Modelos de series de tiempo

Proyecto

Fecha de entrega: 11 de diciembre

Instrucciones. El proyecto se realizará en equipos de máximo tres integrantes. Cada equipo trabajará con una de las series de tiempo asignadas. Los resultados deberán presentarse en un reporte de 10 a 12 cuartillas como máximo. Se espera un documento organizado que presente los resultados más relevantes, evitando incluir gráficas, tablas o procedimientos que no contribuyan directamente al análisis. Se deberá justificar la selección del modelo utilizando las herramientas teóricas y estadísticas estudiadas durante el curso, así como evaluar su capacidad predictiva y compararla con métodos alternativos. El código utilizado deberá entregarse en un archivo separado y permitir reproducir los resultados numéricos y las gráficas presentadas en el reporte.

1. Realiza un análisis exploratorio de la serie de tiempo que permita identificar sus principales características, incluyendo la presencia de tendencias, estacionalidad, variaciones en la dispersión y posibles cambios estructurales en su comportamiento. Para ello

   (a) Representa gráficamente la serie de tiempo e interpreta sus características principales.

   (b) Determina si es necesario aplicar transformaciones o diferenciaciones ordinarias y/o estacionales para obtener una serie aproximadamente estacionaria. Justifica las transformaciones realizadas a partir del ACF, PACF y periodograma.

2. A partir del análisis anterior y de las propiedades teóricas de los modelos ARMA, ARIMA y SARIMA, propón al menos tres modelos candidatos que puedan describir adecuadamente la serie (sin utilizar la función auto.arima). Para cada modelo:

   (a) Especifica su expresión matemática utilizando el operador de rezago (B).

   (b) Justifica la elección de los órdenes autorregresivos, de medias móviles y, cuando corresponda, de diferenciación y estacionalidad.

3. Estima los parámetros de los modelos propuestos utilizando las herramientas computacionales apropiadas.

(a) Reporta las estimaciones de los parámetros y sus errores estándar.

(b) Calcula los criterios de información AIC y BIC para comparar los modelos candidatos.

(c) Verifica las condiciones de causalidad e invertibilidad correspondientes a las componentes ARMA estacionarias de los modelos considerados.

(d) Realiza un análisis de residuales mediante gráficas, funciones de autocorrelación y pruebas estadísticas apropiadas, incluyendo la prueba de Ljung–Box.

(e) Discute si los residuales son compatibles con las propiedades esperadas de un proceso de ruido blanco y si los modelos proporcionan una representación adecuada de la serie.

(f) Con base en los resultados anteriores, selecciona un modelo principal, justificando tu decisión.

4. Divide cronológicamente la serie de tiempo en dos conjuntos:

   (a) Conjunto de entrenamiento: todas las observaciones, excepto las últimas (h).

   (b) Conjunto de prueba: las últimas (h) observaciones, donde (h) corresponde al horizonte de pronóstico establecido para la serie de tiempo asignada.

   El valor de (h) dependerá de la frecuencia de la serie: si es mensual $h = 12$, trimestral $h = 8$ y anual $h = 5$.

   (a) Identifica, estima y selecciona un modelo utilizando exclusivamente el conjunto de entrenamiento.

   (b) Utiliza el modelo seleccionado para obtener pronósticos fuera de muestra y evalúa su precisión mediante la raíz del error cuadrático medio (RMSE).

   (c) Compara los resultados obtenidos con un modelo seleccionado automáticamente mediante auto.arima() y de manera opcional con un modelo de espacio de estados.

   (d) Presenta una tabla comparativa con los valores de RMSE e interpreta los resultados. Discute si el modelo seleccionado manualmente presenta ventajas frente a los métodos alternativos.

5. Una vez evaluado el modelo, vuelve a estimar sus parámetros utilizando la serie completa. Obtén pronósticos para el horizonte indicado en el inciso anterior.

   (a) Reporta los pronósticos puntuales y construye sus intervalos al 95%.

   (b) Interpreta los resultados, discutiendo la evolución esperada de la serie, el comportamiento de la incertidumbre y las limitaciones del modelo para realizar pronósticos.

6. Presenta una discusión de los resultados más importantes del análisis. En particular, explica si el modelo seleccionado describe adecuadamente el comportamiento de la serie, qué tan satisfactoria fue su capacidad predictiva y cuáles son sus principales limitaciones.