



Tarea 1 Módulo 2 Series de Tiempo

Anselmo Daniel Suarez Muñoz

21 de octubre de 2024

1. PROBLEMA 1

(15 pts) Supongamos que tienes un vector $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{4t})$ donde Y_{3t} tiene un cambio estructural, Y_{1t} tiene estacionalidad, Y_{2t} es estacionaria y Y_{4t} es $I(1)$. Te piden que estimes un modelo VAR, ¿cómo procederías a hacerlo?

1.1. SOLUCIÓN

De acuerdo al problema planteado la forma de proceder a estimar el modelo en base a la teoría vista, primero lo que haría sería tratar de hacer estacionarias las series que no lo son y tratar de eliminar las características que nos causarían problemas a la hora de la estimación, la interpretación de resultados y las futuras predicciones, claro desde el punto de vista univariado.

Con lo anterior me refiero a, por ejemplo, para la serie Y_{1t} que tiene estacionalidad aplicaría una diferenciación en el periodo de estacionalidad para eliminar la estacionalidad, por ejemplo, supongamos que la estacionalidad empieza en un tiempo t_0 y termina en t_1 , entonces llamemos a ese periodo $s = t_1 - t_0$ por lo que la diferenciación quedaría de la siguiente manera

$$Y'_{1t} = Y_{1t} - Y_{1(t-s)}$$

donde Y'_{1t} es la nueva serie univariada.

Un razonamiento similar se haría para la serie Y_{4t} se aplicaría también una diferenciación para evitar la integración, la nueva serie sería

$$\Delta Y_{4t} = Y_{4t} - Y_{4(t-1)}$$

Para el caso de la Y_{3t} deberíamos analizar ese cambio estructural, puede que sea estacional o puede que tenga algún cambio significativo, dependiendo de dicho cambio se procedería a hacer algún ajuste

para tratar dicha serie, se podrían analizar la ACF y ver su comportamiento apartir de allí hacer alguna diferenciación para volver la serie estacionaria. Tal vez agregar una variable dummy que capture el cambio. Así el nuevo vector de series de tiempo estaría compuesto por las nuevas series transformadas

$$Z_t = (Y'_{1t}, Y_{2t}, Y'_{3t}, \Delta Y_{4t})$$

Después procedería a identificar el número de rezagos del modelo, es decir el orden. Posterior a esto procedemos a estimar los coeficientes con algún método como GLS el cual nos resulta más eficiente para los modelos VAR, por último realizar un análisis de los resultados de ajuste del modelo

2. PROBLEMA 2

Para un modelo $VAR(p)$, genera un procedimiento para obtener matrices A_i estacionarias controlando la estabilidad del modelo como:

- i) leve,
- ii) moderada,
- iii) fuerte.

Hint: utiliza la representación $VAR(1)$ para un modelo VAR de orden p y estudia los valores propios de la solución de $|I_{Kp} - zA|$.

2.1. SOLUCIÓN

Para un modelo $VAR(p)$, es necesario garantizar la estacionariedad mediante el control de la estabilidad. Esto se puede lograr utilizando la representación del modelo $VAR(1)$ y analizando los valores propios de la matriz que define el sistema. Un modelo $VAR(p)$ se puede reescribir en forma de $VAR(1)$ como sigue. Un modelo $VAR(p)$ se representa como:

$$\mathbf{Y}_t = A_1 \mathbf{Y}_{t-1} + A_2 \mathbf{Y}_{t-2} + \cdots + A_p \mathbf{Y}_{t-p} + \mathbf{u}_t$$

Donde A_i son matrices de coeficientes y \mathbf{u}_t es el vector de términos de error. Esta ecuación se puede reformular como un modelo $VAR(1)$ de la siguiente forma:

$$\mathbf{Z}_t = \begin{pmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \\ \vdots \\ \mathbf{Y}_{t-p+1} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & \cdots & A_p \\ I & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix} \mathbf{Z}_{t-1} + \begin{pmatrix} \mathbf{u}_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Llamemos a la matriz del lado derecho \mathbf{A} . El sistema ahora tiene la forma $\mathbf{Z}_t = \mathbf{A}\mathbf{Z}_{t-1} + \mathbf{u}_t$, donde la estabilidad del modelo depende de los valores propios de la matriz \mathbf{A} .

Para controlar la estabilidad del modelo $VAR(p)$, es necesario que los valores propios de la matriz \mathbf{A} estén dentro del círculo unitario.

Para generar las matrices A_i estacionarias primero llevemos la matriz \mathbf{A} a una matriz semejante \mathbf{A}' por medio de operaciones elementales a la matriz de bloques \mathbf{A} , dichas operaciones son similares a las operaciones elementales entre filas y columnas de una matriz común.

Supongamos que $\det(A_1) \neq 0$ observamos que al multiplicar por $-A_1^{-1}$ a la primera fila de la matriz por bloques \mathbf{A} y sumamos a la segunda obtenemos la siguiente matriz:

$$\begin{pmatrix} A_1 & A_2 & \cdots & A_p \\ 0 & -A_1^{-1}A_2 & \cdots & -A_1^{-1}A_p \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix}$$

De forma similar si al multiplicar por $(A_1^{-1}A_2)^{-1} = A_2^{-1}A_1$ a la segunda fila de la matriz por bloques \mathbf{A} y sumamos a la tercera fila obtenemos la siguiente matriz:

$$\begin{pmatrix} A_1 & A_2 & \cdots & A_p \\ 0 & -A_1^{-1}A_2 & \cdots & -A_1^{-1}A_p \\ 0 & 0 & \cdots & -A_2^{-1}A_p \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix}$$

Realizando este procedimiento de forma iterativa se eliminarán las matrices I y obtendremos la siguiente matriz:

$$\begin{pmatrix} A_1 & A_2 & \cdots & A_p \\ 0 & -A_1^{-1}A_2 & \cdots & -A_1^{-1}A_p \\ 0 & 0 & \cdots & -A_2^{-1}A_p \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -A_{p-1}^{-1}A_p \end{pmatrix}$$

Después si multiplicamos por A_1 a la segunda fila y la sumamos a la primera fila obtenemos la siguiente matriz

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & -A_1^{-1}A_2 & \cdots & -A_1^{-1}A_p \\ 0 & 0 & \cdots & -A_2^{-1}A_p \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -A_{p-1}^{-1}A_p \end{pmatrix}$$

Así mismo, si multiplicamos por A_2 a la tercera fila y la sumamos a la segunda fila obtenemos la siguiente matriz

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & -A_1^{-1}A_2 & \cdots & 0 \\ 0 & 0 & \cdots & -A_2^{-1}A_p \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -A_{p-1}^{-1}A_p \end{pmatrix}$$

Nuevamente realizando este procedimiento de forma iterativa se eliminarán las matrices fuera de la diagonal y obtendremos la siguiente matriz:

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & -A_1^{-1}A_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -A_{p-1}^{-1}A_p \end{pmatrix}$$

Así la matriz resultante será la matriz obtenida de multiplicar por -1 apartir de la segunda fila hasta la última, es decir se obtiene:

$$\mathbf{A}' = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_1^{-1}A_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{p-1}^{-1}A_p \end{pmatrix}$$

Así por ser matrices semejantes \mathbf{A} y \mathbf{A}' los valores propios de la primera corresponden o son iguales a los valores propios de la segunda, de esta forma para calcular los valores propios de \mathbf{A}' es equivalente a calcular el polinomio característico de dicha matriz o equivalentemente

$$P_{\mathbf{A}'}(s) = P_{A_1}(s)P_{A_1^{-1}A_2}(s) \cdots P_{A_{p-1}^{-1}A_p}(s)$$

Luego sabemos que un modelo $VAR(p)$ es estable si los valores propios de la matriz \mathbf{A} caen dentro del círculo unitario, lo que equivale a que todas las raíces del polinomio característico de \mathbf{A}' sean menores o iguales a 1 en módulo, una forma de garantizar esto último es que los valores propios de A_1 y cada matriz de la forma $B = A_{i-1}^{-1}A_i$ para $i = 1, \dots, p$ también sean menores o iguales a 1 en módulo.

Luego la forma de generar dichas matrices A_i se describe a continuación:

- 1 Primero generamos A_1 de tal forma que sus valores propios sean menores o iguales a 1, además de que $\det(A_1) \neq 0$. Para controlar la estabilidad del modelo, en particular para esta matriz A_1 podemos generar una matriz diagonal D de forma aleatoria con valores en la diagonal menores que 1 y cercanos a 1 para evitar tener desestabilidad en las siguientes matrices. También generar vectores v_i columna aleatorios, de forma que el determinante de la matriz Q generada por dichos vectores columna sea distinto de cero, en notación se tendría lo siguiente:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$$Q = (v_1|v_2|\dots|v_n)$$

Así por la descomposición espectral se tiene que

$$A_1 = Q^{-1}DQ$$

2 Generar A_i siguiendo el paso 1, de tal forma que $\det(A_i) \neq 0$ y demás los valores propios de $B = A_{i-1}^{-1}A_i$ sean menores o iguales que 1, para $i = 2, \dots, p$

Así podremos decir que entre los valores propios de \mathbf{A} más cercanos a 1 estén diremos que la estabilidad es débil, y entre más lejos estén diremos que la estabilidad es fuerte.

3. PROBLEMA 3

Para un $VAR(p)$, donde K es el número de series de tiempo y T el número de observaciones, propón un criterio empírico para determinar $p_{\max} = f(K, T)$ en la selección de rezagos óptimos y comenta brevemente sus propiedades para algún conjunto de datos simulados.

3.1. SOLUCIÓN

Si el objetivo es la predicción, es lógico seleccionar el orden del modelo VAR (Vector Autoregression) de manera que se minimice el error cuadrático medio (MSE) de pronóstico. Akaike propuso utilizar el MSE de pronóstico aproximado para determinar este orden, que se expresa como:

$$\Sigma_y(1) = \frac{T + Km + 1}{T} \Sigma_u,$$

donde:

- m es el orden del proceso VAR ajustado a los datos,
- T es el tamaño de la muestra,
- K es la dimensión de la serie temporal.

Para hacer que este criterio sea operativo, la matriz de covarianza del ruido blanco Σ_u debe ser reemplazada por una estimación. Akaike sugirió usar un estimador de mínimos cuadrados (LS) con ajuste de grados de libertad:

$$\Sigma_u(m) = \frac{T}{T - Km - 1} \hat{\Sigma}_u(m),$$

donde $\hat{\Sigma}_u(m)$ es el estimador de máxima verosimilitud (ML) de Σ_u obtenido al ajustar un modelo VAR(m). El criterio resultante se denomina criterio de error de predicción final (FPE), que se define como:

$$FPE(m) = \det \left(\frac{T + Km + 1}{T} \begin{bmatrix} T & \\ & T - Km - 1 \end{bmatrix} \hat{\Sigma}_u(m) \right) = \left(\frac{T + Km + 1}{T - Km - 1} \right)^K \det \hat{\Sigma}_u(m).$$

En este contexto, al aumentar el orden m , el determinante del estimador ML $\det(\hat{\Sigma}_u(m))$ disminuye, mientras que el término multiplicativo $\left(\frac{T + Km + 1}{T - Km - 1} \right)$ aumenta. La estimación del orden VAR se obtiene al equilibrar estas dos fuerzas.

La estimación del orden $\hat{p}(FPE)$ se elige de modo que:

$$FPE[\hat{p}(FPE)] = \min\{FPE(m) \mid m = 0, 1, \dots, M\}.$$

Es decir, se estiman modelos VAR de órdenes $m = 0, 1, \dots, M$ y se calculan los valores correspondientes de $FPE(m)$. El orden que minimiza los valores de FPE se elige como estimación para p .

4. PROBLEMA 4

(22.5 pts) Se te adjuntan dos bases de datos: `Clima.csv` y `Datos.csv`. La primera incluye datos de temperatura y precipitación promedio en México de 1985 a noviembre de 2022, y la segunda información de la inflación, actividad económica mensual, índice de precios y cotizaciones, y tipo de cambio nominal, todas para México de 1980 a mayo de 2022. Contesta lo siguiente:

- ¿La temperatura y la precipitación son causales en el sentido de Granger?
- ¿Hay relación causal entre alguna de las series de tiempo climáticas con alguna de las económicas o financieras? ¿Tiene sentido lo encontrado?
- Genera un modelo $VAR(p)$ de predicción para la temperatura y evalúa su funcionamiento fuera de muestra en los últimos 24 meses. ¿Funciona mejor que un $SARIMA(p, d, q)(P, D, Q)$?

4.1. SOLUCIÓN

- Para ver si hay causalidad en el sentido de Granger usamos la prueba para las series propuestas, al sacar las respectivas series temporales analizaremos los resultados obtenidos al contruir las series y ver la prueba de causalidad en el sentido de Granger, los resultados se adjuntan en la tabla(4.1)

Hipótesis	F-Statistic	p-value
Temperatura no causa a Precipitación	6.2339	1.309×10^{-10}
Precipitación no causa a Temperatura	3.9569	5.786×10^{-6}

Cuadro 4.1: Resultados de la prueba de causalidad de Granger entre temperatura y precipitación

Ambas series (temperatura y precipitación) tienen una relación bidireccional en el sentido de Granger ya que en ambos casos el p -valor es menor al nivel de significancia 0,05. Es decir, tanto los valores pasados de la temperatura ayudan a predecir la precipitación como los valores pasados de la precipitación ayudan a predecir la temperatura. Este tipo de relación podría indicar que existe algún tipo de interdependencia entre los fenómenos climáticos representados por las series.

- Se llevaron a cabo pruebas de causalidad de Granger entre los datos climáticos y las variables económicas. Los resultados se presentan a continuación:

Causalidad de Temperatura a Inflación, IGAE, IPC y TC

- **F-Test:** 1.8095
- **p-value:** 0.0002072

La hipótesis nula H_0 establece que "la Temperatura no causa Granger a la Inflación (o IGAE, IPC, TC)". Un p-value tan bajo (0.0002072) sugiere que **rechazamos la hipótesis nula**. Esto indica que hay evidencia significativa de que la temperatura **causa Granger** a las variables económicas mencionadas. Es decir, cambios en la temperatura parecen influir en la inflación, IGAE, IPC y el tipo de cambio, al menos en términos de relación temporal.

Causalidad de Precipitación a Inflación, IGAE, IPC y TC

- **F-Test:** 1.2626
- **p-value:** 0.08891

En este caso, la hipótesis nula H_0 establece que "la Precipitación no causa Granger a la Inflación (o IGAE, IPC, TC)". Aunque el p-value (0.08891) es mayor que el nivel de significancia comúnmente utilizado (0.05), está relativamente cerca. Esto sugiere que hay **poca evidencia para rechazar la hipótesis nula**. Por lo tanto, **no podemos concluir que la Precipitación causa Granger** a las variables económicas en este análisis. Sin embargo, podría ser un indicio de una relación débil que merece ser investigada más a fondo.

Conclusiones Generales

- **Relación Significativa:** La temperatura parece tener un efecto causal claro sobre las variables económicas analizadas, lo que puede tener implicaciones importantes para la formulación de políticas y la planificación económica en el contexto de cambios climáticos.
- c) Los resultados del modelo VAR se adjuntan en la tabla(4.2) y los resultados del modelo SARIMA en la tabla(4.3)

Métrica	ME	RMSE	MAE	MPE	MAPE
Test set	-0.3686884	0.6962002	0.561752	-2.011086	2.86729

Cuadro 4.2: Resultados del Modelo VAR

Métrica	ME	RMSE	MAE	MPE	MAPE
Test set	-0.4741299	0.7728234	0.6343892	-2.35145	3.129418

Cuadro 4.3: Resultados del Modelo SARIMA

El modelo VAR presenta las siguientes características:

- **Error Medio (ME):** -0.3686884, indicando que el modelo tiende a subestimar la temperatura en promedio.
- **Raíz del Error Cuadrático Medio (RMSE):** 0.6962002, lo que sugiere un buen ajuste en las predicciones.
- **Error Absoluto Medio (MAE):** 0.561752, indicando una precisión promedio en las predicciones.
- **Porcentaje Medio de Error (MPE):** -2.011086, que muestra que las predicciones son un 2.01 % inferiores a los valores reales.
- **Porcentaje Medio Absoluto de Error (MAPE):** 2.86729, lo que significa que, en promedio, las predicciones se desvían un 2.87 %.
- **Autocorrelación del Residuos (ACF1):** 0.356954, sugiriendo que los errores no están autocorrelacionados, lo que es positivo.

El modelo SARIMA muestra:

- **Error Medio (ME):** -0.4741299, también indicando subestimación.
- **Raíz del Error Cuadrático Medio (RMSE):** 0.7728234, mayor que el RMSE del VAR, lo que indica un menor ajuste.
- **Error Absoluto Medio (MAE):** 0.6343892, mayor que el del VAR, sugiriendo menor precisión.
- **Porcentaje Medio de Error (MPE):** -2.35145, con predicciones un 2.35 % inferiores.
- **Porcentaje Medio Absoluto de Error (MAPE):** 3.129418, que indica una desviación mayor en comparación con el VAR.
- **Autocorrelación del Residuos (ACF1):** 0.3109, sugiriendo autocorrelación en los residuos, lo que es negativo.

La siguiente gráfica muestra los resultados de las predicciones donde se observa que la roja se acerca de forma más natural a la azul la cual corresponde a las verdaderas predicciones

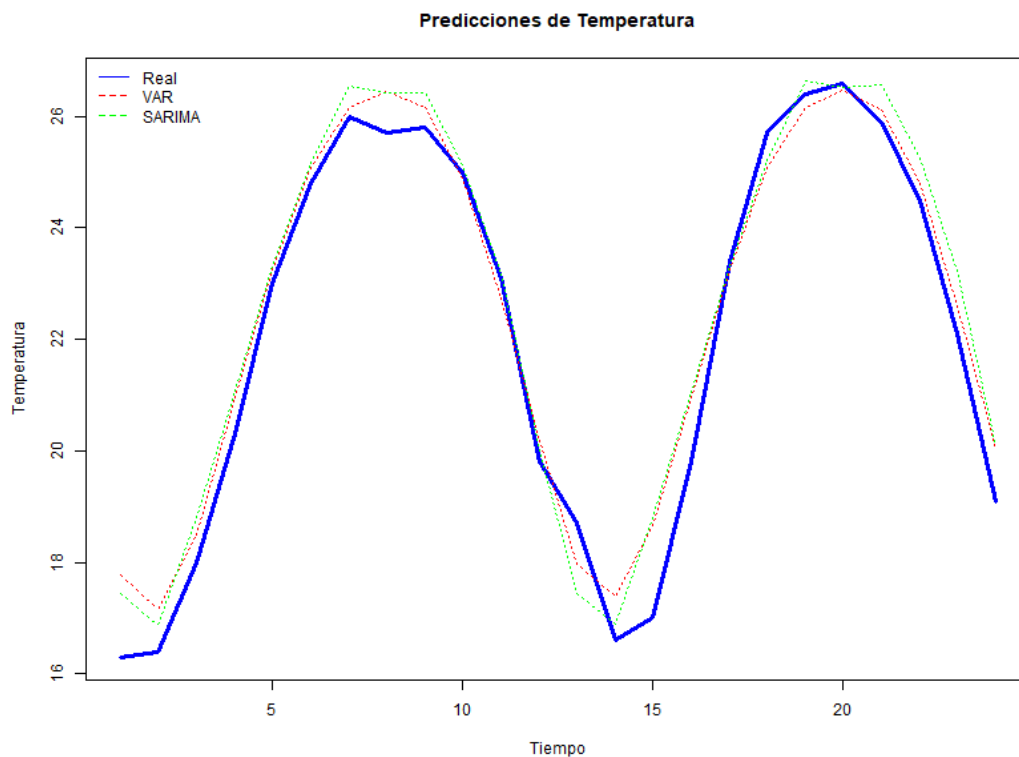


Figura 4.1: Gráfica de resultados de modelo VAR y SARIMA