Análisis Bayesiano en series temporales. Modelos BVAR

Bayesian analysis in time series. BVAR models

Cruz, Stefano[1]; Granda, Gabriel[1] ✉; Guamán, Ronny[1]; Sosa, Jorge [1];

[1] *Escuela Politécnica Nacional, Facultad de Ciencias, Quito, Ecuador*

✉ william.granda@epn.edu.ec

Fecha de envío: 15 de agosto 2021

Resumen: Los modelos VAR son conocidos por tener una buena habilidad predictiva. No obstante, cuando se utiliza una cantidad considerable de variables, estos modelos son susceptibles al problema de sobre-parametrización, lo cual conduce a una estimación deficiente de parámetros y a un pobre desempeño en términos de calidad de pronóstico. Por esta razón, se propone el uso de modelos bayesianos en la estimación de vectores auto-regresivos como una alternativa al problema de sobre-parametrización propio de los modelos VAR. Estos modelos reciben el nombre de VAR Bayesianos (BVAR).

Palabras clave: Modelo VAR, Estadística Bayesiana, Economía de Ecuador

Abstract: VAR models are known to have good predictive ability. However, when a considerable number of variables are used, these models are susceptible to the problem of over-parameterization, which leads to poor parameter estimation and poor performance in terms of forecast quality. For this reason, the use of Bayesian models in the estimation of auto-regressive vectors is proposed as an alternative to the over-parameterization problem typical of VAR models. These models are called Bayesian VAR (BVAR).

Key Words: VAR Model, Bayesian Statistics, Economics of Ecuador

1. Introducción

En [2] se encuentra que la regresión multivariante de un VAR viene dada por

$$Y_{T\times N} = X_{T\times k}B_{k\times N} + V_{T\times N} \qquad \text{y} \qquad V \sim MN(0, \Sigma_{N\times N}, I_T),$$

donde MN representa la distribución matricial normal. Luego, vectorizando, se obtiene que

$$y = (I_N \otimes X)\beta + v \qquad \text{y} \qquad v \sim \mathcal{N}(0, \Sigma \otimes I_T).$$

En un VAR bayesiano, la matriz B de coeficientes es aleatoria y se puede especificar como sigue

$$\beta \sim \mathcal{N}(\beta_0, \Omega_0),$$

donde, el vector β_0 es la media a priori, la matriz Ω_0 es la varianza alrededor de β_0 y mide la incertidumbre que tenemos sobre nuestras creencias anteriores.

Consideraremos el siguiente a priori:

Estimador mixto de Theil: El caso más simple es aquel en el que la varianza del error se estima en un paso preliminar y se trata como conocida. La distribución a posterior es:

$$\beta|y,\Sigma \sim \mathcal{N}(\beta_1,\Omega_1),$$

donde, la media y la varianza a posteriori son:

$$\Omega_1^{-1} = \Omega_0^{-1} + (\hat{\Sigma}^{-1} \otimes X'X) \qquad \text{y} \qquad \beta_1 = \Omega_1(\Omega_0^{-1}\beta_0 + (\hat{\Sigma}^{-1} \otimes X')y)$$

El llamado a priori de Minnesota es un caso especial del estimador mixto cuando Σ es función de un pequeño número de hiperparámetros. A priori, cada variable $t = 1,\ldots,T$, en el VAR es

$$y_t = 0 + Iy_{t-1} + 0y_{t-2} + \cdots + 0y_{t-p} + u_t$$

Además, se supone que Ω_0 es una matriz diagonal, los elementos diagonales correspondientes a las j-ésimas variables endógenas en la i-ésima ecuación en el retardo l vienen dadas por:

$$\Omega_{0(i,j)}^{l} = \begin{cases} \frac{\lambda_1}{l^{\lambda_3}} & \text{si} \quad i = j, \\ \frac{\lambda_1\lambda_2\sigma_i}{l^{\lambda_3}\sigma_j} & \text{si} \quad i \neq j, \end{cases}$$

Con $\lambda_1 \in [0,+\infty[$(mide la rigidez del prior), $\lambda_2 \in (0,1)$ (caracteriza la importancia relativa de los retardos de las otras variables con respecto a los propios retardos) y λ_3 (mide el grado de reducción de la varianza en función del retardo) hiperparámetros y $\frac{\sigma_i}{\sigma_j}$ se estiman a partir de regresiones AR univariadas. Normalmente los valores sugeridos para los hiperparámetros son:

$$\lambda_1 = 0,2, \lambda_2 = 1 \text{ y } \lambda_3 = 1.$$

2. METODOLOGÍA

En este apartado se tratará una comparación entre los modelos VAR y VAR Bayesianos, exponiendo los problemas de sobre parametrización de los modelos VAR y como solución a esto, la implementación de los modelos BVAR. Se considerarán cuatro series de datos económicos de Ecuador:

- Demanda Interna X_{1t}
- Formación Bruta de Capital Fijo X_{2t}
- Exportaciones X_{3t}
- PIB X_{4t}

Se dispone de 77 datos trimestrales desde el primer trimestre del 2001 hasta el primer trimestre del 2020. La gráfica de las series empleando el software Eviews es:

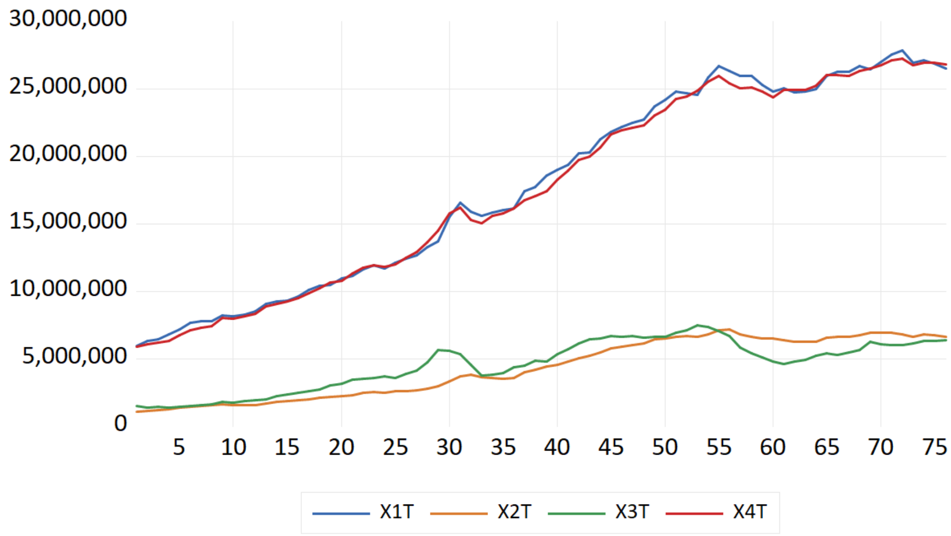

Figura 1: Serie de los datos económicos del Ecuador

Como podemos ver, estos datos son no estacionarios; por esta razón, se trabajará con las variaciones trimestrales de las series. Así, consideremos las siguientes notaciones:

- Y_{1t} := Variación trimestral de la Demanda Interna:
- Y_{2t} := Variación trimestral de la Formación Bruta de Capital Fijo;
- Y_{3t} := Variación trimestral de las Exportaciones; y,
- Y_{4t} := Variación trimestral del PIB.

El gráfico de las variaciones trimestrales de las series es:

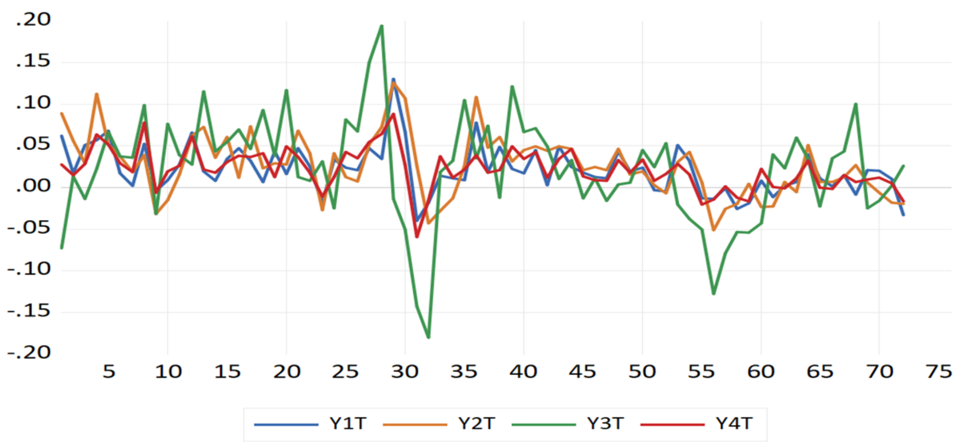

Figura 2: Gráfico de las variaciones trimestrales de las series

Notemos que trabajamos con 72 datos, dado que, al calcular las variaciones, se pierde el primer dato. Para comprobar que las variaciones de las series son estacionarias presentamos la prueba de Dickey - Fuller:

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-6.463121	0.0000
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

Figura 3: :Prueba Dickey y Fuller para Y_{1t}

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.847010	0.0001
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

Figura 4: :Prueba Dickey y Fuller para Y_{2t}

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.165628	0.0000
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

Figura 5: :Prueba Dickey y Fuller para Y_{3t}

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.506465	0.0000
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

Figura 6: :Prueba Dickey y Fuller para Y_{4t}

Considerando un nivel de confianza $\alpha = 0{,}05$, dado que todas las probabilidades son menores a α, rechazamos la hipótesis nula del test; es decir, las variaciones trimestrales de las series son estacionarias.

Para efectos de comparación se trabajará únicamente con 73 datos y se dejará los últimos 4 datos para comparar con predicciones posteriores (segundo trimestre del 2019 – primer trimestre del 2020).

El modelo VAR obtenido que cumple con las pruebas de estabilidad, pruebas de independencia de los residuos (prueba de Portmanteau) y la prueba de normalidad es un $VAR(4)$ con los siguientes retardos: 1,2,4,12.

El error medio cuadrático obtenido con este modelo es:

Modelo / Variable	Y_{1t}	Y_{2t}	Y_{3t}	Y_{4t}
VAR	0,02772395	0,03599093	0,06602036	0,02347054

Tabla 1: Error medio cuadrático del Modelo VAR

Modelo BVAR

Con el fin de mejorar las predicciones, se aplicará el enfoque Bayesiano. Para ello, se utilizará el a priori de Litterman Minnesota. El paquete Eviews entrega valores por defecto en los hiperparámetros ($\lambda_1 = 0{,}1; \lambda_2 = 0{,}99; \lambda_3 = 1$)

Tras realizar varias pruebas con el paquete Eviews, el mejor modelo es el obtenido al estimar la matriz de covarianza con un modelo VAR Diagonal y tomando $\lambda_1 = 0{,}1$.

Estimación de la matriz de covarianzas	λ_1	Estimación del Error Cuadrático Medio			
		Y_{1t}	Y_{2t}	Y_{3t}	Y_{4t}
AR univariante	0.1	0.03112	0.03679	0.06474	0.02385
	0.5	0.02796	0.03412	0.06360	0.02365
	2	0.02323	0.02956	0.05385	0.01947
VAR diagonal	0.1	0.02006	0.02662	0.04929	0.01693
	0.5	0.02747	0.03439	0.06313	0.02370
	2	0.0283	0.03693	0.06228	0.02456
Modelo VAR	0.1	0.027	0.03489	0.06397	0.02342
	0.5	0.02718	0.03432	0.06279	0.02363
	2	0.02762	0.03467	0.06315	0.02377
AR(1)	0.1	0.03009	0.03592	0.06471	0.02365
	0.5	0.02778	0.03409	0.06346	0.02361
	2	0.02843	0.03484	0.06406	0.02442

Tabla 2: Error cuadrático medio para varios modelos

A continuación, presentamos los valores reales, y las predicciones obtenidas por el modelo VAR y el modelo BVAR:

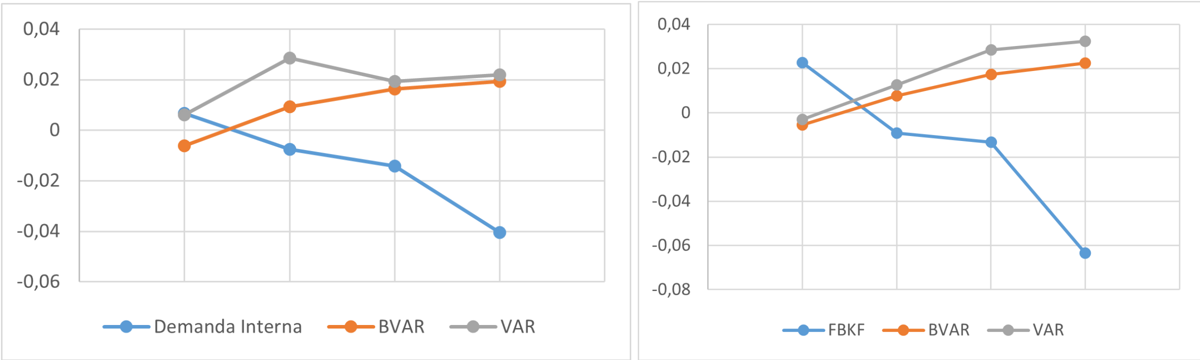

Figura 7: Comparaciones entre el modelo BVAR y el modelo VAR para Y_{1t} (Demanda Interna) y Y_{2t} (FBKF)

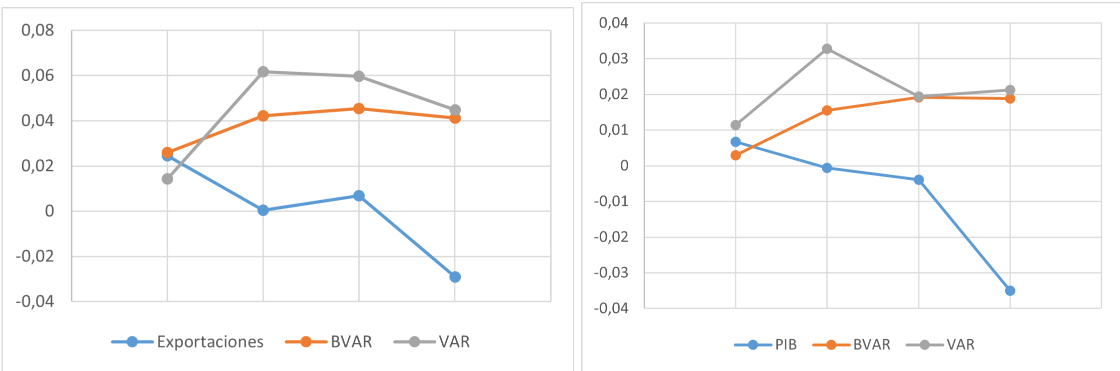

Figura 8: Comparaciones entre el modelo BVAR y el modelo VAR para Y_{3t} (Exportaciones) y Y_{4t} (PIB)

Para la demanda interna podemos observar como el modelo BVAR presenta mejores predicciones que el modelo VAR. Sin embargo, ambas predicciones están alejadas del valor verdadero, el cual pertenece al segundo trimestre del 2020 (periodo en el que se experimentó una crisis económica). Intepretaciones análogas se obtienen para el FBKF, las exportaciones y el PIB.

3. Conclusiones

- El modelo BVAR con el a priori de Litterman Minnesota mejora las predicciones obtenidas con el modelo VAR.
- En los modelos VAR y BVAR al incluir variables con causalidad (X_{1t}, X_{2t}, X_{3t} y X_{4t}) el modelo cuenta con más información para explicar las variables individuales, lo que reduce la incertidumbre y permite mejores pronósticos.
- Podemos notar que los valores reales y las predicciones difieren, esta gran diferencia se debe a los impactos de las medidas y restricciones tomadas por parte de Estado para evitar la propagación del COVID-19. De hecho, hasta mayo del 2020 el país experimentó grandes pérdidas económicas y desempleos masivos.

Referencias

[1] E.L. Lehmann, George Casella *Theory of Point Estimation.* Segunda edición.

[2] ENRIQUE M. QUILIS (2002), *Modelos BVAR: Especificación, estimación e inferencia*, España, pp 9.

[3] HELMUT LÜTKEPOHL. *New Introduction to Multiple Time Series Analysis.*

[4] HOLGER CAPA *Modelización de series temporales.*