

Estudio del comportamiento de representaciones de tiempo frecuencia en presencia de ruido heterocedástico dependiente

Sofía Nieva

Facultad de Ciencias Exactas y Naturales
Universidad Nacional de Buenos Aires

Tesis de Licenciatura
17 de Diciembre de 2021

Tabla de Contenidos

- 1 Introducción
- 2 Métodos para descomponer señales
- 3 Teoría de Wavelets
- 4 Método no paramétrico
- 5 Simulaciones
- 6 Un ejemplo real
- 7 Conclusiones

Tabla de Contenidos

① Introducción

- ▶ Preliminares
- ▶ Objetivos

② Métodos para descomponer señales

③ Teoría de Wavelets

④ Método no paramétrico

⑤ Simulaciones

⑥ Un ejemplo real

⑦ Conclusiones

Descomposición de señales

Serie de tiempo o *señal*: conjunto de observaciones x_t , cada una registradas en un momento t específico

Descomposición de señales

Serie de tiempo o *señal*: conjunto de observaciones x_t , cada una registradas en un momento t específico

Modelo aditivo: $X_t = T_t + S_t + Y_t$

- T_t : Tendencia
- S_t : Estacionalidad
- Y_t : Residuo o ruido

Descomposición de señales

Serie de tiempo o *señal*: conjunto de observaciones x_t , cada una registradas en un momento t específico

Modelo aditivo: $X_t = T_t + S_t + Y_t$

- T_t : Tendencia
- S_t : Estacionalidad
- Y_t : Residuo o ruido

Características:

- ▶ Heterocedástico
- ▶ Dependiente

Descomposición de señales

Serie de tiempo o señal: conjunto de observaciones x_t , cada una registradas en un momento t específico

Modelo aditivo: $X_t = T_t + S_t + Y_t$

- T_t : Tendencia
- S_t : Estacionalidad
- Y_t : Residuo o ruido

Características:

- ▶ Heterocedástico
- ▶ Dependiente

Ejemplos:

- ▶ ARMA
- ▶ GARCH

Métodos

Tres dominios:

- Tiempo
- Frecuencia
- Tiempo-Frecuencia \rightarrow Representaciones de tiempo-frecuencia

Métodos

Tres dominios:

- Tiempo
- Frecuencia
- Tiempo-Frecuencia → Representaciones de tiempo-frecuencia

Modelos:

- Paramétricos
 - ▶ TBATS
- No paramétricos
 - ▶ EMD
 - ▶ SST

Preguntas

1. Si hay varios componentes oscilatorios dentro de la señal, ¿cómo detectarlos y estimarlos?
2. Si existe una tendencia además de los componentes oscilatorios, ¿cómo extraerla?
3. Si el patrón de los componentes oscilatorios varía en el tiempo, ¿cómo cuantificarlo/identificarlo?
4. Dado que la longitud de los datos observados se alarga con el tiempo, ¿qué tan sensible es el estimador a la longitud de la serie de tiempo observada?
5. Si los errores son dependientes, o si la varianza del error cambia según el tiempo, ¿Es el estimador robusto ante tales errores heterocedásticos dependientes?

Identificabilidad

En estadística, un modelo se dice *identificable* si diferentes valores de los parámetros generar diferentes distribuciones de probabilidad de las variables observables. Con lo cual, es *no identificable* si dos o más parametrizaciones son observacionalmente equivalentes.

Identificabilidad

En estadística, un modelo se dice *identificable* si diferentes valores de los parámetros generar diferentes distribuciones de probabilidad de las variables observables. Con lo cual, es *no identificable* si dos o más parametrizaciones son observacionalmente equivalentes.

En otras áreas, el término *identificabilidad* puede tener otros significados. En descomposición de señales vamos a ver condiciones analíticas que permitan dar confianza de que los componentes encontrados son los correctos (estaban en el proceso y no los generó el método).

Tabla de Contenidos

- ① Introducción
- ② Métodos para descomponer señales
 - ▶ TBATS
 - ▶ EMD
- ③ Teoría de Wavelets
- ④ Método no paramétrico
- ⑤ Simulaciones
- ⑥ Un ejemplo real
- ⑦ Conclusiones

Formulación BATS

Box-Cox (B), errores ARMA (A), tendencia (T) y componentes estacionales (S)

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^\omega - 1}{\omega} & \omega \neq 0 \\ \log y_t & \omega = 0 \end{cases}$$

$$y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^K s_{t-m_i}^{(i)} + d_t$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha d_t$$

$$b_t = (1 - \phi)b + \phi b_{t-1} + \beta d_t$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$$

$$d_t = \sum_{i=1}^p \varphi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Formulación BATS

- y_t es la observación a tiempo t
- $y_t^{(\omega)}$ es la observación transformada con Box-Cox con parámetro ω
- m_1, \dots, m_K períodos constantes de las componentes estacionales
- ℓ_t es el nivel local
- b es la tendencia a largo plazo
- b_t es la tendencia a corto plazo
- $s_t^{(i)}$ es el i -ésimo componente estacional
- d_t es un proceso ARMA(p, q)
- ε_t es un proceso de ruido blanco Gaussiano (media 0, varianza σ^2)
- Parámetros de suavizado: ϕ , α , β y γ_i para $i = 1, \dots, K$.
- Argumentos: BATS($\omega, \phi, p, q, m_1, m_2, \dots, m_K$)

BATS

Ventajas y Desventajas

- Funciona para múltiples componentes estacionales
- Identifica la tendencia

BATS

Ventajas y Desventajas

- Funciona para múltiples componentes estacionales
- Identifica la tendencia
- No funciona con períodos no enteros o anidados
- No funciona para estacionalidad dinámica
- Requiere estimar muchos parámetros
- Produce componentes estacionales ruidosas

Formulación TBATS

Representación trigonométrica (T) + BATS

- Mas flexible que BATS
- Descomposición de la señal más suave
- Representación trigonométrica de los componentes estacionales, basada en series de Fourier:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t$$

Formulación TBATS

- $\gamma_1^{(i)}$ y $\gamma_2^{(i)}$ son parámetros de suavizado
- $\lambda_j^{(i)} = 2\pi j / m_i$ (notar que $\lambda_j^{(i)}$ está fijo)
- $s_{j,t}^{(i)}$ es el nivel estocástico del i -ésimo componente estacional
- $s_{j,t}^{*(i)}$ es el crecimiento estocástico del nivel
- k_i es el número de armónicos requeridos
- Argumentos: $\text{TBATS}(\omega, \phi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \dots, \{m_K, k_K\})$

Formulación TBATS

- $\gamma_1^{(i)}$ y $\gamma_2^{(i)}$ son parámetros de suavizado
- $\lambda_j^{(i)} = 2\pi j/m_i$ (notar que $\lambda_j^{(i)}$ está fijo)
- $s_{j,t}^{(i)}$ es el nivel estocástico del i -ésimo componente estacional
- $s_{j,t}^{*(i)}$ es el crecimiento estocástico del nivel
- k_i es el número de armónicos requeridos
- Argumentos: $\text{TBATS}(\omega, \phi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \dots, \{m_K, k_K\})$

El método TBATS se obtiene reemplazando el componente estacional $s_t^{(i)}$ en las ecuaciones de BATS por la formulación estacional trigonométrica

TBATS

Ventajas y Desventajas

- Requiere menos estimaciones iniciales que BATS
- Funciona para múltiples componentes estacionales con períodos no enteros o anidados
- Se puede controlar la suavidad con el número de armónicos
- Tiene en cuenta autocorrelación de los residuos

TBATS

Ventajas y Desventajas

- Requiere menos estimaciones iniciales que BATS
- Funciona para múltiples componentes estacionales con períodos no enteros o anidados
- Se puede controlar la suavidad con el número de armónicos
- Tiene en cuenta autocorrelación de los residuos
- La estimación de los parámetros depende de toda la serie
- No hay garantías para errores heterocedásticos
- Si bien tiene componentes estacionales estocásticos, los períodos son fijos

Descomposición Empírica en Modos (EMD)

Funciones de moda intrínseca (IMF)

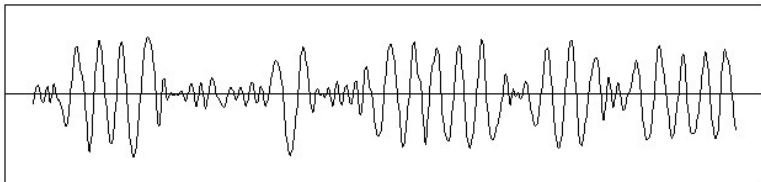
Dada una señal $f(t)$, el método la descompone en varias funciones de moda intrínseca (IMF):

$$f(t) = \sum_{k=1}^K f_k(t)$$

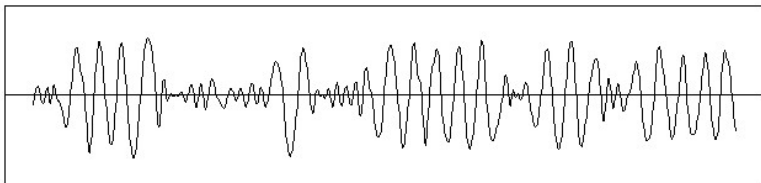
$$f_k(t) = A_k(t) \cos(\phi_k(t)), \text{ con } A_k(t), \phi'_k(t) > 0 \forall t$$

Cada IMF f_k es básicamente una función que oscila alrededor de 0, no necesariamente con frecuencia o amplitud constante, pero que en general varían lentamente

Ejemplo IMF



Ejemplo IMF



Condiciones IMF

1. En todo el conjunto de datos, el número de extremos locales y el número de ceros de $f_k(t)$ deben ser iguales o diferir máximo en uno
2. En cualquier momento t , el valor de la envolvente definida por los mínimos locales de $f_k(t)$ es el negativo de la envolvente correspondiente definida por los máximos locales.

Sifting

Construcción de la descomposición

1. Empezar con la señal original
2. Construir las envolventes inferior y superior interpolando con splines cúbicos
3. Calcular la curva media (promedio de las envolventes)
4. Restar la media a la señal y repetir hasta obtener un IMF
5. Restar este IMF a la señal original y repetir los pasos anteriores para encontrar el siguiente hasta satisfacer criterio de parada

Descomposición con EMD

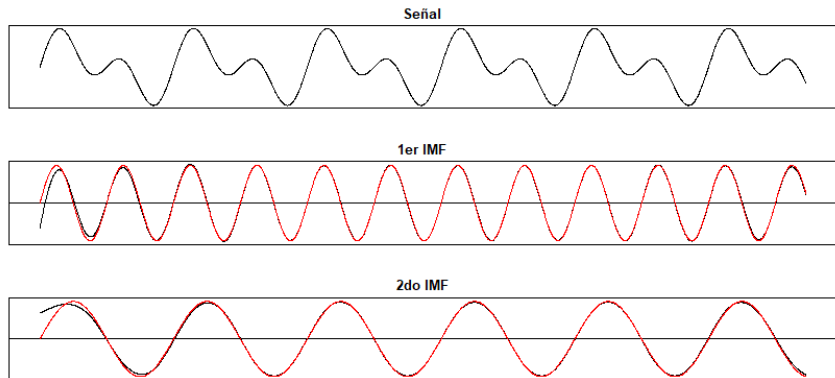


Figura: Descomposición de señal sin ruido suma de dos sinusoides. En rojo el verdadero valor de los componentes y en negro la estimación obtenida con EMD.

Inestabilidad y mezcla de modos

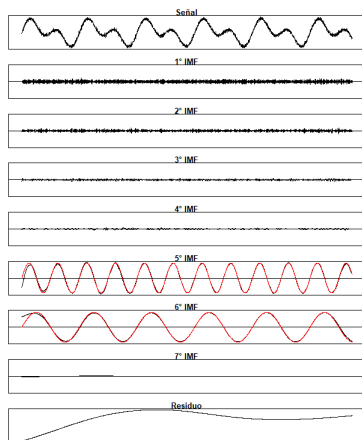


Figura: Descomposición de la misma señal con ruido blanco.

Inestabilidad y mezcla de modos

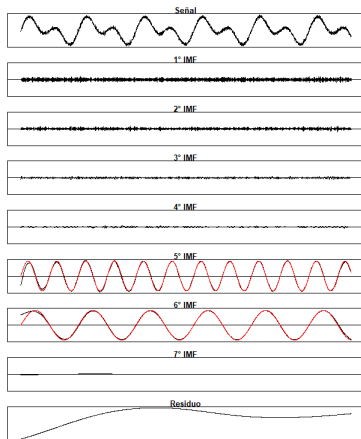


Figura: Descomposición de la misma señal con ruido blanco.

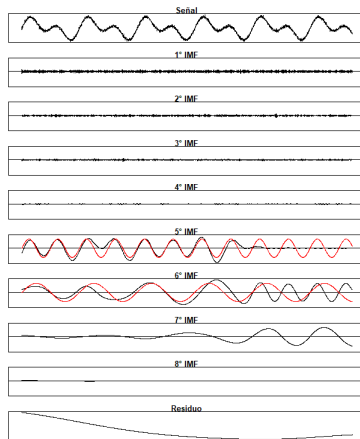


Figura: Descomposición de la misma señal con otra realización del ruido.

EMD

Ventajas y Desventajas

- Funciona para series con múltiples componentes estacionales que varían en el tiempo.
- Funciona para series no estacionarias
- Es local y adaptativo
- Uso eficaz de los datos
- Identifica la tendencia

EMD

Ventajas y Desventajas

- Funciona para series con múltiples componentes estacionales que varían en el tiempo.
- Funciona para series no estacionarias
- Es local y adaptativo
- Uso eficaz de los datos
- Identifica la tendencia

- Inestable frente a ruido
- Difícil obtener garantías teóricas

Tabla de Contenidos

- 1 Introducción
- 2 Métodos para descomponer señales
- 3 Teoría de Wavelets
 - ▶ Wavelets
 - ▶ Transformada Wavelet Continua (CWT)
 - ▶ Transformada Synchrosqueezing (SST)
- 4 Método no paramétrico
- 5 Simulaciones
- 6 Un ejemplo real

Wavelets

Pequeñas funciones con forma de onda localizada. Se utilizan para transformar una señal del dominio del tiempo a una representación de tiempo-frecuencia que presente la información de forma más útil.

Wavelets

Pequeñas funciones con forma de onda localizada. Se utilizan para transformar una señal del dominio del tiempo a una representación de tiempo-frecuencia que presente la información de forma más útil.

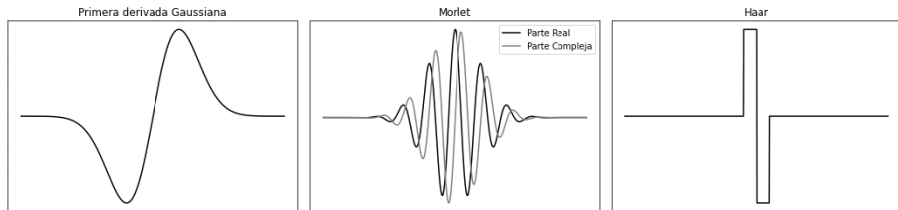


Figura: Ejemplos de wavelets

Requisitos

- Debe tener energía finita:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$$

Requisitos

- Debe tener energía finita:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$$

- $\hat{\psi}(f)$, la transformada de Fourier debe cumplir:

$$C_g = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty$$

Requisitos

- Debe tener energía finita:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$$

- $\hat{\psi}(f)$, la transformada de Fourier debe cumplir:

$$C_g = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty$$

- Para wavelets complejas, la transformada de Fourier debe ser real y decaer para las frecuencias negativas.

Ubicación y Escala

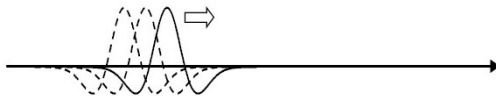


Figura: Distintas ubicaciones de una wavelet

Ubicación y Escala

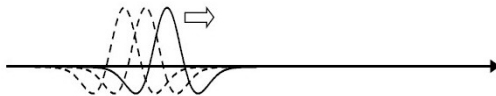


Figura: Distintas ubicaciones de una wavelet

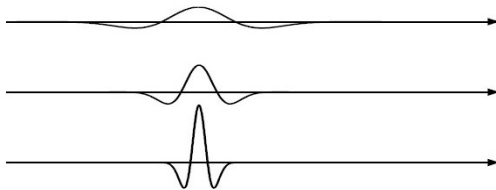


Figura: Distintas escalas de una wavelet

Wavelet madre

Es la forma básica de la wavelet de la que luego se derivan las versiones dilatadas y trasladadas que se utilizan en la transformación

Wavelet madre

Es la forma básica de la wavelet de la que luego se derivan las versiones dilatadas y trasladadas que se utilizan en la transformación

Ejemplo (Sombrero Mexicano)

Wavelet madre:

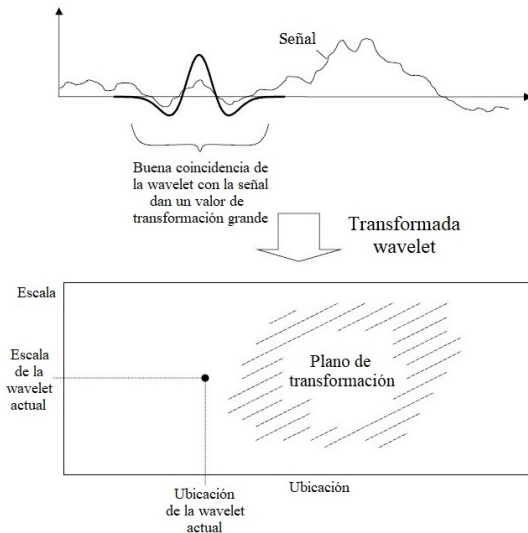
$$\psi(t) = (1 - t^2)e^{-t^2/2}$$

Versión dilatada y trasladada:

$$\psi\left(\frac{t-b}{a}\right) = \left(1 - \left(\frac{t-b}{a}\right)^2\right) e^{-\frac{1}{2}[(t-b)/a]^2}$$

a parámetro de escala, b parámetro de ubicación

Esquema CWT



Formulación

Transformada respecto a una wavelet ψ

$$T(a, b) = w(a) \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - b}{a} \right) dt$$

donde $x(t)$ es la señal, $w(a)$ es una función de peso y $*$ indica el conjugado

Formulación

Transformada respecto a una wavelet ψ

$$T(a, b) = w(a) \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - b}{a} \right) dt$$

donde $x(t)$ es la señal, $w(a)$ es una función de peso y $*$ indica el conjugado

Notación

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t - b}{a} \right)$$

$$T(a, b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^*(t) dt$$

Interrogación de una señal

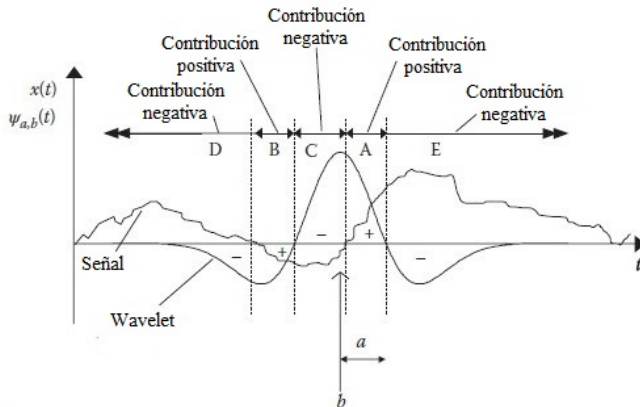


Figura: Wavelet con una escala y ubicación específica sobre una señal.

Gráfico de la Transformada Contorno

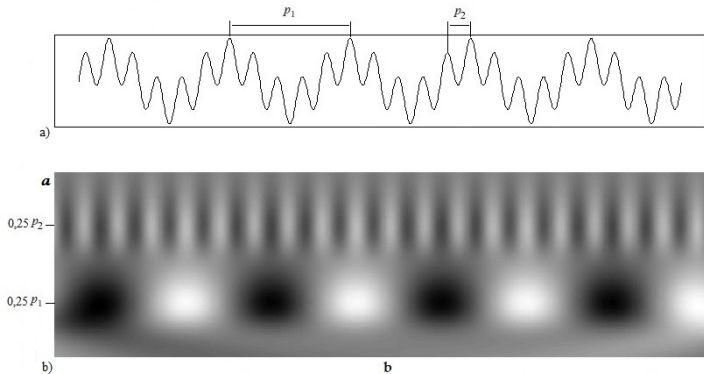


Figura: **a)** Señal compuesta por dos sinusoides **b)** Gráfico de contorno de $T(a, b)$.
(wavelet madre: Sombrero Mejicano)

Gráfico de la Transformada

Fase y módulo

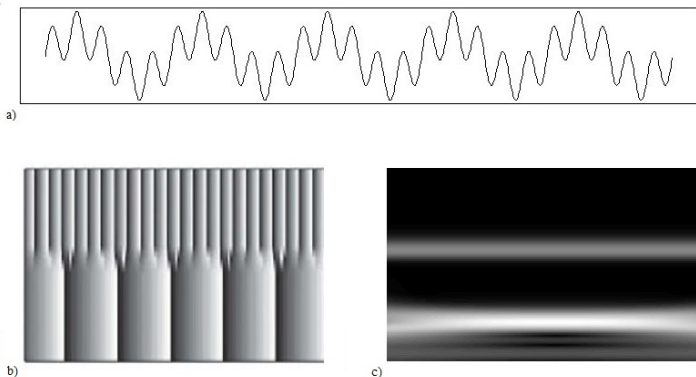


Figura: **a)** Señal compuesta por dos sinusoides. **b)** Gráfico de la fase $\phi(a, b)$ de $T(a, b)$. **c)** Gráfico del módulo de $T(a, b)$. (wavelet madre: Morlet)

Transformada inversa

Transformada wavelet inversa

Permite recuperar la señal original a partir de su transformada wavelet al integrar en todas las escalas y ubicaciones, a y b :

$$x(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_0^{\infty} T(a, b) \psi_{a,b}(t) \frac{da db}{a^2}$$

Transformada inversa

Transformada wavelet inversa

Permite recuperar la señal original a partir de su transformada wavelet al integrar en todas las escalas y ubicaciones, a y b :

$$x(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_0^{\infty} T(a, b) \psi_{a,b}(t) \frac{dad b}{a^2}$$

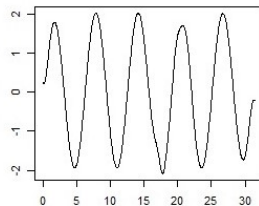
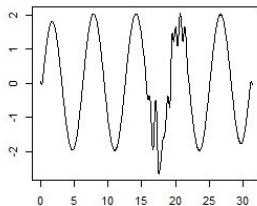
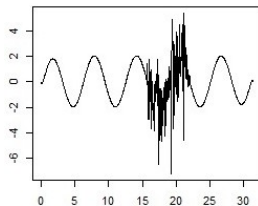
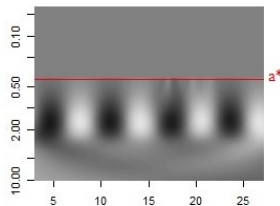
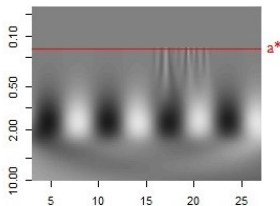
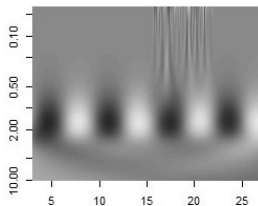
Filtrado básico

Fijamos una escala a^* como umbral y asignamos $T(a, b) = 0, \forall a < a^*$

$$x(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_{a^*}^{\infty} T(a, b) \psi_{a,b}(t) \frac{dad b}{a^2}$$

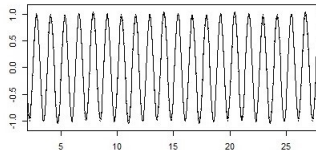
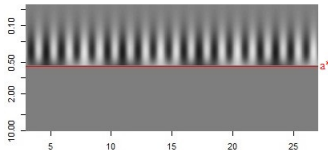
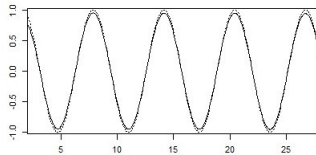
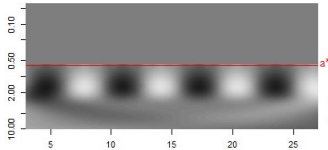
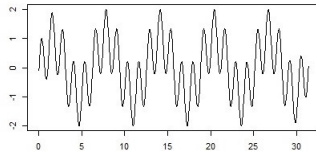
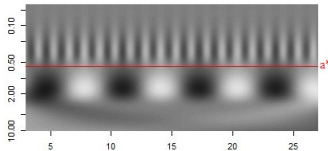
Aplicaciones

Filtrado de ruido



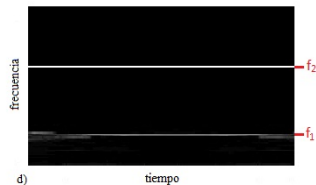
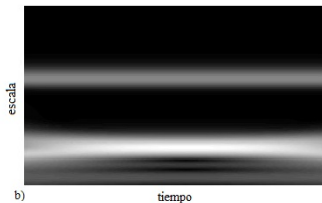
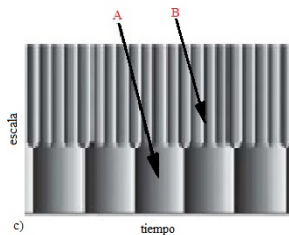
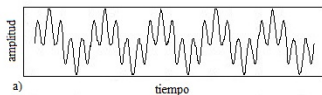
Aplicaciones

Reconstrucción de componentes



Métodos de reasignación

Synchrosqueezing



Clase \mathcal{A}_ϵ

Definición

Una función $f : \mathbb{R} \rightarrow \mathbb{C}$, se dice que es de tipo moda intrínseca con precisión $\epsilon > 0$ si f y $A := |f|$ tienen las siguientes propiedades:

$$f(t) = A(t)e^{2\pi i\phi(t)}, \text{ donde}$$

$$A \in C^1(\mathbb{R}), \phi \in C^2(\mathbb{R}), \inf_{t \in \mathbb{R}} \phi'(t) > 0$$

$$|A'(t)|, |\phi''(t)| \leq \epsilon |\phi'(t)|, \forall t \in \mathbb{R}, \text{ y } M'' := \sup_{t \in \mathbb{R}} |\phi''(t)| < \infty$$

Llamamos función de amplitud a $A(t)$, función de fase a $\phi(t)$ y frecuencia instantánea a $\phi'(t)$

Clase $\mathcal{A}_{\epsilon,d}$

Definición

Una función $f : \mathbb{R} \rightarrow \mathbb{C}$, se dice que es una superposición de componentes de moda intrínseca bien separados con precisión $\epsilon > 0$ y separación $d > 0$ si puede escribirse como:

$$f(t) = \sum_{k=1}^K f_k(t)$$

donde las f_k son funciones de tipo moda intrínseca y sus respectivas funciones de fase ϕ_k satisfacen:

$$\phi'_k(t) > \phi'_{k-1}(t) \text{ y } |\phi'_k(t) - \phi'_{k-1}(t)| \geq d [\phi'_k(t) + \phi'_{k-1}(t)], \forall t \in \mathbb{R}$$

Algoritmo SST

1. Elegir wavelet madre $\psi \in \mathcal{S}$ tal que $\text{sop}(\hat{\psi}) \subset [1 - \Delta, 1 + \Delta]$, $\Delta \ll 1$ y calcular $T_f(a, b)$
2. Calcular la regla de reasignación:

$$\omega_f(a, b) := \begin{cases} \frac{-i\partial_b T_f(a, b)}{2\pi T_f(a, b)} & \text{si } |T_f(a, b)| \neq 0 \\ \infty & \text{si } |T_f(a, b)| = 0 \end{cases}$$

3. La SST de $f(t)$ con umbral ϵ se define como:

$$S_{f, \epsilon}(b, \xi) := \lim_{\alpha \rightarrow 0} \int_{\{(a, b) : |T_f(a, b)| \geq \epsilon\}} h_{\alpha}(|\omega_f(a, b) - \xi|) T_f(a, b) a^{-3/2} da$$

donde $(b, \xi) \in \mathbb{R} \times \mathbb{R}^+$, $\alpha, \epsilon > 0$, $h_{\alpha}(b) := \frac{1}{\alpha} h\left(\frac{b}{\alpha}\right)$, $h \in L^1(\mathbb{R})$, y $h_{\alpha} \rightarrow \delta$ débilmente cuando $\alpha \rightarrow 0$ con δ la función delta de Dirac.

Resultados teóricos

En ausencia de ruido

Teorema

Sean $f \in \mathcal{A}_{\epsilon, d}$, $\tilde{\epsilon} := \epsilon^{1/3}$, ψ tal que $\text{sop}(\hat{\psi}) \subset [1 - \Delta, 1 + \Delta]$, con $\Delta < d/(1 + d)$, y sea $\mathcal{R}_\psi = \int \hat{\psi}(\zeta) \zeta^{-1} d\zeta$. Consideramos la CWT $T_f(a, b)$ de f respecto a la wavelet ψ , y la SST $S_{f, \tilde{\epsilon}}(b, \omega)$, obtenida de T_f , con umbral $\tilde{\epsilon}$. Entonces, dado ϵ suficientemente chico, vale lo siguiente:

- (I) $|T_f(a, b)| > \tilde{\epsilon}$ si para algún k , $(a, b) \in Z_k = \{(a, b); |a\phi'_k(b) - 1| < \Delta\}$
- (II) Para cada $k \in \{1, \dots, K\}$, y para cada par $(a, b) \in Z_k$ para el cual $|T_f(a, b)| > \tilde{\epsilon}$, se tiene que $|\omega_f(a, b) - \phi'_k(b)| \leq \tilde{\epsilon}$
- (III) Además, para cada $k \in \{1, \dots, K\}$, existe una constante C_1 tal que, para cualquier $b \in \mathbb{R}$,

$$\left| \left(\mathcal{R}_\psi^{-1} \int_{|\omega - \phi'_k(b)| < \tilde{\epsilon}} S_{f, \tilde{\epsilon}}(b, \omega) d\omega \right) - A_k(b) e^{2\pi i \phi_k(b)} \right| \leq C_1 \tilde{\epsilon}$$

Resultados teóricos

En presencia de ruido blanco Gaussiano

Teorema

Sea $f \in \mathcal{A}_{\epsilon,d}$, ϵ, ψ, Δ y Z_k como en el teorema anterior, con $\psi \in \mathcal{S}$ y $|\langle \psi, \psi' \rangle| < \|\psi\|_{L^2} \|\psi'\|_{L^2}$. Sea $g = f + N$, donde N es ruido blanco Gaussiano con densidad espectral ϵ^{2+p} para algún $p > 0$. Para cada k , sea $M_k \geq 1$ dado por $M_k = \max\left(\frac{1}{1-\Delta} \|\phi'_k\|_{L^\infty}, (1+\Delta) \left\|\frac{1}{\phi'_k}\right\|_{L^\infty}\right)$, entonces:

- (I) Si $a \in [M_k^{-1}, M_k]$, para cada punto $(a, b) \in Z_k$ con $|T_f(a, b)| > \tilde{\epsilon}$, $\exists E_1$ y C'_2 ctes tales que con probabilidad $1 - e^{-E_1 \epsilon^{-p}}$, $|\omega_g(a, b) - \phi'_k(b)| \leq C'_2 \tilde{\epsilon}$. Si $(a, b) \notin Z_k$ para ningún k , entonces $\exists E_2$ cte, tal que con probabilidad $1 - e^{-E_2 \epsilon^{-p}}$, $|T_g(a, b)| \leq \tilde{\epsilon} + \frac{1}{2}\epsilon$.
- (II) $\exists C'_3$ cte tal que con probabilidad $1 - e^{-E_1 \epsilon^{-p}}$, $\forall b \in \mathbb{R}$ se tiene:
- $$\left| \left(\mathcal{R}_\psi^{-1} \int_{|\omega - \phi'_k(b)| \leq C'_2 \tilde{\epsilon}} S_{f, M_k^{1/2} \epsilon + \tilde{\epsilon}}(b, \omega) d\omega \right) - A_k(b) e^{2\pi i \phi_k(b)} \right| \leq C'_3 \tilde{\epsilon}$$

Resultados teóricos

Ideas

- El método de synchrosqueezing tiene éxito en descomponer funciones arbitrarias en la clase $\mathcal{A}_{\epsilon,d}$
- Identifica la frecuencia instantánea y la amplitud de cada componente
- La SST está completamente concentrada, en el plano (b, ω) , en bandas estrechas alrededor de las curvas de frecuencia instantánea $\omega = \phi'_k(b)$
- Para reconstruir el k -ésimo componente $A_k(b)e^{2\pi i\phi_k(b)}$, invertimos $S_{f,\tilde{\epsilon}}$ restringiendo el dominio a la k -ésima banda concentrada
- Obtenemos una descomposición adaptativa con garantías teóricas sobre la precisión de la estimación de los componentes

Tabla de Contenidos

- ① Introducción
- ② Métodos para descomponer señales
- ③ Teoría de Wavelets
- ④ Método no paramétrico
 - ▶ Identificabilidad
 - ▶ Modelo
 - ▶ Resultados teóricos
- ⑤ Simulaciones
- ⑥ Un ejemplo real

Problema

Con SST podemos aproximar bien los componentes estacionales de $f \in \mathcal{A}_{\epsilon,d}$. Pero, ¿es esa la única representación de f dentro de la clase?

Problema

Con SST podemos aproximar bien los componentes estacionales de $f \in \mathcal{A}_{\epsilon,d}$. Pero, ¿es esa la única representación de f dentro de la clase?

Ejemplo

$$\frac{1}{4} \cos([\Omega - \gamma]t) + \frac{5}{2} \cos(\Omega t) + \frac{1}{4} \cos([\Omega + \gamma]t) = \left(2 + \cos^2 \left[\frac{\gamma}{2}t\right]\right) \cos(\Omega t)$$

Problema

Con SST podemos aproximar bien los componentes estacionales de $f \in \mathcal{A}_{\epsilon,d}$. Pero, ¿es esa la única representación de f dentro de la clase?

Ejemplo

$$\frac{1}{4} \cos([\Omega - \gamma]t) + \frac{5}{2} \cos(\Omega t) + \frac{1}{4} \cos([\Omega + \gamma]t) = (2 + \cos^2 [\frac{\gamma}{2}t]) \cos(\Omega t)$$

En general, la expresión $f(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t))$ no es única y por lo tanto existe el problema de la identificabilidad de f . Soluciones:

- Modelar $A(t)$ y $\phi(t)$ de forma paramétrica \rightarrow Restrictivo
- Considerar nuevas clases de funciones ($\mathcal{A}_{\epsilon}^{c_1, c_2}$ y $\mathcal{A}_{\epsilon,d}^{c_1, c_2}$) que imponen supuestos no paramétricos sobre $A(t)$ y $\phi(t)$

Clase $\mathcal{A}_\epsilon^{c_1, c_2}$

Definición

Dados $0 < \epsilon \ll 1$ y $\epsilon \ll c_1 < c_2 < \infty$, el espacio $\mathcal{A}_\epsilon^{c_1, c_2}$ de funciones de moda intrínseca consiste de funciones $f : \mathbb{R} \rightarrow \mathbb{R}$, $f \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ de la forma

$$f(t) = A(t) \cos(2\pi\phi(t))$$

donde $A : \mathbb{R} \rightarrow \mathbb{R}$ y $\phi : \mathbb{R} \rightarrow \mathbb{R}$ satisfacen las siguientes condiciones para todo $t \in \mathbb{R}$:

$$A \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \inf_{t \in \mathbb{R}} A(t) > c_1, \sup_{t \in \mathbb{R}} A(t) < c_2$$

$$\phi \in C^2(\mathbb{R}), \inf_{t \in \mathbb{R}} \phi'(t) > c_1, \sup_{t \in \mathbb{R}} \phi'(t) < c_2$$

$$|A'(t)| \leq \epsilon \phi'(t), \quad |\phi''(t)| \leq \epsilon \phi'(t)$$

Clase $\mathcal{A}_{\epsilon, d}^{c_1, c_2}$

Definición

Fijado $0 < d < 1$, el espacio $\mathcal{A}_{\epsilon, d}^{c_1, c_2}$ de superposiciones de funciones de moda intrínseca consiste en funciones f de la forma

$$f(t) = \sum_{k=1}^K f_k(t)$$

para algún $K > 0$ finito y para cada $k = 1, \dots, K$, $f_k(t) \in \mathcal{A}_{\epsilon}^{c_1, c_2}$ tal que ϕ_k satisface

$$\phi'_k(t) > \phi'_{k-1}(t) \quad \text{y} \quad \phi'_k(t) - \phi'_{k-1}(t) \geq d [\phi'_k(t) + \phi'_{k-1}(t)]$$

Clase $\mathcal{A}_\epsilon^{c_1, c_2}$

Teorema

Supongamos que $a(t) \cos \phi(t) \in \mathcal{A}_\epsilon^{c_1, c_2}$ puede representarse de una forma diferente que también pertenece a $\mathcal{A}_\epsilon^{c_1, c_2}$, es decir,

$$a(t) \cos \phi(t) = A(t) \cos \varphi(t) \in \mathcal{A}_\epsilon^{c_1, c_2}$$

Definimos $t_m \in \mathbb{R}, m \in \mathbb{Z}$ tal que $\phi(t_m) = (m + 1/2)\pi$, $s_m \in \mathbb{R}, m \in \mathbb{Z}$ tal que $\phi(s_m) = m\pi$, $\alpha(t) = A(t) - a(t)$ y $\beta(t) = \varphi(t) - \phi(t)$. Entonces $\alpha \in C^2(\mathbb{R})$, $\beta \in C^1(\mathbb{R})$, $\alpha(t_m) = 0 \forall m$, $\beta(s_m) \geq 0 \forall m$ y $\beta(s_m) = 0$ si y solo si $\alpha(s_m) = 0$. Además, tenemos que $|\alpha'(t)| \leq 3\pi\epsilon$, $|\alpha(t)| \leq \frac{4\pi^2\epsilon}{c_1}$ y $|\beta(t)| < 3\pi\epsilon$ para todo $t \in \mathbb{R}$

Clase $\mathcal{A}_{\epsilon, d}^{c_1, c_2}$

Teorema

Supongamos que $f(t) = \sum_{l=1}^N a_l(t) \cos \phi_l(t) \in \mathcal{A}_{\epsilon, d}^{c_1, c_2}$ puede representarse de una forma diferente que también pertenece a $\mathcal{A}_{\epsilon, d}^{c_1, c_2}$, es decir,

$$f(t) = \sum_{l=1}^N a_l(t) \cos \phi_l(t) = \sum_{l=1}^M A_l(t) \cos \varphi_l(t) \in \mathcal{A}_{\epsilon, d}^{c_1, c_2}$$

entonces $M = N$, $|\phi_l(t) - \varphi_l(t)| \leq E_l \epsilon$, $|\phi'_l(b) - \varphi'_l(b)| \leq E_l \epsilon$ y $|a_l(t) - A_l(t)| \leq E_l \epsilon$ para todo $l = 1, \dots, N$, donde $E_l > 0$ es una constante universal finita que depende de c_1 , c_2 y d .

Caso Continuo

Se modela a un proceso aleatorio $Y(t)$ con múltiples componentes estacionales y tendencia contaminadas por errores heterocedásticos dependientes de la siguiente manera:

$$Y(t) = f(t) + T(t) + \sigma(t)\Phi(t)$$

Caso Continuo

Se modela a un proceso aleatorio $Y(t)$ con múltiples componentes estacionales y tendencia contaminadas por errores heterocedásticos dependientes de la siguiente manera:

$$Y(t) = f(t) + T(t) + \sigma(t)\Phi(t)$$

- La estacionalidad $f(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t))$ está en $\mathcal{A}_\epsilon^{c_1, c_2}$ o en $\mathcal{A}_{\epsilon, d}^{c_1, c_2}$ cuando $K > 1$
- La tendencia $T(t)$ es una función real C^1
- $\Phi(t)$ es algún proceso aleatorio estacionario
- $\sigma \in C^\infty \cap L^\infty$ es una función suave que se utiliza para modelar la heterocedasticidad del término de error, con $\sigma(t) > 0$

Caso Discreto

En la práctica, solo podemos acceder al proceso de tiempo continuo $Y(t)$ en puntos de muestreo discretos $n\tau$, donde $n \in \mathbb{Z}$ y $\tau > 0$ es el intervalo de muestreo. Entonces, consideramos el siguiente modelo de tiempo discreto:

$$Y_n = f(n\tau) + T(n\tau) + \sigma(n\tau)\Phi_n, n \in \mathbb{Z} \quad (1)$$

Caso Discreto

En la práctica, solo podemos acceder al proceso de tiempo continuo $Y(t)$ en puntos de muestreo discretos $n\tau$, donde $n \in \mathbb{Z}$ y $\tau > 0$ es el intervalo de muestreo. Entonces, consideramos el siguiente modelo de tiempo discreto:

$$Y_n = f(n\tau) + T(n\tau) + \sigma(n\tau)\Phi_n, n \in \mathbb{Z} \quad (1)$$

- f , T y σ son como en el caso continuo
- $\Phi_n, n \in \mathbb{Z}$ es una serie estacionaria con media cero, por ejemplo, un proceso ARMA

Teoremas de Chen

Caso continuo y discreto

Muestran la robustez de SST para señales suma de estacionalidad más tendencia contaminada además por un proceso aleatorio heterocedástico y dependiente

Teoremas de Chen

Caso continuo y discreto

Muestran la robustez de SST para señales suma de estacionalidad más tendencia contaminada además por un proceso aleatorio heterocedástico y dependiente

Ventajas:

- Modelo no paramétrico junto con la técnica adaptativa de análisis de tiempo-frecuencia SST conforman un método capaz de responder satisfactoriamente las 5 preguntas
- Además tenemos los resultados de identificabilidad que nos dan confianza en las estimaciones encontradas

Tabla de Contenidos

- ① Introducción
- ② Métodos para descomponer señales
- ③ Teoría de Wavelets
- ④ Método no paramétrico
- ⑤ Simulaciones
 - ▶ Configuración
 - ▶ Parámetros
 - ▶ Experimentos
- ⑥ Un ejemplo real

Funciones

$$s_{1,1}(t) := 2,5 \cos(2\pi t), \quad s_{1,2} := 3 \cos(2\pi^2 t)$$

$$A_1(t) := (1 + 0,1 \cos(t)) \arctan\left(\frac{1132 - 200t}{87}\right) / 2 + 2$$

$$A_2(t) := 3,5\chi_{[0, 7,5]}(t) + 2\chi_{(7,5, 10]}(t)$$

$$\phi_1(t) := t + 0,1 \sin(t), \quad \phi_2(t) := 3,4t - 0,02|t|^{2,3}$$

$$s_{2,1}(t) := A_1(t) \cos(2\pi\phi_1(t)), \quad s_{2,2} := A_2(t) \cos(2\pi\phi_2(t))$$

Funciones

$$s_{1,1}(t) := 2,5 \cos(2\pi t), \quad s_{1,2} := 3 \cos(2\pi^2 t)$$

$$A_1(t) := (1 + 0,1 \cos(t)) \arctan\left(\frac{1132 - 200t}{87}\right) / 2 + 2$$

$$A_2(t) := 3,5\chi_{[0, 7,5]}(t) + 2\chi_{(7,5, 10]}(t)$$

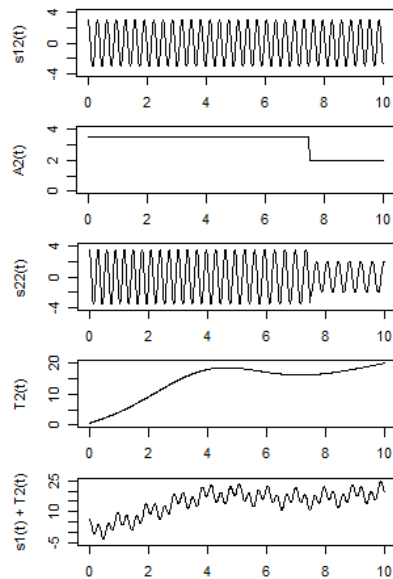
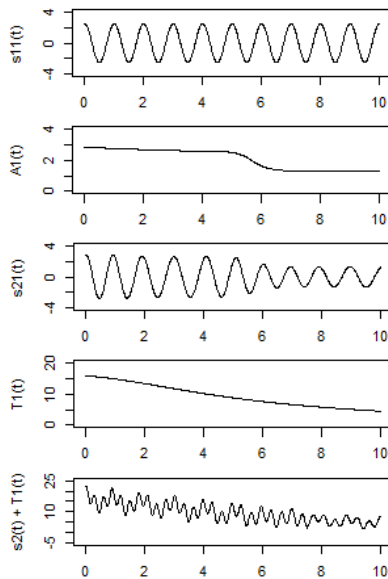
$$\phi_1(t) := t + 0,1 \sin(t), \quad \phi_2(t) := 3,4t - 0,02|t|^{2,3}$$

$$s_{2,1}(t) := A_1(t) \cos(2\pi\phi_1(t)), \quad s_{2,2} := A_2(t) \cos(2\pi\phi_2(t))$$

Estacionalidad y Tendencia

$$s_1(t) := s_{1,1}(t) + s_{1,2}(t), \quad s_2(t) := s_{2,1}(t) + s_{2,2}(t)$$

$$T_1(t) := 8 \left(\frac{1}{1 + (\frac{t}{5})^2} + e^{-\frac{t}{10}} \right), \quad T_2(t) := 2t + 10e^{-\frac{(t-4)^2}{6}}$$



Ruido

- $X_1(n) := 2\sigma(n\tau)X_{\text{ARMA1}}(n)$
- $X_2(n) := \sigma(n\tau) \left(4X_{\text{ARMA1}}(n)\chi_{[1, \frac{N}{2}]}(n) + X_{\text{ARMA2}}(n)\chi_{[\frac{N}{2}+1, N]}(n) \right)$
- $X_3(n) := 2X_{\text{GARCH}}(n)$

Ruido

- $X_1(n) := 2\sigma(n\tau)X_{\text{ARMA1}}(n)$
- $X_2(n) := \sigma(n\tau) \left(4X_{\text{ARMA1}}(n)\chi_{[1, \frac{N}{2}]}(n) + X_{\text{ARMA2}}(n)\chi_{[\frac{N}{2}+1, N]}(n) \right)$
- $X_3(n) := 2X_{\text{GARCH}}(n)$

Métrica

Raíz del error cuadrático medio

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{t=1}^T (\hat{y}_t - y_t)^2}{T}}$$

Muestreo

- Período de tiempo $[0, 10]$
- Intervalo de muestreo de longitud $\tau = 1/100$
- Resultan $N = 1001$ puntos de muestreo

Muestreo

- Período de tiempo $[0, 10]$
- Intervalo de muestreo de longitud $\tau = 1/100$
- Resultan $N = 1001$ puntos de muestreo

Experimentos

Vamos a probar los algoritmos en series de tiempo de la forma:

$$Y_{i,j,k,\sigma_0} := \mathbf{s}_i + \mathbf{T}_j + \sigma_0 \mathbf{X}_k$$

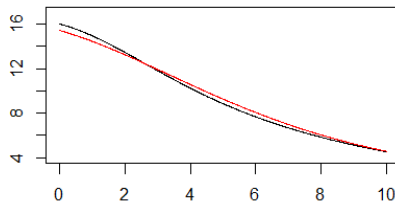
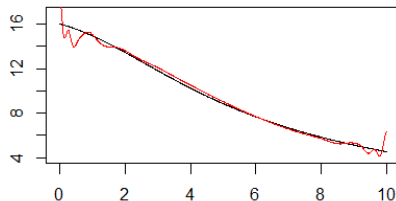
donde $i = 1, 2$, $j = 1, 2$, $k = 1, 2, 3$ y $\sigma_0 \geq 0$

Elección de parámetros

- Wavelet madre: “hhhat” (función analítica de Hilbert del sombrero hermitiano)
- Cantidad de componentes estacionales $K = 2$
- Umbral Γ , valor por defecto
- Número de voces $n_v = 32$

Elección de parámetros

- Wavelet madre: “hhhat” (función analítica de Hilbert del sombrero hermitiano)
- Cantidad de componentes estacionales $K = 2$
- Umbral Γ , valor por defecto
- Número de voces $n_v = 32$
- Lidiamos con los efectos de frontera



Experimento 1

Señal sin ruido $s_2 + T_1$

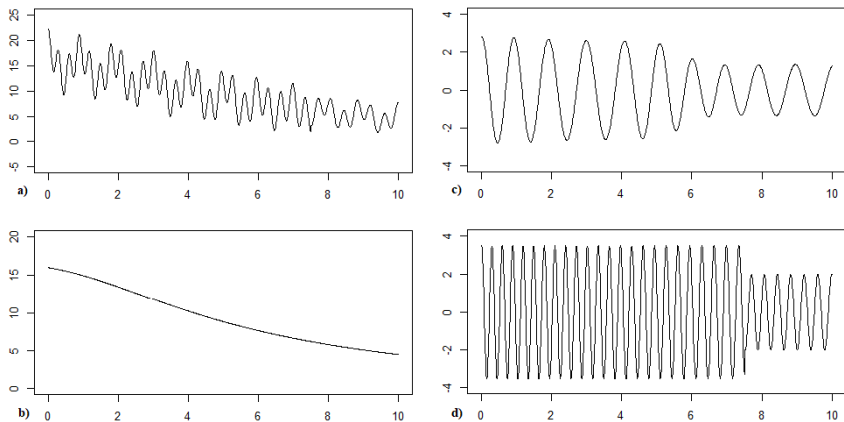
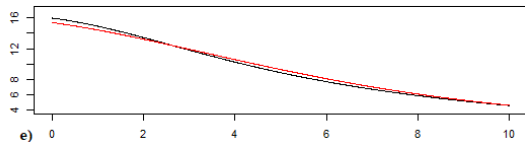
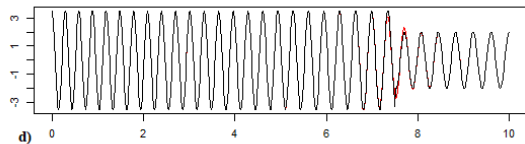
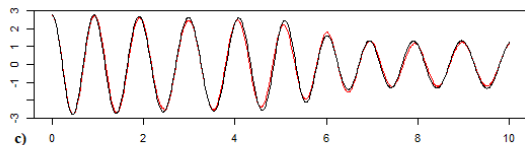
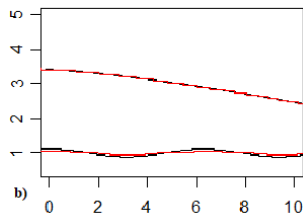
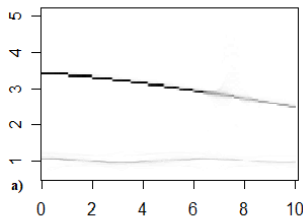


Figura: a) $s_2 + T_1$ b) T_1 c) $s_{2,1}$ d) $s_{2,2}$

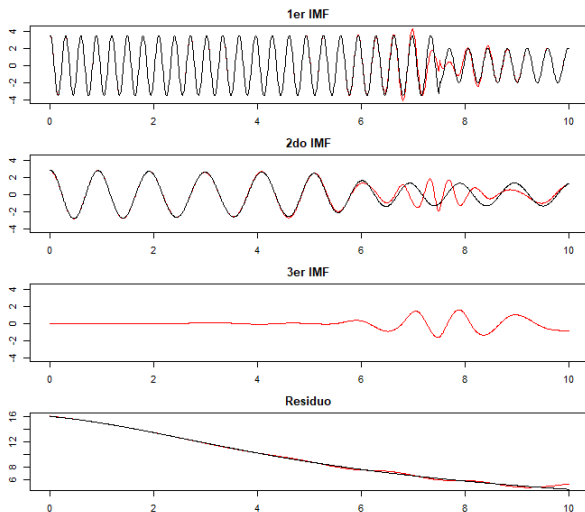
Experimento 1

Resultado SST



Experimento 1

Resultado EMD



Experimento 1

Resultados

Método	$\widetilde{s_{2,1}}$	$\widetilde{s_{2,2}}$	$\widetilde{T_1}$
SST	0,170	0,084	0,325
EMD	0,662	0,418	0,155

Tabla: Resultados de SST y EMD para la señal $s_2 + T_1$

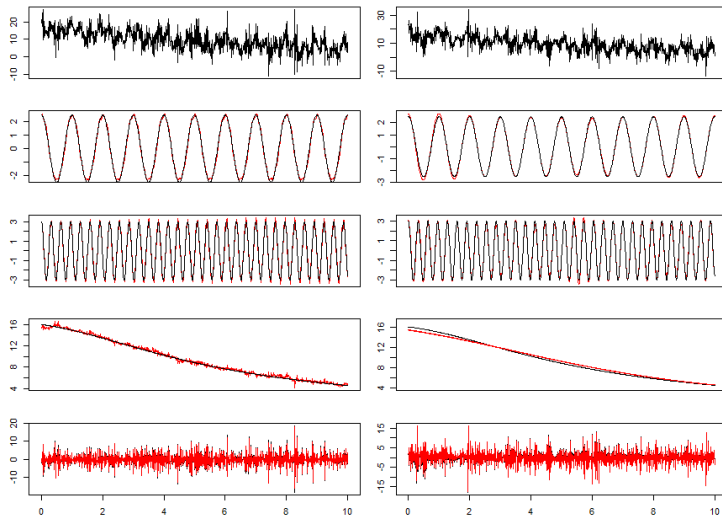
Experimento 2

Señal con estacionalidad no dinámica $Y_{1,1,1,1} := s_1 + T_1 + X_1$

- 101 realizaciones
- Analizamos con SST y TBATS
- TBATS necesita como parámetro los períodos de los componentes estacionales, 100 y $100/\pi$
- Graficamos la descomposición hecha por cada método para la realización que tuvo como RMSE la mediana entre todas las realizaciones.

Experimento 2

Resultados TBATS/SST



Experimento 2

Resultados

Método	$\widetilde{s}_{1,1}$	$\widetilde{s}_{1,2}$	\widetilde{T}_1	\widetilde{r}	Tiempo
SST	$0,103 \pm 0,023$	$0,181 \pm 0,024$	$0,327 \pm 0,020$	$0,391 \pm 0,024$	$2,761 \pm 0,223$
TBATS	$0,140 \pm 0,036$	$0,204 \pm 0,068$	$0,431 \pm 0,655$	$0,518 \pm 0,649$	$7,714 \pm 1,819$

Tabla: Resultados de SST y TBATS para la señal $s_1 + T_1 + X_1$.

Experimento 3

Señales con estacionalidad dinámica $Y_{2,j,k,\sigma_0} := s_2 + T_j + \sigma_0 X_k$

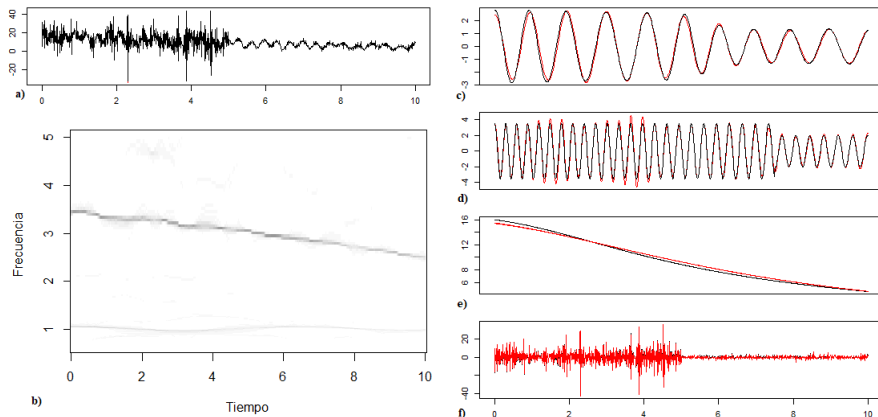


Figura: Resultado SST para la señal $Y_{2,1,2,1} := s_2 + T_1 + X_2$

Experimento 3

Resultados SST

(j, k, σ_0)	$\widetilde{s_{1,1}}$	$\widetilde{s_{1,2}}$	$\widetilde{T_1}$	\widetilde{r}	Tiempo
(1, 2, 0.5)	$0,206 \pm 0,018$	$0,164 \pm 0,020$	$0,327 \pm 0,015$	$2,960 \pm 0,199$	$2,378 \pm 0,270$
(2, 2, 0.5)	$0,205 \pm 0,018$	$0,164 \pm 0,020$	$2,124 \pm 0,016$	$3,629 \pm 0,166$	$2,945 \pm 0,494$
(1, 2, 1)	$0,241 \pm 0,035$	$0,279 \pm 0,043$	$0,333 \pm 0,029$	$0,499 \pm 0,037$	$2,765 \pm 0,262$
(2, 2, 1)	$0,241 \pm 0,035$	$0,279 \pm 0,043$	$2,124 \pm 0,031$	$2,156 \pm 0,030$	$2,661 \pm 0,169$
(1, 3, 0.5)	$0,253 \pm 0,042$	$0,279 \pm 0,032$	$0,332 \pm 0,029$	$1,967 \pm 0,078$	$2,482 \pm 0,085$
(2, 3, 0.5)	$0,253 \pm 0,042$	$0,279 \pm 0,032$	$2,125 \pm 0,029$	$2,877 \pm 0,083$	$2,758 \pm 0,146$
(1, 3, 1)	$0,368 \pm 0,076$	$0,504 \pm 0,064$	$0,348 \pm 0,058$	$0,719 \pm 0,069$	$2,824 \pm 0,128$
(2, 3, 1)	$0,368 \pm 0,076$	$0,504 \pm 0,064$	$2,128 \pm 0,058$	$2,220 \pm 0,059$	$2,411 \pm 0,120$

Tabla: Resultados de aplicar SST a señales $Y_{2,j,k,\sigma_0} := s_2 + T_j + \sigma_0 X_k$

Experimento 3

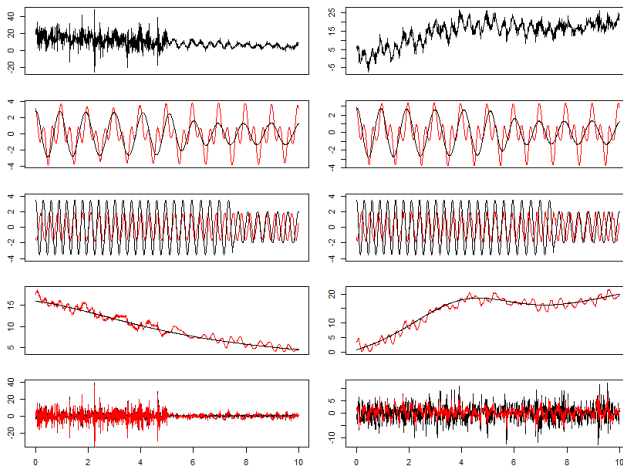


Figura: Resultado TBATS para las señales $Y_{2,1,2,1} := s_2 + T_1 + X_2$ (izquierda) y $Y_{2,2,3,0,5} := s_2 + T_2 + 0,5X_3$ (derecha)

Experimento 3

Resultados TBATS

(j, k, σ_0)	$\widetilde{s}_{1,1}$	$\widetilde{s}_{1,2}$	\widetilde{T}_1	\widetilde{r}	Tiempo	Fallas
(1, 2, 0.5)	$1,393 \pm 0,072$	$2,098 \pm 0,016$	$1,187 \pm 0,851$	$3,726 \pm 0,796$	$8,926 \pm 2,690$	1
(2, 2, 0.5)	$1,412 \pm 0,048$	$2,096 \pm 0,023$	$1,363 \pm 0,615$	$3,738 \pm 0,598$	$8,876 \pm 2,862$	1
(1, 2, 1)	$1,411 \pm 0,046$	$2,116 \pm 0,022$	$0,865 \pm 0,199$	$1,897 \pm 0,147$	$10,447 \pm 3,394$	3
(2, 2, 1)	$1,411 \pm 0,045$	$2,112 \pm 0,019$	$1,293 \pm 1,379$	$1,433 \pm 1,303$	$9,618 \pm 2,615$	3
(1, 3, 0.5)	$1,443 \pm 0,078$	$2,047 \pm 0,187$	$1,361 \pm 0,550$	$2,671 \pm 0,188$	$7,477 \pm 2,994$	4
(2, 3, 0.5)	$1,443 \pm 0,061$	$2,061 \pm 0,161$	$1,575 \pm 0,449$	$2,691 \pm 0,177$	$8,441 \pm 2,989$	2
(1, 3, 1)	$1,509 \pm 0,119$	$2,103 \pm 0,124$	$1,608 \pm 0,556$	$1,775 \pm 0,210$	$5,767 \pm 1,756$	1
(2, 3, 1)	$1,517 \pm 0,131$	$2,103 \pm 0,141$	$1,713 \pm 0,439$	$1,782 \pm 0,214$	$5,716 \pm 2,098$	0

Tabla: Resultados de aplicar TBATS a señales $Y_{2,j,k,\sigma_0} := s_2 + T_j + \sigma_0 X_k$.

Tabla de Contenidos

- ① Introducción
- ② Métodos para descomponer señales
- ③ Teoría de Wavelets
- ④ Método no paramétrico
- ⑤ Simulaciones
- ⑥ Un ejemplo real
 - ▶ Datos
 - ▶ CWT y SST
 - ▶ Descomposición

Movimiento de codornices japonesas

- Períodos de descanso y de movimiento
- Ritmos circadianos de 24 hs y ritmos ultradianos, por ejemplo de 12, 8 o 6 hs
- Patrones fractales
- Correlación temporal de largo alcance

Estudiaremos los ritmos, con períodos de varias horas, utilizando SST

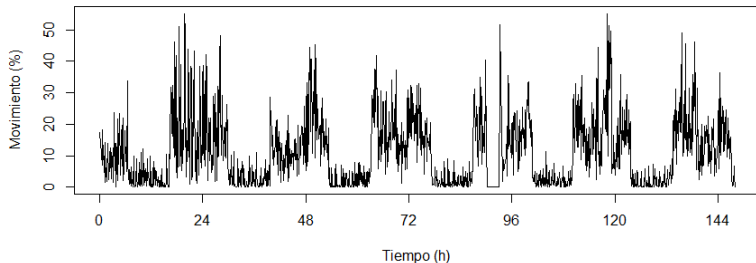


Recolección y procesamiento de los datos

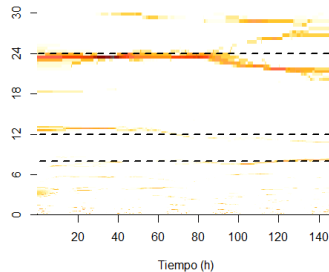
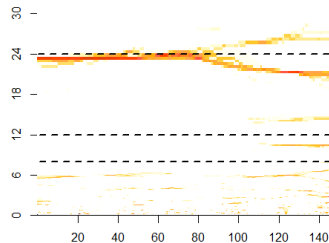
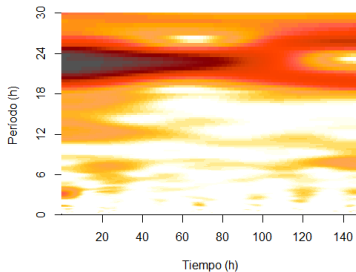
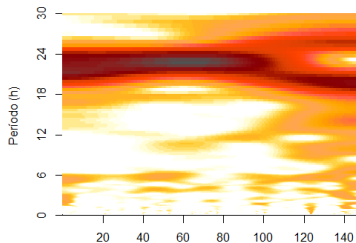
- Los datos se registraron de forma continua durante 6,5 días
- Frecuencia de muestreo de 0,5 s ($> 10^6$ datos)
- Serie de estados mutuamente excluyentes: móvil (1)/ inmóvil (0)
- Cada día, durante 30 minutos, se realizaban tareas de mantenimiento

Recolección y procesamiento de los datos

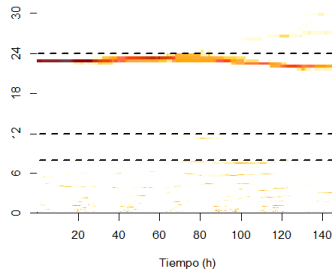
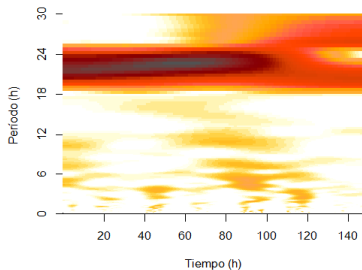
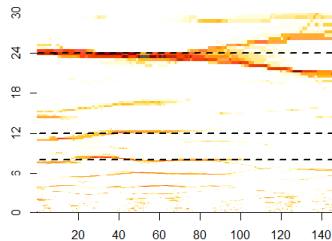
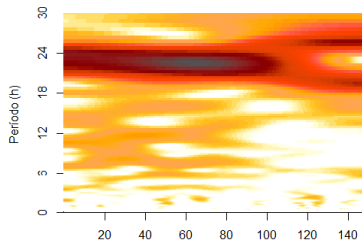
- Los datos se registraron de forma continua durante 6,5 días
- Frecuencia de muestreo de 0,5 s ($> 10^6$ datos)
- Serie de estados mutuamente excluyentes: móvil (1)/ inmóvil (0)
- Cada día, durante 30 minutos, se realizaban tareas de mantenimiento
- Se agrupan y promedian los datos de a 6 minutos
- Obtenemos una serie de largo 1479 con valores entre 0 y 1



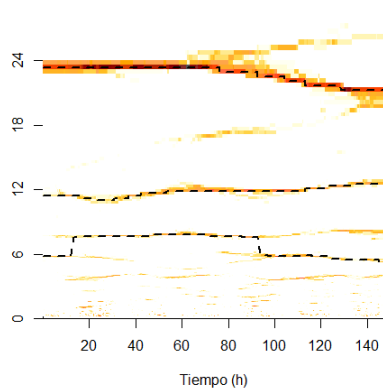
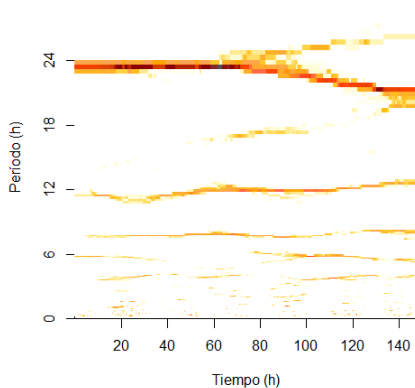
Codornices 2 y 4



Codornices 7 y 11



SST para la codorniz 5



Descomposición para la codorniz 5

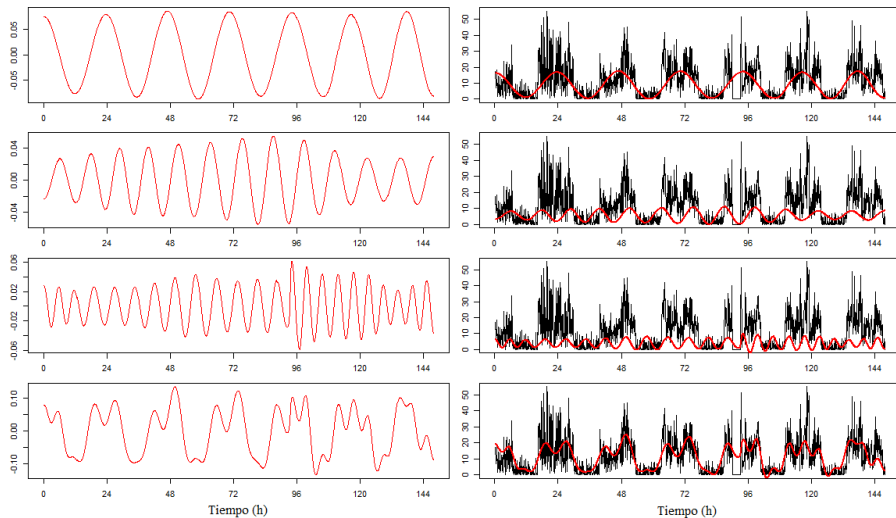


Tabla de Contenidos

- 1 Introducción
- 2 Métodos para descomponer señales
- 3 Teoría de Wavelets
- 4 Método no paramétrico
- 5 Simulaciones
- 6 Un ejemplo real
- 7 Conclusiones**

Generales

Existen muchos métodos en la literatura estadística para descomponer señales que tienen en cuenta tanto la estacionalidad como la tendencia. En este trabajo estudiamos tres de ellos: SST, EMD y TBATS.

Generales

Existen muchos métodos en la literatura estadística para descomponer señales que tienen en cuenta tanto la estacionalidad como la tendencia. En este trabajo estudiamos tres de ellos: SST, EMD y TBATS.

Sobre el problema de la identificabilidad vimos teoremas que nos permiten afirmar que si f está en cierta clase de funciones la diferencia entre distintas representaciones de f es despreciable, con lo que podemos decir que f es “identificable”, y que al aplicar SST estamos encontrando la única función que explica nuestros datos y que no hay otra observacionalmente equivalente.

Comparación de métodos

- TBATS al ser paramétrico puede ser demasiado restrictivo
- Métodos no paramétricos como EMD o SST responden mejor a un sistema dinámico, incluso si las frecuencias y las amplitudes que varían en el tiempo son desconocidas
- SST es robusto a errores, incluso heterocedásticos y dependientes, en la práctica y la teoría, mientras que EMD puede ser muy inestable y difícil de analizar matemáticamente
- SST requiere algunos argumentos, EMD ninguno

En definitiva, no hay un método universal que funcione perfecto bajo cualquier circunstancia. Depende que tipo de serie queremos estudiar y que pretendemos obtener del análisis de esa serie. Pero es importante entender las fortalezas y debilidades de cada uno para saber porque pueden fallar y cuál conviene usar en cada caso

En definitiva, no hay un método universal que funcione perfecto bajo cualquier circunstancia. Depende que tipo de serie queremos estudiar y que pretendemos obtener del análisis de esa serie. Pero es importante entender las fortalezas y debilidades de cada uno para saber porque pueden fallar y cuál conviene usar en cada caso

¡Muchas Gracias!