat{\sigma}_{wk}^{-2} = \frac{N}{N-k} \hat{\sigma}_{wk}^2$$

debiendo eliminarse el valor medio de los datos antes de aplicar el criterio.

Ninguno de los estimadores del orden del modelo es óptimo. Por ejemplo, el FPE tiende a subestimar el orden, y el AIC es estadísticamente inconsistente cuando $N \rightarrow \infty$. En general, el orden del modelo es $\leq 2N^{1/2}$. Otros estudios muestran que cuando N es pequeño, el orden debe seleccionarse en el rango N/3 a N/2 para obtener buenos resultados. Por tanto, el método a seguir sería calcular varios criterios e interpretar los resultados obtenidos para seleccionar el orden final.

- COEFICIENTES DEL MODELO.

La estimación de los coeficientes (parámetros) del modelo se realiza a partir de las muestras de la señal. Como regla general, si tenemos N muestras de la señal, suele utilizarse la mitad para estimar los coeficientes y el resto para validar el modelo. Supongamos que disponemos de una secuencia de muestras S_j , $j = 0, 1, \dots, (k-1)$. Estimaremos la muestra S_k mediante el estimador:

$$\hat{S}_k = - \sum_{i=1}^{\hat{p}} \hat{a}_i S_{k-i}$$

donde p tiene un valor inicial estimado previamente. En el instante k, podemos calcular el error e(k). El método de mínimos cuadrados determina los parámetros

estimados minimizando el valor esperado del error cuadrático para cada coeficiente. Suponiendo que la secuencia es estacionaria, obtenemos p ecuaciones lineales:

$$\mathbf{r}_{i-j} = E\{S_{k-j}S_{k-i}\} = \mathbf{r}_{j-i}$$

$$\sum_{j=1}^{\hat{p}} \hat{\mathbf{a}}_j \mathbf{r}_{i-j} = -\mathbf{r}_i; i = 1, \dots, \hat{p}$$

Estas ecuaciones son las de Yule-Walker o normales. Pueden resolverse para los $\hat{\mathbf{a}}_j$ si se dan los coeficientes de correlación r_j , $j=1, \dots, p$. Dichos coeficientes se estiman a partir de la secuencia finita $\{S_k\}$. Para $k=0, 1, \dots, (N-1)$ podemos obtenerlos mediante:

$$\hat{\mathbf{r}}_i = \frac{1}{N} \sum_{k=1}^{N-1-i} S_k S_{k+i}$$

y utilizaremos los coeficientes de correlación estimados en las expresiones anteriores. Las ecuaciones de Yule-Walker pueden escribirse de forma matricial como $\mathbf{R}\mathbf{a} = \mathbf{r}$, donde \mathbf{R} es la matriz de correlación, \mathbf{r} es el vector de coeficientes de correlación y \mathbf{a} el vector de coeficientes del modelo. La solución directa de la ecuación se obtiene mediante inversión de la matriz de correlación: $\mathbf{a} = \mathbf{R}^{-1} \mathbf{r}$.

Se han desarrollado diversos algoritmos para obtener los coeficientes del modelo. Los más utilizados son: a) *Forward-backward*; b) *Least-squares*; c) Yule-Walker; d) Burg; e) *Geometric lattice*. Los dos últimos utilizan la estructura AR lattice (representación como filtro de celosía de la función de transferencia del modelo AR) para obtener los coeficientes.

4. Examinar las propiedades del modelo obtenido.

Una vez disponemos del modelo, el siguiente paso es comprobar si realmente describe el proceso que representa. Como hemos visto, la elección del orden p no es exacta. Una subestimación de p produce una baja resolución en el espectro, mientras que una sobreestimación hace aparecer componentes espúreas (el método de Burg es el más sensible a este efecto, mientras que el LS no presenta estas componentes). No existe un método definitivo de validación del modelo, por lo que se aplican diversas técnicas empíricas de ayuda a la decisión final:

- **PREDICCIÓN DEL MODELO.** Se puede comparar la serie original con la predicha por el modelo, comprobando el grado de semejanza.
- **ANÁLISIS DE RESIDUOS.** Los residuos entre la señal original y la predicha son idealmente blancos, ya que están relacionados con el ruido aleatorio de

entrada al modelo. Por tanto, un análisis de su autocorrelación y de la correlación cruzada entre ambos permiten comprobar si la hipótesis se cumple para el modelo.

- **COMPARACIÓN ENTRE MODELOS DE DIFERENTES ÓRDENES.** Se pueden comparar funciones de transferencia o diagramas de ceros y polos (casos ARMA) entre diferentes modelos con pequeñas variaciones de p respecto del estimado para comprobar la dependencia con p .
- **COMPARACIÓN ENTRE ESPECTROS.** La concordancia de espectros obtenidos mediante el modelo AR (método paramétrico) y un método basado en la DFT (no paramétrico) puede dar una indicación de lo apropiado del modelo.

Si el modelo no es suficientemente preciso, repetir los pasos 3 y 4.

En la siguiente figura se muestra una señal de respiración muestreada a 12.5 Hz y su periodograma correspondiente. El comportamiento no estacionario de la señal es evidente, por lo que resulta aconsejable segmentar la señal antes de calcular el espectro. En la parte inferior se muestra la señal subdividida en 4 secciones de igual duración. Los espectros inferiores, obtenidos mediante un modelo AR de orden 16 muestran las variaciones espectrales con el tiempo, y por tanto la no estacionariedad de la señal. La necesidad de segmentar implica que las ventanas sobre las que se calcula el estimado espectral contienen pocos datos. Puesto que la resolución de los métodos no paramétricos está basada en la longitud de la muestra de señal, la segmentación automáticamente implica reducir la resolución espectral. Con los métodos paramétricos, secuencias de menos puntos siguen produciendo alta resolución.

El principal problema que presentan estos métodos es su dependencia de la correcta estimación del orden del modelo, produciendo resultados muy diferentes para distintos valores del orden. Esta estimación es frecuentemente subjetiva, y pequeñas desviaciones producen resultados muy diferentes. Además, la ventana de datos analizada debe ser estacionaria. No obstante, puede realizarse una segmentación mayor en este caso que en los métodos paramétricos, ya que en este caso la resolución no depende de la duración de la señal.

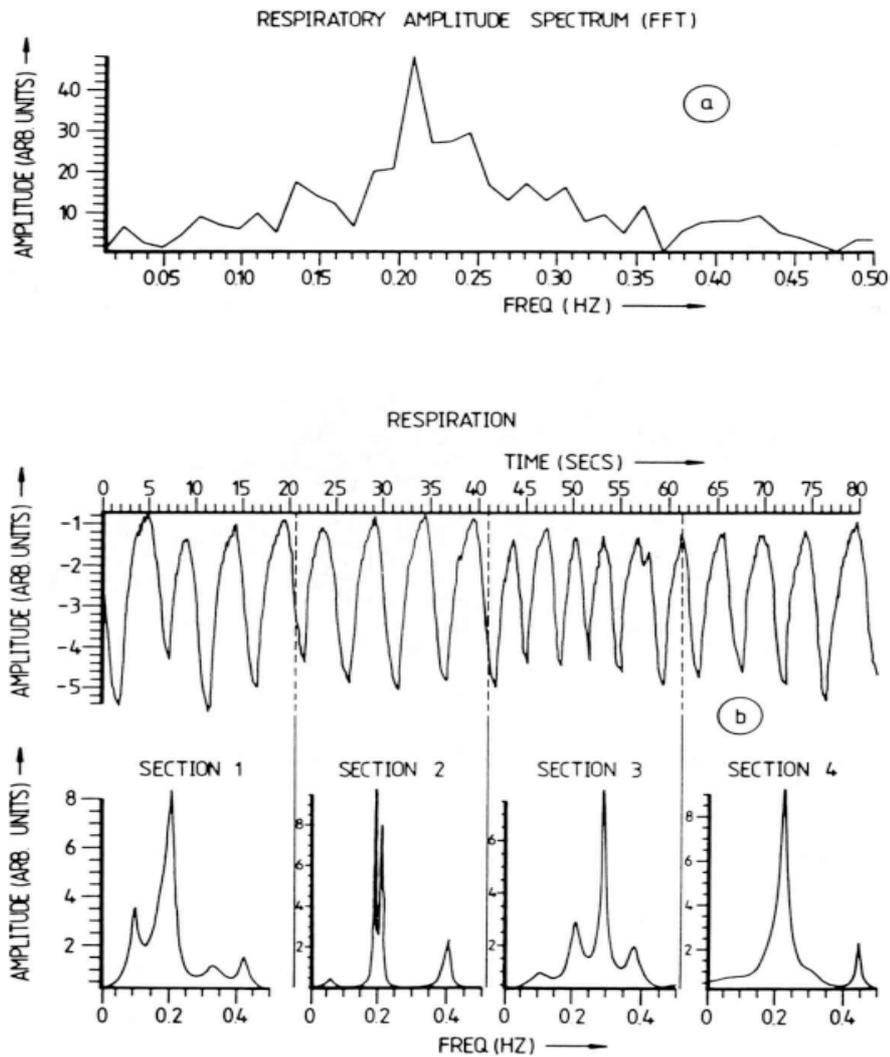


Figura 5.2. Estimación espectral mediante periodograma y modelo AR de una señal respiratoria. (M. Di Rienzo Ed. "Computer Analysis of Cardiovascular Signals". IOS Press.)

5.4. Métodos tiempo-frecuencia.

Tanto los métodos paramétricos como los no paramétricos hacen la suposición de que la señal es estacionaria. En caso contrario, se debe implementar un esquema de segmentación que asegure una ventana de datos libre de transitorios. No obstante, en

señales biomédicas es frecuente la aparición de tales eventos, e incluso pueden contener información relevante desde el punto de vista médico, como ocurre en algunas patologías.

Vamos a comentar algunas de las técnicas que han sido propuestas para el tratamiento de señales transitorias, centrándonos en las representaciones tiempo-frecuencia (TFR). Las TFR se utilizan generalmente como representaciones de distribuciones de energía de una señal dada. Por tanto, su aplicación a señales biológicas proporcionaría un método de detectar transitorios detectando cambios en la distribución de energía de la señal producidos por los mismos en un rango de frecuencias determinado. La figura 5.3 permite ilustrar esto. La imagen superior representa una señal compuesta por dos ondas seno de 50 y 100 Hz, localizadas en 0.3s y 1.3s respectivamente y de duración 0.5s. A la izquierda se muestra el espectro obtenido mediante la TF. Se obtiene información exclusivamente sobre el contenido frecuencial de la señal, pero sin determinar en qué momento se produjo. A la derecha se muestra el plano tiempo-frecuencia y la TFR de la señal. En este caso, se puede conocer simultáneamente tanto su contenido frecuencial como el instante de ocurrencia.

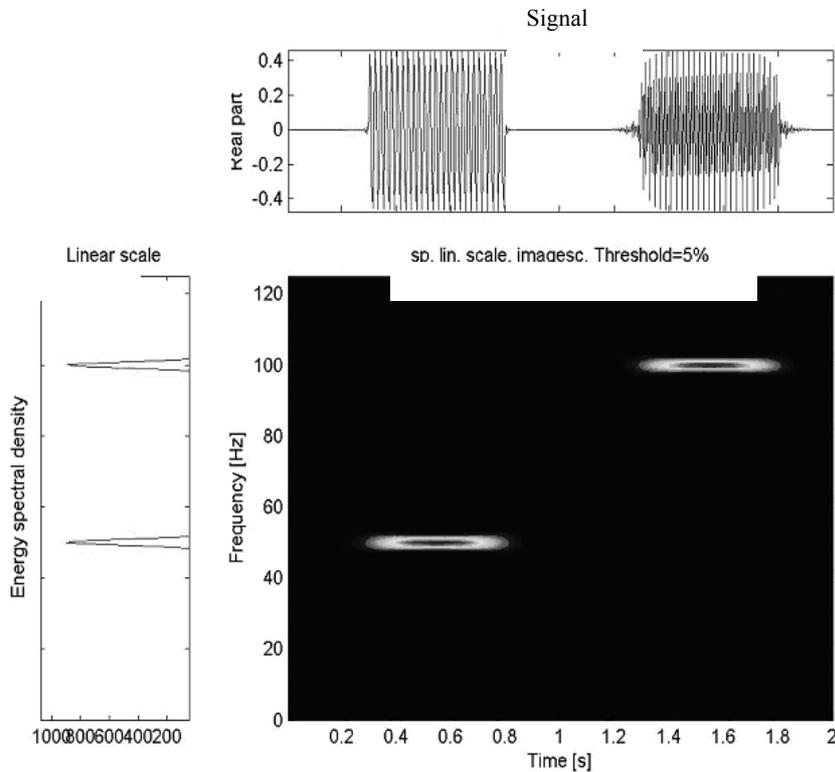


Figura 5.3. Comparación entre el espectro obtenido mediante TF y TFR para una señal localizada en el tiempo. Señal (arriba), periodograma (izquierda) y representación tiempo-frecuencia (derecha).

Podemos dividir las TFR en lineales o cuadráticas. Una TFR se dice que es lineal cuando satisface el principio de superposición lineal, es decir, dada una señal formada por dos componentes, $x(t)=a_1x_1+a_2x_2$, y siendo T_{x1} y T_{x2} las TFR de x_1 y x_2 respectivamente, la TFR total viene dada por:

$$T_x(t, f) = a_1 T_{x1} + a_2 T_{x2}(t, f)$$

Para TFR cuadráticas, se cumple el principio de superposición cuadrático:

$$T_x(t, f) = |a_1|^2 T_{x1}(t, f) + |a_2|^2 T_{x2}(t, f) + a_1 a_2^* T_{x1, x2} + a_2 a_1^* T_{x2, x1}(t, f)$$

donde T_{x1} y T_{x2} son los autotérminos de la transformada de cada señal y los otros son los términos cruzados (o términos de interferencia) de las dos señales. Puede demostrarse que para una señal compuesta por n componentes, existen $n(n-1)/2$ términos cruzados que complican el análisis visual de señales multicomponentes. Puesto que muchas TFR están relacionadas con distribuciones de energía, están dentro de este último caso.

Otro aspecto que limita las prestaciones de las TFR está relacionado con la máxima precisión alcanzable. Dada una señal $x(t)$ con energía finita, puede caracterizarse simultáneamente en el tiempo y en la frecuencia mediante su posición media y su dispersión. En el caso del tiempo, tendríamos el tiempo medio (t_m) y su duración; en el caso de la frecuencia, su frecuencia media (ν_m) y su ancho de banda. Por tanto, la señal estaría caracterizada en el plano t - f por su posición media (t_m, ν_m) y una localización de su energía con área proporcional al producto duración \times ancho de banda ($T \times B$). Este producto cumple la propiedad de estar acotado inferiormente ($T \times B \geq 1$), conocida como desigualdad de Heisember-Gabor, e indica que una señal no puede ser localizada con precisión arbitrariamente grande en el tiempo y la frecuencia simultáneamente. El valor mínimo se obtiene para señales gaussianas.

Clases

Existen múltiples TFR, que pueden agruparse en clases en función de las propiedades que cumplen. Una de las clases más importante es la de Cohen, formada por las TFR que son covariantes en tiempo y frecuencia. Esto implica que cuando la señal sea desplazada en el tiempo o en la frecuencia, su TFR se desplazará también en el plano tiempo-frecuencia. Su expresión general es:

$$C_x(t, w; \phi) = \iiint e^{j(-\theta t - \tau \theta + \omega u)} \phi(\theta, \tau) \cdot x(u + \frac{\tau}{2}) \cdot x^*(u - \frac{\tau}{2}) \cdot du d\tau d\theta$$

donde $x(t)$ es la señal y $\phi(\theta, \tau)$ se denomina *kernel* de la distribución. La elección del *kernel* da lugar a distintas TFR, y permite minimizar los términos de interferencia a costa de una pérdida de resolución.

A continuación comentaremos brevemente algunas de las TFR más usuales en aplicaciones biomédicas.

Transformada corta de Fourier (STFT)

El espectro obtenido mediante la transformada de Fourier proporciona una expansión de la señal $x(t)$ en una familia de ondas infinitas y, por tanto, no localizadas en el tiempo. Por tanto, este tipo de espectro sólo nos proporciona información de las frecuencias presentes en la señal, pero no de su momento de aparición.

La transformada corta de Fourier (Short-Time Fourier Transform: STFT) es una adaptación de la TF introducida por Gabor para incluir la dependencia temporal mediante la utilización de ventanas centradas en distintos instantes temporales de la señal sobre las que se calcula la TF. Las ventanas, $w(t)$, son de duración constante para todas las frecuencias y se supone que la señal es estacionaria en las mismas. La STFT se define como:

$$\text{STFT}_x(t, f) = \int x(\tau) w^*(\tau - t) e^{-j2\pi f\tau} d\tau$$

La resolución temporal de la STFT es proporcional a la duración efectiva de la ventana, y la resolución en frecuencia es proporcional al ancho de banda efectivo de la misma, por lo que para conseguir buena resolución temporal será necesaria una ventana muy estrecha, mientras que para el caso de la frecuencia se necesitará una ventana ancha. Esto está relacionado con la desigualdad de Heisenberg-Gabor.

El espectrograma (el módulo al cuadrado de la STFT) puede interpretarse como una distribución de la energía de la señal. Pertenece a la clase de Cohen, siempre es no negativo, y es una transformación cuadrática, por lo que cumple el principio de superposición cuadrática. Aunque no proporciona una buena resolución tiempo-frecuencia, la interferencia de los términos cruzados sólo es importante en zonas del plano en que los autotérminos se superponen. Si estos últimos están suficientemente separados, la amplitud de los términos cruzados es despreciable frente a los primeros y aparecen en las posiciones en las que aparecen también los autotérminos, por lo que quedan enmascarados. Esto puede verse en la figura 5.4, donde se muestra el espectrograma de dos señales chirp, en el primer caso cercanas por lo que aparecen términos de interferencia entre ellas (líneas verticales), y en el segundo suficientemente separadas, por lo que no aparecen interferencias.

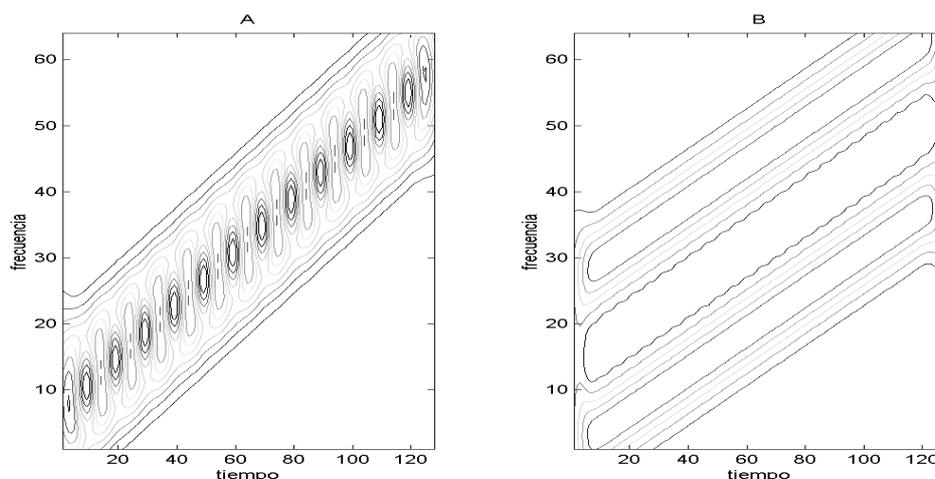


Figura 5.4. Espectrograma de dos señales *chirp*. A: frecuencias cercanas. B: frecuencias separadas.

Transformada *wavelet* (WT)

La Transformada de Fourier descompone la señal según una base de sinusoides de extensión infinita. Aunque esta aproximación es útil para señales estacionarias, crea problemas al analizar las no estacionarias. Una forma de resolver el problema es utilizar una base de funciones de extensión finita y media cero (*wavelets*) obtenidas por translación y expansión de una función prototipo (*wavelet* madre). La *wavelet* se define como:

$$\Psi_{t,a} = |a|^{-1/2} \Psi[(t-b)/a]$$

donde b indica el desplazamiento de la *wavelet* madre y a es el factor de escala, que permite obtener expansiones ($|a|>1$) o compresiones ($|a|<1$) de la *wavelet* madre.

La transformada *wavelet* continua (CWT) se define como:

$$CWT_x(a,b) = \frac{1}{\sqrt{|a|}} \int x(t) \Psi^*((t-b)/a) dt$$

y representa la suma para todo el tiempo de la señal multiplicada por diferentes versiones escaladas y desplazadas de la *wavelet* madre. En contraste con la STFT, que utiliza ventanas constantes para todas las frecuencias, la WT utiliza ventanas de corta duración a

altas frecuencias y de larga duración a bajas frecuencias, proporcionando una alta resolución en frecuencia a bajas frecuencias y una alta resolución en tiempo a altas frecuencias. Al igual que ocurre con la STFT, los términos cruzados no son despreciables cuando los autotérminos se solapan en el plano t-f.

El análisis *wavelet* ha sido utilizado como un análisis tiempo-escala, pero para *wavelets* bien localizadas alrededor de f_0 , puede expresarse en términos tiempo-frecuencia utilizando la relación $f = T_m \cdot f_0/a$ (T_m : periodo de muestreo).

La CWT permite variaciones continuas de desplazamientos y escalados de la *wavelet* madre. Esto implica una carga computacional importante, además de redundancias en los resultados. Otra aproximación es limitar el número de escalas y desplazamientos. La más usual es la basada en tomar sólo valores potencias de 2 (escalas diádicas: $a=2^j$; $b=k \cdot 2^j$; $j \in \mathbb{N}$). Esto da lugar a la transformada *wavelet* discreta (DWT), que puede ser calculada eficientemente con el algoritmo de Mallat (FWT: transformada *wavelet* rápida).

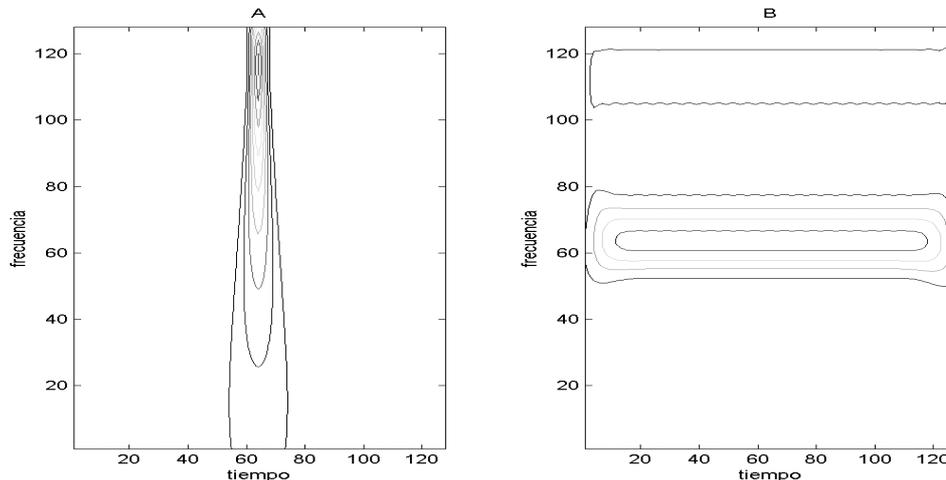


Figura 5.5. Ejemplo de dependencia de la resolución con la frecuencia para escalogramas. A: escalograma de un impulso de Dirac. B: escalograma de dos sinusoides simultáneas.

El escalograma es el módulo al cuadrado de la CWT, proporcionando también una distribución de la energía de la señal en el plano t-f. Pertenece a la clase afín, para la cual se cumple la covarianza para translación en el tiempo y expansión, y es siempre positivo. Su resolución tanto temporal como frecuencial depende de la frecuencia. La figura 5.5 permite ilustrar esta dependencia. En el primer caso, se muestra el escalograma de un impulso de Dirac en el instante $t=64$. La dispersión temporal depende de la frecuencia (a mayor frecuencia, mayor resolución temporal). En el segundo caso se muestra el

escalograma de dos sinusoides simultàneas (65 y 115 Hz). En este caso se observa menor dispersión en frecuencia (mayor resolución) para frecuencias menores.

Distribución Wigner-Ville (WVD)

La WVD pertenece a la clase de Cohen y satisface una serie de propiedades deseables en una TFR, tales como ser real y preservar desplazamientos en tiempo y frecuencia. El *kernel* de la distribución vale 1, por lo que se define como:

$$WVD_x(t, f) = \int x(t + \tau/2)x^*(t - \tau/2)e^{-j2\pi f\tau} d\tau$$

La WVD puede interpretarse como una distribución bidimensional de la energía de la señal sobre el plano tiempo-frecuencia. Se trata de una TFR cuadrática, y presenta términos cruzados no despreciables independientemente de la posición de los autotérminos (por contraposición con lo que ocurre con la STFT y la WT). Los términos cruzados aparecen en una posición dada por la media aritmética de las frecuencias de dos autotérminos y, si éstos son de igual amplitud, con amplitud doble. Esto hace que la interpretación visual de señales multicomponente o no linealmente moduladas sea difícil.

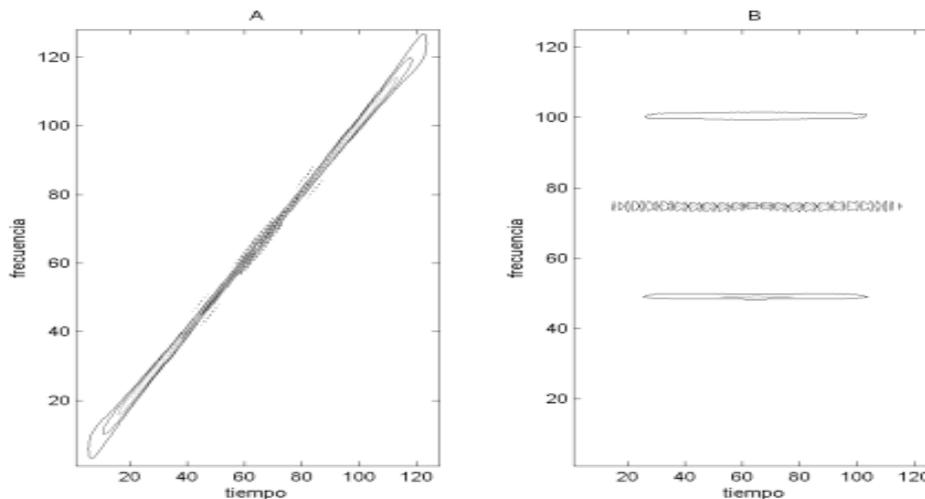


Figura 5.6. Efecto de los términos de interferencia en la WVD. A: Señal chirp. B: señal compuesta por dos sinusoides de distinta frecuencia.

La figura 5.6 muestra un ejemplo de este efecto. En primer lugar se muestra la WVD de una señal chirp de frecuencia normalizada variable entre 0 y 0,5, en la que sólo hay una única componente frecuencial en cada instante temporal, por lo que no aparecen términos de interferencia. Además, puede observarse la alta resolución proporcionada por

la WVD. En el segundo caso, se muestra una señal compuesta por dos sinusoides. Puede observarse la presencia de los autotérminos que representan a las dos componentes de 50 y 100 Hz, y los términos de interferencia entre ambos.

Distribuciones de interferencias reducidas (RID)

La reducción de los términos de interferencia puede realizarse en dos frentes. En primer lugar, puede aplicarse alguna transformación a la señal para reducir redundancias en su representación, lo que produciría también términos cruzados entre dichas redundancias. Para ello, dada una señal $x(t)$, se utiliza en su lugar su representación analítica, $x_a(t)$, definida como:

$$x_a(t) = x(t) + j \text{HT}(x(t))$$

donde $\text{HT}(x)$ es la transformada de Hilbert de x . La función analítica es una señal compleja de la señal real, y su espectro se corresponde con el de la señal original pero con valores estrictamente positivos, por lo que se eliminan los términos de interferencia entre valores negativos y positivos del espectro.

En segundo lugar se aplica promediado o filtrado local de la TFR, reduciendo los términos cruzados a expensas de introducir cierta distorsión en los autotérminos de la señal. Esta aproximación puede implementarse con una adecuada selección del *kernel* de la distribución, y da lugar a un tipo de TFR denominada distribución de interferencias reducidas (RID: *reduced interference distribution*).

La siguiente figura muestra el *kernel* de la WVD (de valor constante unidad) y la representación tiempo-frecuencia correspondiente a una señal compuesta por dos sinusoides de diferentes frecuencias y localizadas en distintos instantes temporales. Puede observarse la aparición de términos de interferencia entre los autotérminos de ambas debido a que el *kernel* de la WVD no produce ninguna modificación de la TFR en el plano tiempo-frecuencia.

La representación del *kernel* RID muestra una función con valor unidad para $\theta=0$ y $\tau=0$ y que tiende a cero para otros valores. La aplicación de este *kernel* sobre la TFR produce un filtrado de los términos cruzados presentes entre los autotérminos, reduciéndolos. La representación de la misma señal con una RID muestra una atenuación importante de los términos de interferencia, aumentando así la interpretabilidad de la misma, aunque con pérdida de resolución (los autotérminos aparecen suavizados).

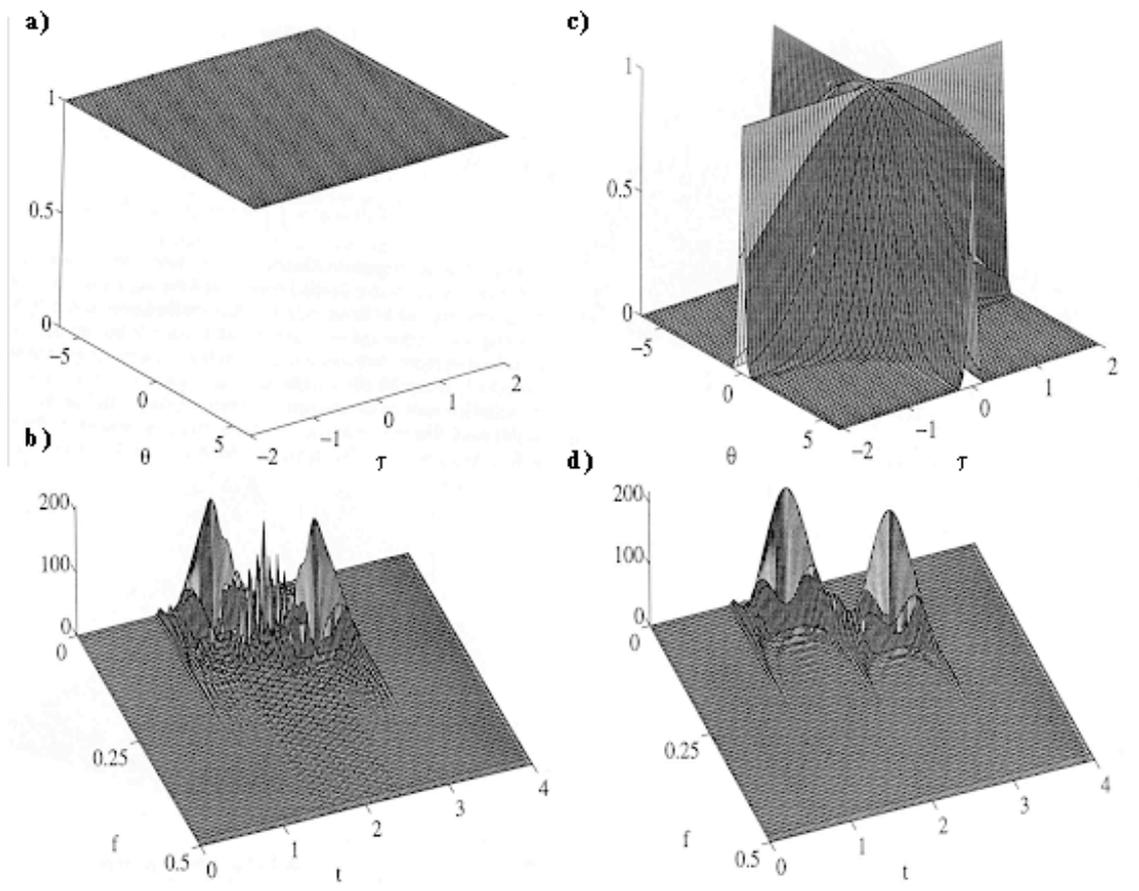


Figura 5.7. *Kernel* y representación tiempo-frecuencia de una señal compuesta por dos sinusoides de diferentes frecuencias y localizadas en distintos instantes temporales. a) y b): Wigner-Ville; c) y d): RID. (M. Akay Ed. "Time Frequency and Wavelets in Biomedical Signal Processing". IEEE Press)

Dentro de este tipo de distribuciones están la distribución de Choi-Williams, cuyo *kernel* es exponencial (por lo que también se conoce como distribución exponencial), y la de Zhao-Atlas-Mark (ZAM), con *kernel* cónico. En ambos casos, el *kernel* es fijo e independiente de la señal. Otra aproximación, que pretende optimizar el compromiso entre resolución y eliminación de términos cruzados, es la utilización de *kernels* adaptativos.

Los métodos tiempo-frecuencia no requieren que la señal sea estacionaria y proporciona información simultánea en tiempo y frecuencia, por lo que se revelan como una alternativa importante de estimación espectral a pesar de que su carga computacional es mucho mayor que la de los otros métodos.