Análisis de series de tiempo

Técnicas básicas y nuevas tendencias

Dr. Raúl Salgado García

Centro de Investigación en Ciencias - UAEM
Instituto Potosino de Investigación Científica y Tecnológica A. C.

07 de julio de 2025

Escuela Nacional de Supercómputo:
07-11 de julio de 2025.

Día 2

Procesos estocásticos

Contenido

Autocorrelación

Función de autocorrelación

Autocorrelación

La **función de autocorrelación** es una herramienta matemática utilizada para medir la relación entre los valores de una señal o serie temporal en distintos puntos en el tiempo. Su principal objetivo es identificar si existe alguna dependencia o patrón repetido en una serie de datos en función del tiempo, y a qué intervalo de tiempo se repiten estos patrones. Se usa mucho en el análisis de series temporales y en disciplinas como la física, la econometría, la ingeniería de señales, entre otras.

- La autocorrelación mide la correlación de una serie consigo misma pero desfasada en el tiempo. Es decir, compara los valores actuales con los valores pasados de la misma serie.
- Si la serie tiene un patrón repetitivo (por ejemplo, estacionalidad o tendencias), la autocorrelación capturará estas relaciones, mostrando cómo los datos en distintos momentos están relacionados entre sí.

Autocorrelación

Función de autocorrelación

Definición: (Función de autocorrelación)

La función de autocorrelación $C : \mathbb{N} \to \mathbb{R}$ de un proceso estocástico estacionario a tiempo discreto $\{X_t : t \in \mathbb{N}_0\}$ se define como:

$$C(\tau) = \frac{\mathrm{Cov}\Big(X_0, X_\tau\Big)}{\sqrt{\mathrm{Var}(X_0)\mathrm{Var}(X_\tau)}},$$

para todo $\tau \in \mathbb{N}$. Es decir, $C(\tau) = \rho(X_0, X_\tau)$

Autocorrelación

Cuando tenemos una observación del proceso hasta un tiempo finito hacemos uso de la técnica del deslizamiento de ventana para obtener una estimación de la función de autocorrelación:

Definición:

Dada una muestra de tamaño n, $\{x_0, x_1, x_2, \ldots, x_n\}$, de un proceso estocástico $\{X_t : t \ \in \mathbb{N}_0\}$ y dado un $\ell < n$, una estimación de la función de autocorrelación se obtiene mediante la expresión:

$$\hat{C}(k) := \frac{\sum_{j=0}^{n-L}(x_j - m_{n-L}(0))(x_{k+j} - m_{n-L}(k))}{\left(\sum_{j=0}^{n-L}(x_j - m_{n-L}(0))^2\right)\left(\sum_{j=0}^{n-L}(x_{k+j} - m_{n-L}(k))^2\right)}$$

para $0 \leq k < \ell$.

Autocorrelación

El valor de la función de autocorrelación $C(\tau)$ puede variar entre -1 y 1:

- $C(\tau) = 1$: Correlación perfecta positiva. Esto indica que los valores de la serie en los puntos t y $t + \tau$ se mueven de manera idéntica.
- $C(\tau) = 0$: No hay correlación entre los valores en esos dos puntos temporales. Significa que el valor en t no proporciona información sobre el valor en $t + \tau$.
- $C(\tau) = -1$: Correlación perfecta negativa. Esto significa que cuando la serie tiene un valor alto en t, es probable que tenga un valor bajo en $t + \tau$, y viceversa.

Autocorrelación

- Detección de estacionalidad: Si la serie de tiempo tiene un comportamiento cíclico o estacional, la autocorrelación mostrará picos en los valores de τ correspondientes a la duración de los ciclos.
- Identificación de patrones repetidos: En muchos procesos, como en señales de audio o datos financieros, la autocorrelación puede ayudar a identificar patrones que se repiten.
- Modelado de series temporales: Modelos como el ARMA (AutoRegressive Moving Average) y el ARIMA (AutoRegressive Integrated Moving Average) utilizan la función de autocorrelación para ajustar modelos predictivos.
- Análisis de ruido: Si una serie de tiempo es puramente aleatoria (ruido blanco), la función de autocorrelación será cercana a cero en todos los desfases τ.

Transformada de Fourier

Transformada de Fourier

Transformada de Fourier

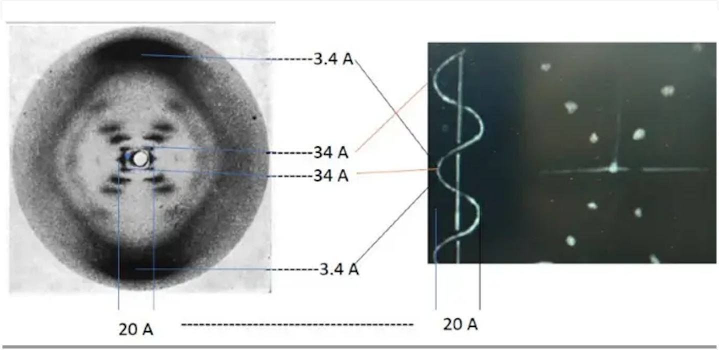
La **transformada de Fourier** es una herramienta matemática fundamental que se utiliza para analizar y descomponer señales o funciones en sus componentes de frecuencia. Es muy común en campos como la física, la ingeniería, la música, la imagenología médica y la electrónica. Cualquier función que varíe con el tiempo o el espacio (por ejemplo, una señal de audio o una imagen) puede descomponerse en una suma de ondas sinusoidales de diferentes frecuencias. Estas ondas son los "bloques básicos"de las señales. La Transformada de Fourier convierte una función del dominio del tiempo (o espacio) al dominio de la frecuencia.

Transformada de Fourier
Aplicaciones

La transformada de Fourier se utiliza en muchos campos:

- **Procesamiento de señales**: Análisis de audio, video y señales de sensores.
- **Imagenología médica**: Conversión de datos de resonancia magnética (RM) en imágenes espaciales.
- **Compresión de datos**: Usada en formatos como JPEG (imágenes) y MP3 (audio).
- **Análisis de vibraciones**: Estudio de la respuesta de sistemas a diferentes frecuencias.
- **Señal en el tiempo**: Si se tiene una señal como una onda de sonido, se puede representar como una función que cambia con el tiempo.
- **Frecuencias componentes**: La Transformada de Fourier permite descubrir qué frecuencias están presentes en esa señal y con qué intensidad.

“Foto 51”



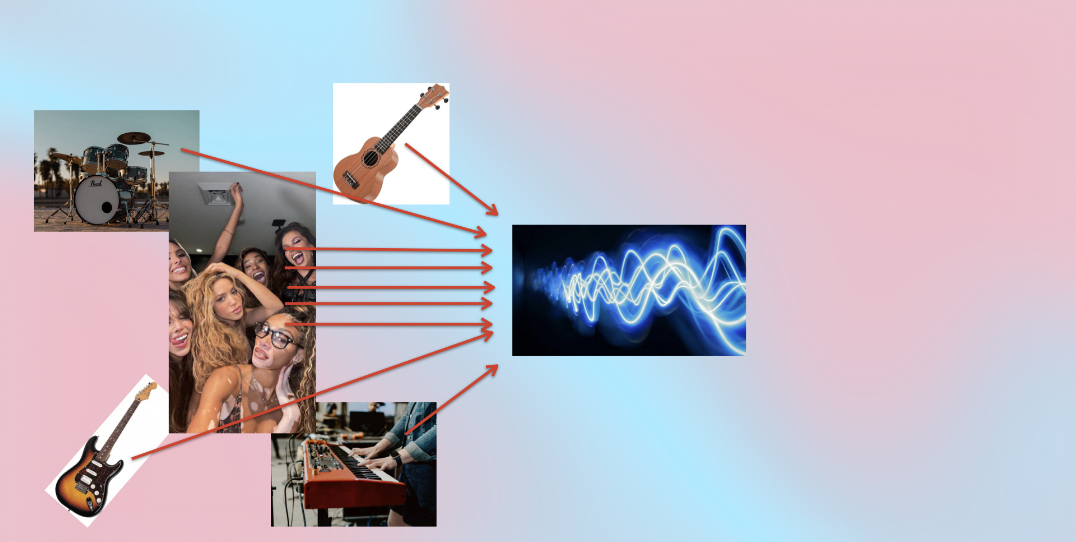
Transformada de Fourier

La transformada de Fourier

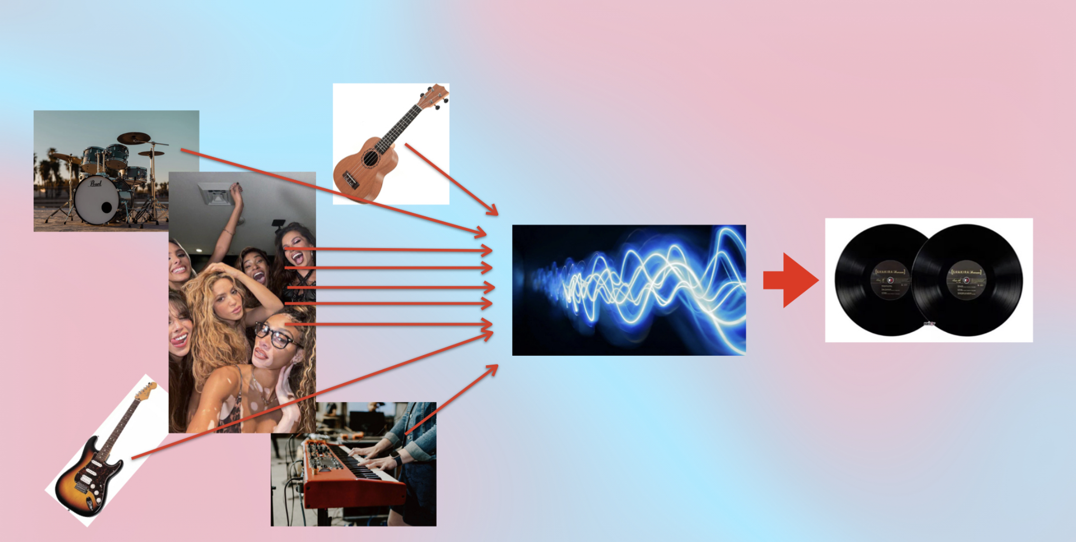
Transformada de Fourier

Transformada de Fourier

Transformada de Fourier

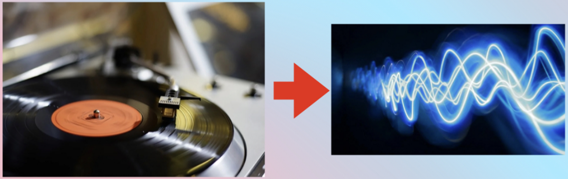


Transformada de Fourier



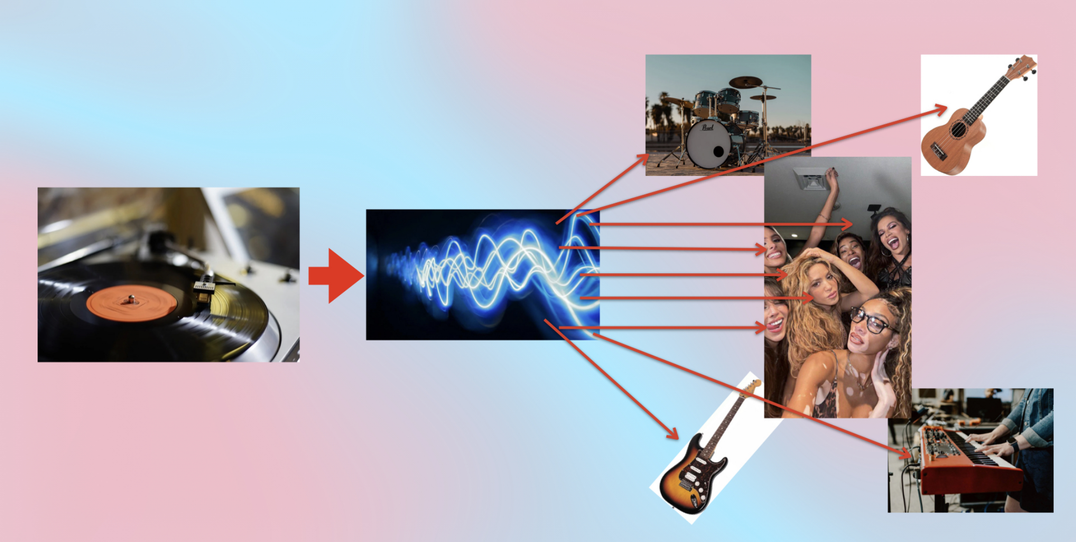
Transformada de Fourier

Transformada de Fourier



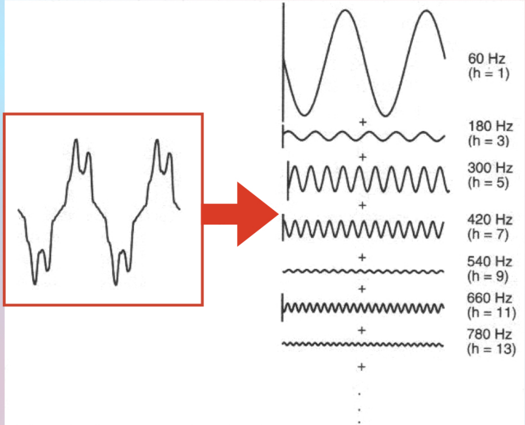
Transformada de Fourier

Transformada de Fourier



Transformada de Fourier

La transformada de Fourier descompone una señal en sus componentes armónicas básicas



Transformada de Fourier

Definición: (Transformada de Fourier)

La **transformada de Fourier** de una función $f : \mathbb{R} \to \mathbb{C}$ se define como:

$$\tilde{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx.$$

La transformación inversa está dada por

$$f(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dx.$$

Transformada de Fourier

Ejemplo:
Consideremos la función

$$f(x) = \frac{2}{\sqrt{2\pi}}(3\sin(2x) + 4\cos(3x))$$

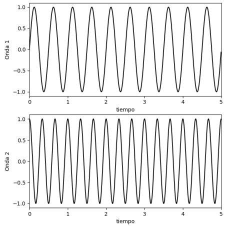
entonces:

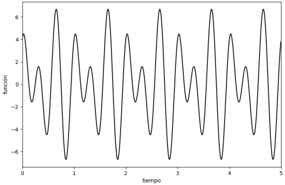
$$\begin{aligned} \tilde{f}(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{2}{\sqrt{2\pi}}\right)(3\sin(2x) + 4\cos(3x))e^{-ikx}dx \\ &= \Big(-3i[\delta(k-2) - \delta(k+2)] + 4[\delta(k-3) + \delta(k+3)]\Big). \end{aligned}$$

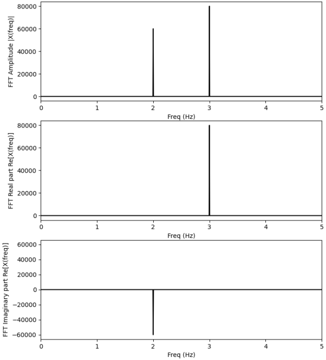
Y las amplitudes están dadas por:

$$|\tilde{f}(k)|^2 = \mathbf{9}\big[\delta(k-2) - \delta(k+2)\big]^2 + \mathbf{16}[\delta(k-3) + \delta(k+3)]^2.$$

Transformada de Fourier

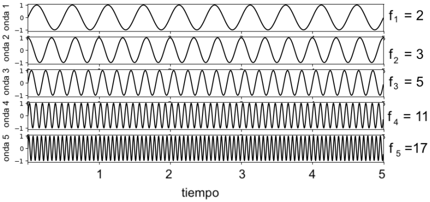


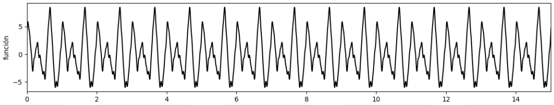


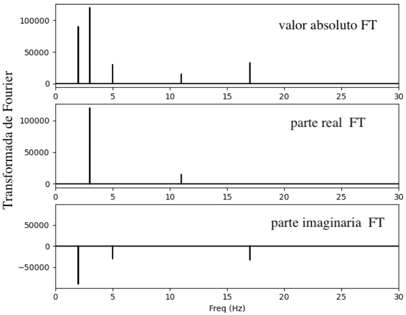


Transformada de Fourier

Otro ejemplo: $f(t) = \mathbf{3}\sin(2\pi f_1 t) + \mathbf{4}\cos(2\pi f_2 t) + \mathbf{0.5}\sin(2\pi f_3 t) + \mathbf{0.3}\cos(2\pi f_4 t) + \mathbf{0.8}\sin(2\pi f_5 t)$







Transformada de Fourier

Teorema de Wiener-Kintchin

El teorema de Wiener-Khintchine establece una relación fundamental entre la función de autocorrelación de un proceso estacionario y su densidad espectral de potencia. Este teorema es muy importante en el análisis de señales, ya que permite estudiar cómo las características temporales de un proceso aleatorio se relacionan con su comportamiento en el dominio de las frecuencias.

Transformada de Fourier

Teorema de Wiener-Kintchin

Para un proceso estocástico a tiempo continuo $\{X(t) : t \geq 0\}$ con promedio cero, $\mathbb{E}[X] = 0$, la densidad espectral $S(f)$ de la señal $X(t)$ se define como :

$$S(f) := \mathbb{E}\left[\int_{-\infty}^{\infty} e^{-2\pi i f t} X(t) dt\right]$$

y el **teorema de Wiener-Kintchin** nos dice que la densidad espectral es proporcional a la transformada de Fourier de la función de autocorrelación $R(\tau)$ no normalizada (i.e. la covarianza):

$$R(\tau) = \mathrm{Cov}(X(t), X(t+\tau))$$

es decir:

$$S(f) = 2\tilde{R}(\omega) := 2\int_{-\infty}^{\infty} e^{-2\pi i f \tau} R(\tau) d\tau$$

Exponente de Hurst

Exponente de Hurst

Exponente de Hurst

Explicación conceptual

- El **exponente de Hurst** (denotado comúnmente por H) es un parámetro utilizado en análisis de series temporales para medir el **grado de persistencia** o **memoria larga** de un proceso estocástico o una serie temporal.

- Fue propuesto por el hidrólogo inglés Harold Edwin Hurst en el contexto de la investigación sobre el comportamiento del río Nilo, y desde entonces se ha aplicado en una amplia variedad de campos, incluyendo la física, economía, biología, entre otras disciplinas.

- El exponente de Hurst toma valores entre 0 y 1 y refleja la tendencia que tiene una serie temporal de permanecer en la misma dirección (**persistencia**) o revertirse (**antipersistencia**).

Exponente de Hurst

Interpretación del Exponente de Hurst de acuerdo a su valor

- $H = 0.5$: Esto corresponde a un camino aleatorio o un proceso completamente no correlacionado (como el **movimiento browniano estándar**). En este caso, no hay memoria ni dependencia en las fluctuaciones pasadas. Cada incremento en la serie es completamente independiente de los anteriores.
- $H > 0.5$: Indica **persistencia** o **dependencia positiva**. Si una serie ha estado creciendo, es más probable que continúe creciendo, y si ha estado disminuyendo, es más probable que continúe disminuyendo. En otras palabras, hay una tendencia a que las fluctuaciones de la serie sigan la misma dirección que en el pasado.
- $H < 0.5$: Indica **antipersitencia** o **dependencia negativa**. En este caso, si una serie ha estado creciendo, es más probable que comience a disminuir y viceversa. Este comportamiento se llama también reversión a la media, donde las fluctuaciones tienden a oscilar alrededor de un valor promedio.

Exponente de Hurst

Definición del exponente de Hurst

El exponente de Hurst, H, se define en términos del comportamiento asintótico del rango reescalado en función del lapso de tiempo de una serie temporal de la siguiente manera:

$$\mathbb{E}\left[\frac{R(n)}{S(n)}\right] = Cn^H, \text{ conforme } n \to \infty.$$

donde:

- $R(n)$ es el rango del primer n desviaciones acumulativas de la media
- $S(n)$ es la serie (suma) de las primeras n desviaciones estándar
- n es el lapso de tiempo de la observación (número de puntos de datos en una serie temporal)
- C es una constante.

Exponente de Hurst

Algoritmo de estimación

Dada una serie de observaciones $\{x_1, x_2, \ldots, x_n\}$ (muestra de tamaño n) el exponente de Hurst se calcula como sigue:

- Estimamos la media muestral:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

- Definimos una nueva serie $\{y_1, y_2, \ldots, y_n\}$ con media muestral nula:

$$y_i = x_i - \overline{x}, \text{ para } 0 \leq i \leq n.$$

- Definimos la serie $\{z_1, z_2, \ldots, z_n\}$ hasta el número de pasos $t \leq n$ de suma acumulada:

$$z_k = \sum_{i=1}^{k} y_i$$

Exponente de Hurst

Algoritmo de estimación (continuación)

- El rango se calcula como

$$R(n) := \max\{z_1, 2_2, \ldots, z_n\} - \min\{z_1, z_2, \ldots, z_n\}$$

- Y la desviación estándar $S(n)$ se calcula como:

$$S(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2}.$$

El exponente de Hurst se estima ajustando una ley de potencia a los valores obtenidos de $R(n)/S(n)$ (promediados sobre diferentes realizaciones) en función de n:

$$\mathbb{E}[R(n)/S(n)] = Cn^H.$$