Instituto Politécnico Nacional

Escuela Superior de Computo

Carrera

Licenciatura en Ciencia de Datos

Alumno

Jiménez Flores Luis Arturo

Profesor

Daniel Jiménez Alcantar

Materia

Análisis de Series de Tiempo

Grupo

6AV1

Practica 8

Periodograma y Transformada discreta de fourier

Fecha de entrega: 13/06/2025

Instrucciones

1.- Elija un Dataset de su elección, deberá acondicionarse de tal manera que pueda realizar el análisis de una serie de tiempo. Desarrolle un reporte técnico que permita observar el trabajo en los siguientes puntos:

1. Introducción
2. Problemática
3. Modelo estadístico
4. Modelo computacional
5. Metodología.
6. Propuesta de solución para el uso del Periodograma y transformación discreta de fourier.
7. Conclusiones por integrante.

Introducción

El presente trabajo constituye un informe técnico detallado sobre el análisis de series de tiempo aplicado a indicadores económicos para México. Bajo el marco de la práctica titulada "Periodograma y Transformada discreta de fourier", se explora la estructura y dinámica de variables macroeconómicas con el objetivo de desentrañar los patrones subyacentes que gobiernan su comportamiento a lo largo del tiempo. La finalidad última de este tipo de análisis es llegar a una comprensión profunda de la estructura de una serie temporal para, eventualmente, poder realizar pronósticos sobre su evolución futura basándose únicamente en sus valores pasados.

Para abordar este desafío, se ha adoptado como pilar metodológico el enfoque de Box-Jenkins, reconocida para la modelización de series temporales. Este procedimiento se articula en cuatro fases secuenciales y rigurosas: la identificación del modelo, la estimación de sus parámetros, la validación de su pertinencia y, finalmente, su utilización para fines de pronóstico. A través de este marco, buscamos responder a preguntas tan relevantes como: "Considerando solo el comportamiento histórico del crecimiento de México, ¿cuál es el pronóstico para el próximo trimestre?".
Los datos que sustentan este análisis comprenden series temporales de frecuencia trimestral que abarcan desde 1993 hasta 2024. Las variables centrales incluyen el Producto Interno Bruto de México (pib_mx), así como sus exportaciones (export) e importaciones (import), todas ellas expresadas como variables reales en millones de pesos constantes de un año base. De forma complementaria, se integra el Producto Interno Bruto de los Estados Unidos (pib_usa), medido en miles de millones de dólares constantes, para contextualizar el entorno económico regional.

Es importante destacar que, dado que las series principales ya se encuentran en términos reales, el índice de precios o deflactor ha sido explícitamente ignorado para evitar redundancias en el análisis.

Si bien un análisis univariado, como el que se presenta, no pretende explicar las complejas interacciones entre las distintas variables económicas, sí constituye un paso fundamental e ineludible. Permite construir un modelo de pronóstico individual para cada serie, sentando las bases para análisis multivariados más complejos en el futuro. Por lo tanto, este informe se centrará en desglosar, paso a paso, el tratamiento de los datos, la transformación de las series para asegurar su estacionariedad, y la aplicación de herramientas de análisis espectral como el periodograma y la transformada de Fourier para identificar ciclos y tendencias que no son evidentes a simple vista. A través de este proceso, se busca ofrecer una perspectiva clara y fundamentada sobre la dinámica intrínseca de la economía mexicana.

Problemática

La problemática central para el análisis y modelización de las series de tiempo económicas presentadas no radica en la ausencia de datos, sino en las características intrínsecas que estos poseen en su estado bruto. Las series económicas como el Producto Interno Bruto (PIB), las exportaciones o las importaciones, por su naturaleza, son inherentemente no estacionarias. Esto significa que sus propiedades estadísticas fundamentales, como su media y su varianza, no son constantes a lo largo del tiempo, sino que evolucionan siguiendo tendencias y ciclos económicos. Esta condición de no estacionariedad viola un supuesto clave de muchos modelos econométricos, incluyendo el modelo ARIMA, cuya aplicación sobre series en su nivel original conduciría a resultados espurios y poco fiables.

A este desafío estructural se suma una complicación de carácter práctico: la heterogeneidad de las unidades y divisas en las que se miden las variables. Mientras que los datos de la economía mexicana (pib_mx, export, import) están expresados en millones de pesos constantes, el indicador de la economía estadounidense (pib_usa) se presenta en miles de millones de dólares constantes. Esta disparidad impide cualquier tipo de comparación directa o modelización conjunta que busque entender las interrelaciones entre ambas economías. Por lo tanto, antes de cualquier análisis, surge la necesidad ineludible de estandarizar los datos a una unidad común, que para este estudio se estableció en millones de dólares.

En consecuencia, la problemática fundamental a resolver es doble. Primeramente, las conversiones de moneda y transformaciones logarítmicas para estabilizar la varianza y linealizar las relaciones, y la aplicación de diferencias para alcanzar la estacionariedad, es un prerrequisito no negociable. Segundo, una vez obtenida una serie estacionaria, el reto se desplaza hacia la correcta identificación de su estructura de dependencia temporal. Es aquí donde el análisis debe determinar, a través de herramientas como los gráficos de autocorrelación (ACF) y autocorrelación parcial (PACF), los órdenes p y q que configurarán un modelo ARIMA adecuado para capturar la dinámica de la serie transformada. Abordar esta problemática de manera rigurosa es esencial para garantizar la validez estadística del modelo final y la pertinencia de sus conclusiones.

Modelo estadístico

Para abordar la complejidad de las series de tiempo económicas, este estudio emplea un enfoque dual que combina el análisis en el dominio de la frecuencia con la modelización en el dominio del tiempo.

El pilar del análisis de frecuencia lo constituyen la **Transformada Discreta de Fourier (TDF)** y el **Periodograma**, herramientas poderosas para descomponer la serie y descubrir sus ciclos inherentes como una especie de filtrado.

La **Transformada Discreta de Fourier (TDF)** es el principal instrumento matemático utilizado para migrar la representación de la serie del dominio del tiempo (donde se observa el valor en cada trimestre) al dominio de la frecuencia. Este cambio de perspectiva permite visualizar la serie no como una secuencia de eventos, sino como una superposición de ondas sinusoidales de distintas frecuencias y amplitudes. La gran ventaja de este método es que habilita la descomposición empírica de la serie original en sus componentes fundamentales mediante la aplicación de filtros:

Filtro Pasa-Baja (Extracción de la Tendencia): Para aislar la tendencia de largo plazo, se aplica un filtro pasa-baja. El proceso consiste en calcular la TDF de la serie (log_pib_mx_usd), establecer una frecuencia de corte (por ejemplo, 0.33 ciclos por año, para eliminar ciclos que duren menos de 3 años) y anular todas las frecuencias que superen dicho umbral. Al aplicar la transformada inversa sobre esta versión filtrada, se reconstruye una nueva serie temporal suavizada que representa exclusivamente la tendencia o el ciclo de largo plazo de la economía.

Filtro Pasa-Alta (Extracción del Ciclo de Corto Plazo): De manera análoga, para aislar el componente cíclico de alta frecuencia, se anulan todas las frecuencias que estén "por debajo" de la frecuencia de corte. La serie resultante tras la transformada inversa representa las fluctuaciones de corto plazo, es decir, las desviaciones de la economía respecto a su tendencia de largo plazo. Este componente aislado es, por definición, estacionario y puede ser analizado de forma independiente.

Por su parte, el **Periodograma** funciona como un detector de ciclicidad. Su objetivo es estimar la densidad espectral de potencia de la serie, identificando de manera explícita cuáles son las frecuencias (y, por lo tanto, los periodos) que contribuyen con mayor dominancia al comportamiento de la serie. Su cálculo requiere dos insumos clave: la serie de tiempo a analizar y la frecuencia de muestreo (fs), que para datos trimestrales se establece en fs=4 (cuatro datos por año). El análisis arrojó dos resultados reveladores:

1. Al aplicar el periodograma a la serie de nivel (log_pib_mx_usd), se encontró una frecuencia dominante extremadamente baja, correspondiente a un periodo de **32.00 años**, lo cual captura la tendencia de muy largo plazo de la serie.

2. Al aplicarlo a la serie de crecimiento ya estacionaria (crecimiento_pib_mx_usd), el resultado fue muy distinto: la frecuencia dominante fue de **1.3858 ciclos por año**, lo que equivale a un ciclo que se repite cada **0.72 años (o 2.89 trimestres)**. Este hallazgo evidencia la existencia de un importante componente cíclico sub-anual en la tasa de crecimiento de la economía.

De forma complementaria a este análisis espectral, se utiliza la **metodología Box-Jenkins** para modelar la estructura de dependencia en el dominio del tiempo. Mientras que la TDF descompone la serie, el modelo ARIMA busca capturar la correlación estocástica entre una observación y sus valores pasados. Siguiendo las fases de identificación, estimación y validación, y utilizando criterios de información como el AIC, se determinó que el modelo **ARIMA(1,0,1)** es el más adecuado para describir el comportamiento de la serie de crecimiento estacionaria. Este modelo captura la "memoria de corto plazo" de la serie, complementando la visión de largo plazo y cíclica que ofrece el análisis de Fourier.

Modelo computacional

La implementación de este análisis se llevó a cabo íntegramente en el lenguaje de programación Python, apoyándose en un ecosistema de librerías especializadas en el manejo de datos, cálculo numérico y modelización estadística. A continuación, se detalla el flujo computacional seguido en el estudio.

1. **Carga y Preparación del Entorno**: El proceso inicia con la importación de las librerías esenciales: pandas para la manipulación de datos estructurados, numpy para operaciones numéricas, matplotlib y seaborn para la visualización gráfica, y los módulos específicos de statsmodels y scipy para el análisis de series de tiempo y el procesamiento de señales. Los datos fueron cargados desde un archivo de Microsoft Excel (politica_comercial.xlsx) a un DataFrame de pandas. Posteriormente, se realizó un paso crítico que consistió en establecer la columna periodo como el índice del DataFrame, una operación indispensable para que pandas reconozca la estructura de serie temporal de los datos. Finalmente, se ejecutó una verificación para confirmar la ausencia de valores nulos, asegurando la integridad del conjunto de datos.

2. **Transformación y Homogeneización de Datos**: Dado que los modelos estadísticos requieren datos homogéneos y estacionarios, se implementó una secuencia de transformaciones:

- **Conversión de Divisa**: Para estandarizar las unidades, las variables de México (PIB, exportaciones e importaciones) se convirtieron a dólares estadounidenses dividiendo sus valores en pesos por la serie del tipo de cambio (tcn). A su vez, el PIB de EE. UU., que estaba en miles de millones, se ajustó a millones multiplicándolo por 1,000.
- **Transformación Logarítmica**: Se aplicó la función numpy.log a las series ya dolarizadas. Este paso es fundamental para estabilizar la varianza y permitir que las diferencias se interpreten como tasas de crecimiento porcentual, un concepto conocido como elasticidad.
- **Diferenciación**: Para alcanzar la estacionariedad, se calculó la primera diferencia de las series logarítmicas utilizando el método ".diff()" de pandas. Esta operación da como resultado la serie de crecimiento trimestral, que es la base para la modelización ARIMA.

3. **Implementación de Modelos Estadísticos**: Se utilizó la función adfuller de statsmodels para realizar la prueba de raíz unitaria y confirmar la estacionariedad. Las funciones plot_acf y plot_pacf se emplearon para visualizar las funciones de autocorrelación y guiar la elección de los órdenes p y q. Finalmente, se utilizó la clase ARIMA de statsmodels.tsa.arima.model para estimar los modelos candidatos y comparar sus criterios de información (AIC y BIC).
   Análisis de Frecuencia:
   - **Transformada de Fourier y Filtros**: Se emplearon las funciones fft, fftfreq, y ifft de la librería scipy.fft. El cómputo se realizó aplicando fft a la serie, creando máscaras booleanas para anular las frecuencias no deseadas (altas para el filtro pasa-baja, bajas para el filtro pasa-alta) y finalmente aplicando ifft para regresar al dominio del tiempo y obtener las series de tendencia y ciclo, respectivamente.
   - **Periodograma**: Se utilizó la función periodogram de scipy.signal, proporcionándole como argumentos la serie de tiempo y la frecuencia de muestreo (fs=4). El resultado se visualizó con matplotlib, y se usó la función numpy.argmax para identificar programáticamente la frecuencia con la máxima potencia espectral.

A lo largo de todo el proceso, las capacidades gráficas de matplotlib fueron extensivamente utilizadas para generar los gráficos que permitieron la inspección visual de las series, la interpretación de los resultados de los modelos y la comunicación efectiva de los hallazgos.

Metodología.

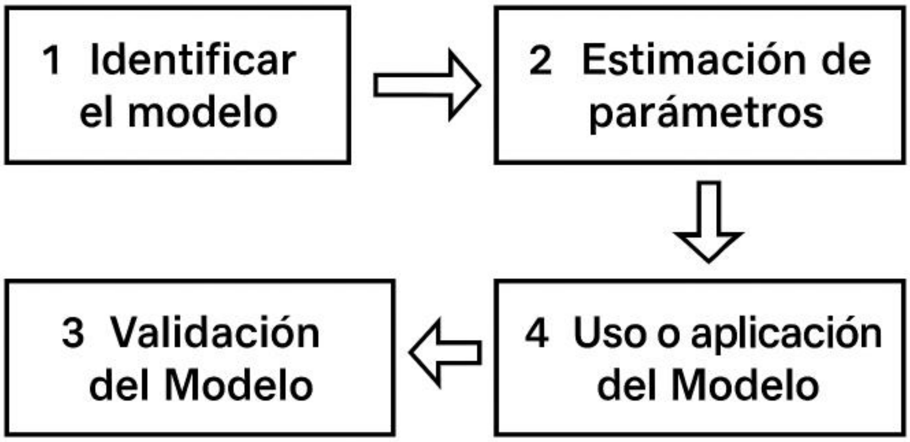


Figura 1. Proceso metodológico de Box-Jenkins.

Se empleó la **Metodología Box-Jenkins**, con un enfoque iterativo para la construcción de modelos ARIMA, que consta de las siguientes etapas principales:

1. **Identificación del Modelo:**

Análisis exploratorio de datos (EDA): Visualización de la serie, cálculo de estadísticas descriptivas, análisis de tendencia, estacionalidad, ciclos, picos y valles.

Verificación de estacionariedad: Uso de la prueba Dickey-Fuller para determinar si la serie es estacionaria en media y varianza.

- Transformación: Aplicación de diferencias (d) hasta lograr la estacionariedad en este caso no fue necesario porque ya era estacionaria. También se podría considerar transformación logarítmica si la varianza no es constante.
- Análisis ACF y PACF: Estudio de las gráficas de autocorrelación y autocorrelación parcial de la serie *estacionaria* para proponer órdenes iniciales para p (AR) y q (MA).

Análisis en el Dominio de la Frecuencia (Análisis Espectral): Este enfoque se utiliza para descomponer la serie e identificar sus componentes cíclicos.

- Descomposición con TDF: Se aplicó la Transformada Discreta de Fourier sobre la serie de nivel logarítmica (log_pib_mx_usd). Mediante filtros pasa-baja y pasa-alta, se logró separar la serie en sus dos componentes principales: la tendencia de largo plazo y el ciclo de corto plazo.
- Detección de Ciclos con Periodograma: Se utilizó el periodograma sobre la serie de crecimiento estacionaria para identificar las frecuencias cíclicas dominantes. Esto permitió cuantificar la periodicidad de las fluctuaciones económicas, revelando un ciclo principal de aproximadamente 2.89 trimestres.

2. **Estimación de Parámetros:**

Ajuste del modelo ARIMA(p, d, q) propuesto a los datos de entrenamiento (una porción de la serie histórica).
Obtención de los coeficientes del modelo y sus errores estándar, así como criterios de información (AIC, BIC).

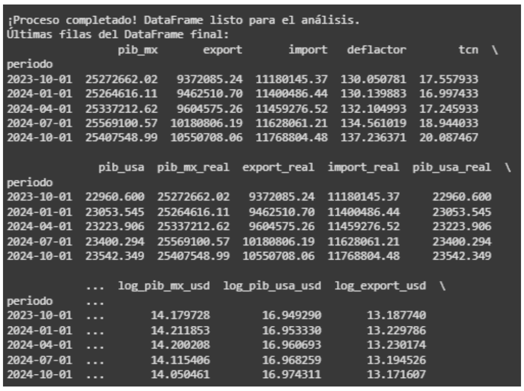
Previo a esto se hizo un preprocesamiento de datos de tal forma:

1. **Preparación y Homogeneización de los Datos**: El primer paso consistió en abordar la heterogeneidad de las variables. Para crear una base de comparación coherente, todas las series económicas relevantes se estandarizaron a una unidad común: millones de dólares estadounidenses. Esto se logró mediante dos operaciones clave:

- Las variables de la economía mexicana (PIB, exportaciones e importaciones), originalmente en pesos, se convirtieron a dólares dividiéndolas por la serie del tipo de cambio nominal (tcn).
- El PIB de EE. UU., medido en miles de millones (billones) de dólares, se ajustó a millones multiplicándolo por 1,000.

2. **Transformación para Alcanzar la Estacionariedad**: Con los datos ya en una unidad común, el siguiente paso se centró en tratar la no estacionariedad, una característica intrínseca de las series de nivel económico.

- Se aplicó una **transformación logarítmica** (np.log) a las series homogeneizadas. Este procedimiento es fundamental para estabilizar la varianza de la serie, linealizar posibles relaciones exponenciales y permitir que los cambios se interpreten como tasas de crecimiento porcentual (elasticidad).
- Posteriormente, se calculó la **primera diferencia** (.diff()) de las series logarítmicas. Esta es la operación crucial que convierte la serie de "nivel" (no estacionaria, con tendencia) en una serie de "cambio" o "tasa de crecimiento" (crecimiento_pib_mx_usd), la cual es teóricamente estacionaria.

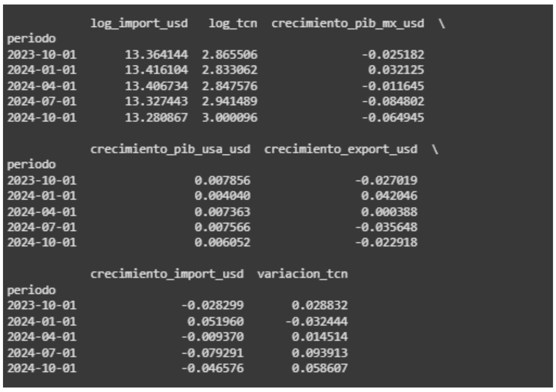


```
¡Proceso completado! DataFrame listo para el análisis.
Últimas filas del DataFrame final:
                pib_mx      export      import   deflactor        tcn  \
periodo
2023-10-01  25272662.02   9372085.24  11180145.37  130.050781  17.557933
2024-01-01  25264616.11   9462510.70  11400486.44  130.139883  16.997433
2024-04-01  25337212.62   9604575.26  11459276.52  132.104993  17.245933
2024-07-01  25569100.57  10180806.19  11628061.21  134.561019  18.944033
2024-10-01  25407548.99  10550708.06  11768804.48  137.236371  20.087467

              pib_usa  pib_mx_real  export_real  import_real  pib_usa_real  \
periodo
2023-10-01  22960.600  25272662.02   9372085.24  11180145.37     22960.600
2024-01-01  23053.545  25264616.11   9462510.70  11400486.44     23053.545
2024-04-01  23223.906  25337212.62   9604575.26  11459276.52     23223.906
2024-07-01  23400.294  25569100.57  10180806.19  11628061.21     23400.294
2024-10-01  23542.349  25407548.99  10550708.06  11768804.48     23542.349

            ...  log_pib_mx_usd  log_pib_usa_usd  log_export_usd  \
periodo     ...
2023-10-01  ...       14.179728        16.949290       13.187740
2024-01-01  ...       14.211853        16.953330       13.229786
2024-04-01  ...       14.200208        16.960693       13.230174
2024-07-01  ...       14.115406        16.968259       13.194526
2024-10-01  ...       14.050461        16.974311       13.171607
```

Figura 2. Tratamiento de Datos parte 1.



```
            log_import_usd   log_tcn  crecimiento_pib_mx_usd  \
periodo
2023-10-01       13.364144  2.865506               -0.025182
2024-01-01       13.416104  2.833062                0.032125
2024-04-01       13.406734  2.847576               -0.011645
2024-07-01       13.327443  2.941489               -0.084802
2024-10-01       13.280867  3.000096               -0.064945

            crecimiento_pib_usa_usd  crecimiento_export_usd  \
periodo
2023-10-01                 0.007856               -0.027019
2024-01-01                 0.004040                0.042046
2024-04-01                 0.007363                0.000388
2024-07-01                 0.007566               -0.035648
2024-10-01                 0.006052               -0.022918

            crecimiento_import_usd  variacion_tcn
periodo
2023-10-01               -0.028299       0.028832
2024-01-01                0.051960      -0.032444
2024-04-01               -0.009370       0.014514
2024-07-01               -0.079291       0.093913
2024-10-01               -0.046576       0.058607
```

Figura 3. Tratamiento de Datos parte 2.

Propuesta de solución para el uso del Periodograma y transformación discreta de fourier.

Para superar las limitaciones de un análisis puramente en el dominio del tiempo y obtener una comprensión más profunda de la dinámica económica, se propone una solución analítica basada en las herramientas del dominio de la frecuencia. La propuesta consiste en utilizar la Transformada Discreta de Fourier (TDF) como un método de descomposición de la serie, y el Periodograma como un detector de ciclos dominantes.

Descomposición de la Serie mediante la Transformada Discreta de Fourier (TDF)

La TDF se propone como una solución para separar la serie de nivel logarítmica (log_pib_mx_usd), que es compleja y no estacionaria, en dos componentes más simples y analíticamente manejables: la tendencia de largo plazo y el ciclo de corto plazo. Esto se logra mediante la aplicación de filtros de frecuencia:

- **Filtro Pasa-Baja para Aislar la Tendencia**: La solución consiste en transformar la serie al dominio de la frecuencia usando la TDF y, a continuación, eliminar las frecuencias altas que representan el ruido y las fluctuaciones de corto plazo. Se define una frecuencia de corte (ej. 0.33 ciclos/año, equivalente a ciclos de 3 años de duración) y se anulan mediante programación todas las frecuencias superiores a este umbral. Al aplicar la transformada inversa de Fourier (ifft), se reconstruye una nueva serie temporal que representa la trayectoria suavizada y fundamental de la economía, es decir, su **tendencia de largo plazo**. Esta serie filtrada permite visualizar y analizar los grandes ciclos económicos sin la distracción del "ruido" trimestral.

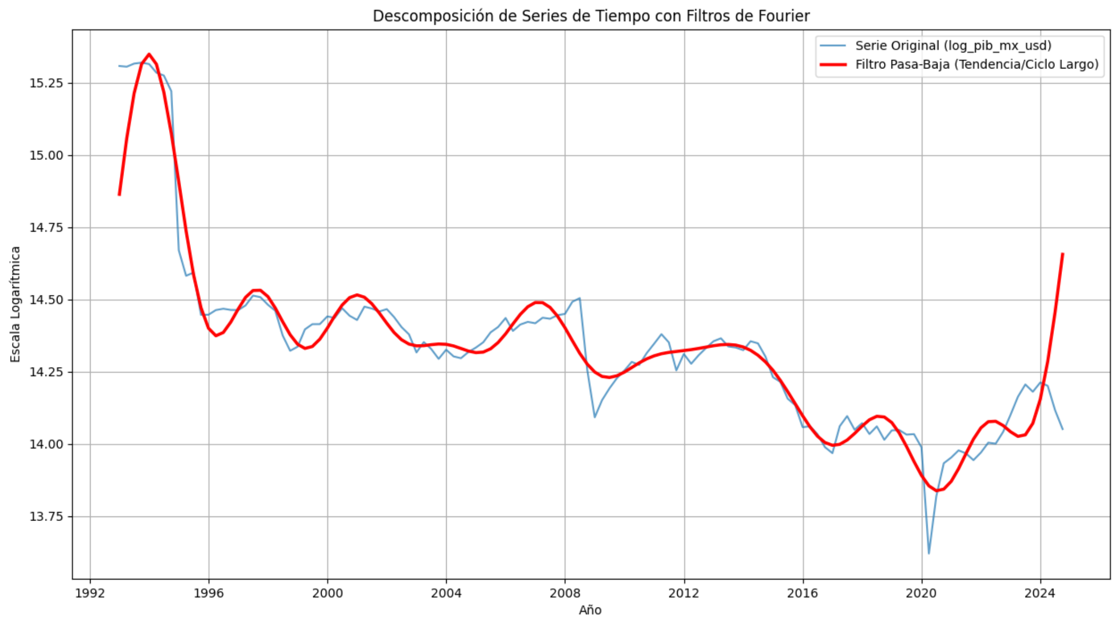


Figura 4. Filtro Pasa-Baja Fourier.

Alumno: Jiménez Flores Luis Arturo
Grupo: 6AV1
Materia: Análisis de Series de Tiempo

- **Filtro Pasa-Alta para Aislar el Ciclo**: De forma complementaria, se propone aislar el componente cíclico de corto plazo. Para ello, se anulan las frecuencias que están *por debajo* de la frecuencia de corte, conservando únicamente las fluctuaciones rápidas. La serie resultante, tras la inversión de la transformada, representa las **desviaciones de la tendencia**. Este componente es estacionario, fluctúa en torno a cero y encapsula el ciclo económico de corto plazo, listo para ser analizado o modelado por separado.

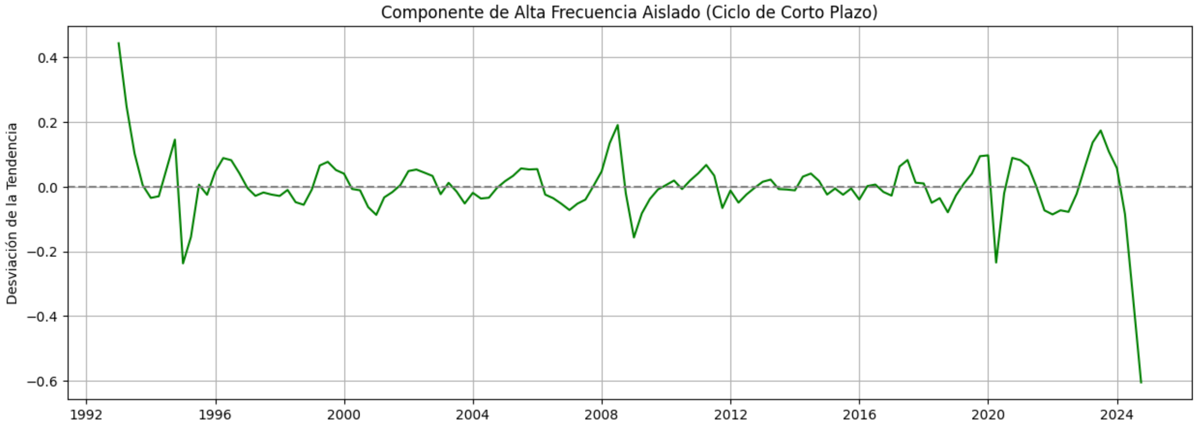


Figura 5. Filtro Pasa-Alta Fourier.

En conjunto, esta técnica de filtrado ofrece una solución empírica y robusta para descomponer una serie no estacionaria en sus partes constitutivas, permitiendo un análisis diferenciado de la dirección fundamental de la economía y sus oscilaciones periódicas.

Detección de Ciclos Dominantes con el Periodograma

Una vez que se ha entendido la importancia de los ciclos, se propone el uso del Periodograma como la herramienta principal para cuantificarlos de manera precisa. La solución es aplicar esta técnica sobre la serie de crecimiento estacionaria (crecimiento_pib_mx_usd), ya que esto permite identificar la periodicidad de las *fluctuaciones* del crecimiento, no de la tendencia general.

El procedimiento consiste en calcular el periodograma de la serie, utilizando una frecuencia de muestreo fs=4 apropiada para los datos trimestrales. El resultado es un gráfico que muestra la "potencia" o importancia de cada frecuencia. La propuesta de solución se centra en el hallazgo clave de este análisis:

- La identificación de un pico de potencia en el periodograma en la frecuencia de **1.3858 ciclos por año**.
- La traducción de esta frecuencia a un periodo tangible y fácil de interpretar: un ciclo económico dominante que se repite cada **0.72 años**, o lo que es equivalente, cada **2.89 trimestres**.

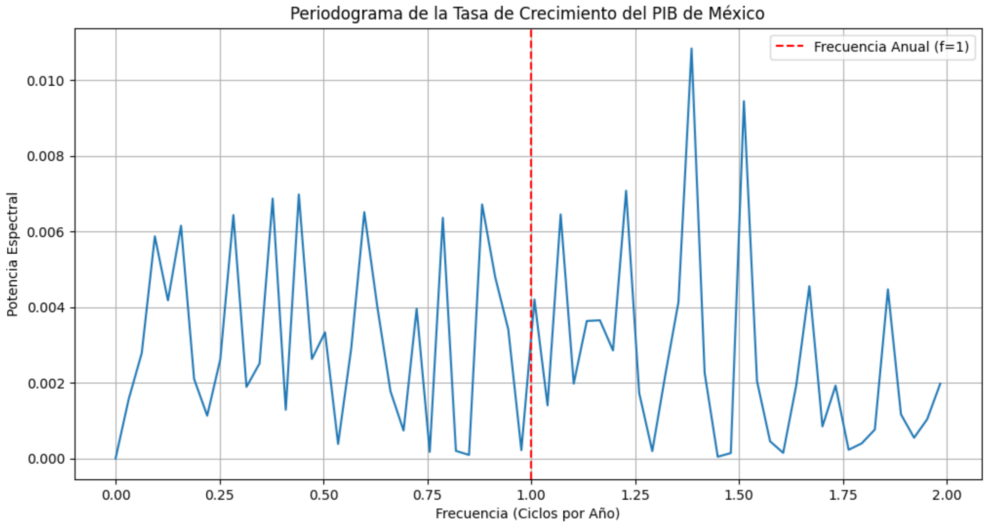


Figura 6. Análisis de Periodograma.

Por lo tanto, la solución que ofrece el periodograma es una respuesta cuantitativa y específica a la pregunta sobre la regularidad de los ciclos económicos. Proporciona una medida concreta de la periodicidad más importante en la tasa de crecimiento del PIB, una información de gran valor para el monitoreo económico y la anticipación de fluctuaciones a corto plazo.

Conclusiones por integrante.

Jiménez Flores Luis Arturo

Al finalizar este exhaustivo análisis de las series de tiempo económicas de México, con fecha del 13 de junio de 2025, es posible concluir que se han alcanzado con éxito los objetivos propuestos.

En una primera instancia, el estudio demostró la importancia crítica de un adecuado pre-procesamiento de los datos. La transformación de las series mediante la conversión a una moneda común (dólares estadounidenses), la aplicación de logaritmos para estabilizar la varianza y la diferenciación para alcanzar la estacionariedad, fueron pasos indispensables que permitieron construir una base sólida para cualquier modelización posterior. Sobre esta base, la metodología Box-Jenkins nos condujo a la identificación de un modelo ARIMA(1,0,1) como el más eficiente para capturar la dependencia estocástica a corto plazo en la tasa de crecimiento del PIB (tentativamente).

Sin embargo, la aportación más reveladora de este estudio reside en la aplicación de las herramientas de análisis espectral. La **Transformada Discreta de Fourier (TDF)** demostró ser mucho más que un simple ejercicio técnico; se erigió como una poderosa lente conceptual dada su capacidad para descomponer, mediante filtros pasa-baja y pasa-alta, una única y compleja serie económica en sus dos narrativas fundamentales la **tendencia de largo plazo** y el **ciclo de corto plazo**. Esta separación nos permite estudiar de forma aislada la trayectoria estructural de la economía y, por otro lado, las fluctuaciones de negocio que ocurren a su alrededor, dos fenómenos que responden a fuerzas distintas.

Complementariamente, el **Periodograma** le puso un ritmo tangible y cuantificable a esas fluctuaciones. En lugar de una apreciación visual subjetiva de los ciclos, esta herramienta nos proporcionó una medida precisa: un ciclo dominante en la tasa de crecimiento económico que se repite, en promedio, cada **2.89 trimestres**. Este hallazgo dota de una periodicidad concreta a la volatilidad de la economía, ofreciendo un dato de valor para el monitoreo y la anticipación de cambios a corto plazo.

Dados estos resultados se puede concluir que la práctica para el uso de la transformada discreta de Fourier para series de tiempo resultó ser un éxito.