UNIVERSIDAD MARISTA

Escuela de actuaría

MODELOS AUTORREGRESIVOS VECTORIALES (VAR)

Estadística III

Autor:
Robles Reyes Erika Regina
Román Crispín Nidya Andrea

Mayo 6 , 2021

1 Modelos de Autorregresión Vectorial

El modelo de autorregresión vectorial (VAR) es un modelo empleado para el análisis de las series de tiempo multivariadas, es decir, cuándo tenemos más de una serie temporal. Se caracteriza por ser uno de los modelos más exitosos, simple, útil y flexible para modelar los rendimientos de los activos, describir el comportamiento dinámico de las series de tiempo económicas y financieras.

Así mismo se utiliza para realizar pronósticos superiores a los de los modelos univariados, los pronósticos realizados por el modelo VAR resultan ser flexibles debido a que se puede ir condicionando las posibles trayectorias futuras de las variables especificadas en el modelo.

El modelo VAR también se utiliza para la inferencia estructural y el análisis de políticas. En el análisis estructural, se imponen ciertos supuestos sobre la estructura causal de los datos investigados, y se resumen los impactos causales resultantes de choques o innovaciones inesperadas en variables específicas sobre las variables del modelo. Estos impactos causales generalmente se resumen con las funciones de respuesta de impulso y la descomposición de la varianza de error prevista.

Utilizamos un modelo del tipo VAR cuando queremos caracterizar las interacciones simultáneas entre un grupo de variable, es decir, cuándo se quiere captar las dependencias dinámicas que puede existir entre las series temporales a estudiar. Un VAR es un modelo de ecuaciones simultáneas formado por un sistema de ecuaciones de forma reducida sin restringir.

1.1 Definición

Una serie de tiempo multivariada r_t es un proceso VAR de orden 1 o VAR (1), si sigue el modelo:

$$r_t = \phi_0 + \Phi\ r_{t-1} + a_t\ ...(1)$$

Dónde:

ϕ_0 es un vector k-dimensional
Φ es una matriz de k x k
a_t es una secuencia de vectores aleatorios no seriados, no correlacionados, con media de cero y la matriz de covarianza $\sum$.

Se requiere que la matriz de covarianza $\sum$ sea definida positiva, de lo contrario r_t se puede reducir. Por otra parte a_t a menudo se asume que es normal multivariada.

Considerando el caso en que el VAR sea bivariado, es decir:

k = 2

$r_t = (r_{1t},\ \ r_{2t})'$
$a_t = (a_{1t},\ a_{2t})'$

Tendremos que el modelo VAR(1) será:

$$r_{1t} = \phi_{10} + \Phi_{11}\ r_{1,t-1} + \Phi_{12}\ r_{2,t-1} + a_{1t}\ ...(2)$$
$$r_{2t} = \phi_{20} + \Phi_{21}\ r_{1,t-1} + \Phi_{22}\ r_{2,t-1} + a_{2t}\ ...(3)$$

Dónde:

ϕ_{ij} es el elemento (i,j) de Φ
ϕ_{i0} es el i-ésimo elemento t de Φ_0

Dentro de la ecuación (2) notamos que Φ_{12} denota dependecia lineal de r_{1t} en $r_{2,t-1}$ en presencia de $r_{1,t-1}$.

Φ_{12} está condicionado de $r_{2,t-1}$ en r_{1t} dado por $r_{1,t-1}$

Sin embargo, si $\Phi_{12} = 0$ entonces r_{1t} no dependendería de $r_{2,t-1}$ y el modelo mostraría que r_{1t} sólo depende de su propio pasado, es decir no es dependiente de ningún r_t anterior. Para $\Phi_{21} = 0$ es similar.

Considerando ambas ecuaciones en conjunto tenemos que:

1. Si $\Phi_{12} = 0$ y $\Phi_{21} \neq 0 \Rightarrow$ hay una relación unidireccional de r_{1t} a r_{2t}
2. Si $\Phi_{12} = \Phi_{21} = 0 \Rightarrow r_{1t}$ y r_{2t} están desvinculados.
3. Si $\Phi_{12} \neq 0$ y $\Phi_{21} \neq 0 \Rightarrow$ existe una relación entre ambas series.

1.2 Formas Reducidas y Estructurales

Que sean ecuaciones de forma reducida quiere decir que los valores contemporáneos de las variables del modelo no aparecen como variables explicativas en ninguna de las ecuaciones. Por el contrario, el conjunto de variables explicativas de cada ecuación está constituido por un bloque de retardos de cada una de las variables del modelo.
Que sean ecuaciones no restringidas significa que aparece en cada una de ellas el mismo grupo de variables explicativas.

El modelo estructural puede incorporar asimismo un vector de variables explicativas exógenas z_t en cada ecuación, que pueden aparecer asimismo con retardos. Un ejemplo de este tipo de variables serían una tendencia determinista, o variables ficticias estacionales. También podrían ser variables que se determinan claramente fuera de la influencia de r_{1t} y r_{2t}, de modo que pueda justificarse que $E(Z_{t-s}a_{1t}) = E(Z_{t-s}a_{2t}) = 0\ \forall s$

De forma resumida, la representación matricial del modelo estructural de primer orden puede escribirse:

$$Br_t = \Gamma_0 + \Gamma_1 r_{t-1} + Gz_t + a_t\ ...(4)$$

El modelo estructural VAR presenta dos dificultades para su estimación:

a) la simultaneidad, al aparecer cada una de las dos variables como variable explicativa en la ecuación de la otra, lo que genera inconsistencia del estimador Minimos Cuadrados Ordinarios (MCO)

b) si los términos de error tuviesen autocorrelación, las estimaciones MCO serían inconsistentes, al tratarse de un modelo dinámicos.

La primera dificultad podría resolverse estimando por variables instrumentales, siempre que contemos con instrumentos adecuados, lo cual no es sencillo de justificar. Por eso, para evitarla, transformamos el modelo. Pero la segunda dificultad persistir, y se debe resolver tratando de ampliar la estructura dinámica del modelo hasta lograr que los términos de error carezcan de autocorrelación.

Supongamos que la matriz B tiene inversa, lo cual requiere que det(B) $\neq$ 0, tendríamos entonces:

$$B^{-1}r_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 r_{t-1} + B^{-1}Gz_t + B^{-1}a_t = \phi_0^* + \Phi^* r_{t-1} + Mz_t + u_t \quad ...(5)$$

Dónde:
$B^{-1} = (w_{k1}, w_{k2}, ..., w_{k,k-1}, 1)$
$\phi_0^* = B^{-1}\Gamma_0$ es un vector k-dimensional
$\Phi^* = B^{-1}\Gamma_1$ es una matriz de k x k
$M = B^{-1}G$ es el vector de las variables exógenas
$u_t = (u_{1t}, ..., u_{kt}) = B^{-1}a_t$

Tendremos la k-ésima ecuación del modelo (5) :

$$r_{kt} + \sum_{i=1}^{k-1} w_{ki} r_{it} = \phi_{k,0}^* + \sum_{i=1}^{k} \phi_{ki}^* r_{i,t-1} + b_{kt} \quad ...(6)$$

Dónde:
$\phi_{k,0}^*$ es el k-elemento de Φ_0^*
ϕ_{ki}^* es el (k,i) elemento de Φ^*

De este modo se pasaría la forma reducida.

1.3 Condición Estacionaria y Momentos de un Modelo VAR(1)

Asumiendo que el modelo VAR(1) definido en la ecuación (1) sea débilmente estacionario. Utilizando $E(a_t) = 0$, obtenemos que:

$$E(r_t) = \phi_0 + \Phi E(r_{t-1}) \quad ...(7)$$

Desde que $E(r_t)$ es invariante del tiempo, tenemos:

$$\mu \equiv E(r_t) = \phi_0 (I - \Phi)^- 1 \quad ...(8)$$

siempre que la matriz I - Φ es no-singular, en dónde I es una matriz de identidad k x k.
Usando $\phi_0 = (I - \Phi)\mu$ para reescribir la ecuación (1) quedaría como:

$$(r_t - \mu) = \Phi(r_{t-1} - \mu) + a_t \text{ ...(9)}$$

Permitiendo que $\widetilde{r_t} = r_t - \mu$ ser la media corregida de la serie de tiempo. A continuación, el modelo VAR(1) se convierte en

$$\widetilde{r_t} = \Phi(\widetilde{r}_{t-1}) + a_t \text{ ...(10)}$$

Este modelo puede ser usado para derivar propiedades de un modelo VAR(1).
Por sustituciones repetidas, podemos reescribir la Ecuación (10) como:

$$\widetilde{r_t} = a_t + \Phi a_{t-1} + \Phi^2 a_{t-2} + \cdots \text{ ...(11)}$$

Esta expresión muestra varias características de un proceso VAR(1).
En primer lugar, puesto que a_t es serialmente incorrupto, se sigue que $Cov(a_t, r_{t-1}) = 0$. De hecho, a_t no está correlacionado con $r_{t-\ell}$ para todos $\ell > 0$. Por esta razón, a_t se conoce como el choque o la innovación de la serie en el tiempo t.
Segundo, postergando la expresión por una a_t' , tomando expectativa, y usando el hecho de que no hay correlaciones seriales en el proceso a_t, obtenemos $Cov(r_t, a_t) = \sum$.
En tercer lugar, para un modelo VAR(1), r_t depende de la innovación pasada en a_{t-j} con matriz de coeficiente Φ^j .
Para que tal dependencia sea significativa,Φ^j debe converger a cero como j $\rightarrow \infty$. Esto significa que los valores propios de k de Φ deben ser inferiores a 1 en el módulo; de lo contrario, Φ^j explotará o convergerá a una matriz distinta de cero como j$\rightarrow \infty$.

1.4 Modelos VAR de orden superior

Los modelos anteriores se dicen que son de orden 1 porque en ellos las variables explicativas aparecen únicamente con un retardo. En general, un modelo VAR de orden n , con variables endógenas, se especifica:

$$R_t = \Gamma_0 + \sum_{i=1}^{n} \Gamma_i R_{t-i} + GZ_t + u_t \text{ ...(12)}$$

Dónde:
R_t es un vector de k x 1
n es el orden del modelo VAR o número de retardos de cada variable en cada ecuación
u_t es un vector de k x 1 de innovaciones, es decir, procesos sin autocorrelación, con $Var(u_t) = \sum$, constante
Z_t es un vector de variables exógenas

El elemento (i; j) en la matriz $A_s; 1 \leq s \leq n$ mide el efecto directo o parcial de un cambio unitario en Y_j en el instante t sobre los valores numéricos de Y_i al cabo de s períodos, es decir, sobre el vector $Y_{i;t+s}$. La columna j de la matriz A_s mide el efecto que un cambio unitario en Y_j en el instante t tiene sobre el vector Y_{t+s}. El elemento i-ésimo en u_t es el componente de Y_{it} que no puede ser previsto utilizando el pasado de las variables que integran el vector.

Como puede verse, en un modelo VAR todas las variables son tratadas simétricamente, siendo explicadas por el pasado de todas ellas. El modelo tiene tantas ecuaciones como variables, y los valores retardados de todas las ecuaciones aparecen como variables explicativas en todas las ecuaciones.
Una vez estimado el modelo, puede procederse a excluir algunas variables explicativas, en función de su significación estadística, pero hay razones para no hacerlo. Por un lado, si se mantiene el mismo conjunto de variables explicativas en todas las ecuaciones, entonces la estimación por mínimos cuadrados ordinarios ecuación por ecuación es eficiente, por lo que el proceso de estimación del modelo es verdaderamente sencillo. Por otro, la presencia de bloques de retardos como variables explicativas hace que la colinealidad entre variables explicativas sea importante, lo que hace perder precisión en la estimación del modelo y reduce los valores numéricos de los estadísticos tipo t de Student. Por tanto, no es buena estrategia proceder en varias etapas, excluyendo del modelo las variables cuyos coeficientes resultan estadísticamente no significativos, por cuanto que esto puede ser consecuencia de la colinealidad inherente al modelo, y no tanto de la falta de contenido informativo de las variables.

1.5 Función de respuesta a impulsos

Un modelo $VAR(p)$ puede ser escrito como una función lineal de los eventos pasados, por lo que se denota de la forma,

$$r_t = \mu + a_t + \Psi a_{t-1} + \Psi a_{t-2} + ..., \qquad (1)$$

donde las matrices (n x n) de las medias móviles $\mu = [\Phi(1)]^{-1}\phi_0$ siempre y cuando exista la inversa y las matríces de coeficientes Ψ_i se puedan obtener igualando los coeficientes B_i en la ecuación

$$(I - \Phi_1 B - ... - \Phi_p B^p)(I + \Psi_1 B + \Psi_2 B^2 + ...) = I$$

donde I es la matriz de identidad. De hecho, se trata de una prepresentación de la media móvil de r_t con los coeficientes de la matríz Ψ_i siendo el impacto de la innovación pasada a_{t-i} en r_t. De igual forma, i es el efecto de a_t sobre la observación futura de r_{t+i}.

Sin embargo, dado que los componentes de a_t normalmente están correlacionados, la interpretación de los elementos de la función lineal de las innivaciones pasadas no es sencilla, por lo que se utiliza la descomposición de Cholesky, que transforma las innovaciones de modo que los componentes resultantes no se encuentren correlacionados. Existe una matríz triangular inferior L tal que $L = LGL$, donde G es la matríz diagonal y los elementos diagonales de L son unidad.

Sea $b_t = L^{-1}a_t$, entonces, $Cov(b_t) = G$ de modo que los elementos de b_t no están correlacionados.Se reescribe la ecuación como,

$$r_t = \mu + a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + ...$$

$$= \mu + LL^{-1}a_t + \Psi_1 LL^{-1}a_{t-1} + \Psi_2 LL^{-1}a_{t-2} + ...$$

$$= \mu + \Psi_0^* b_t + \Psi_1^* b_{t-1} + \Psi_2^* b_{t-2} + ..., \tag{2}$$

donde $\Psi_0^* = L$ y $\Psi_i^* = \Psi_i L$. Los coeficientes de las matrices Ψ_i^* son llamadas *función de respuesta a impulsos* de r_t con innovaciones otogonales b_t. Especificamente, el (i,j) elemento de Ψ_ι^*, es decir, $\Psi_{ij}^*(\iota)$, es el impacto de $b_{i,j}$ en las observaciones futuras $r_{i,t+\iota}$.

2 Modelo de promedio móvil de vectores

Un modelo de media móvil vectorial de orden q o $VMA(q)$, es de la forma

$$r_t = \theta_0 + a_t - \Theta_1 a_{t-a} - ... - \Theta_q a_{t-q} \quad ó \quad r_t = \theta_0 + \Theta(B)a_t \tag{3}$$

donde θ_0 es un vector k-dimensional, Θ_i son matrices k x k y $\Theta(B) = I - \Theta_1 B - ... - \Theta_q B^q$ es la matríz polinomio MA en la parte posterios de cambio de operador B.

Los procesos $VMA(q)$ son débilmente estacionarios siempre que la matríz de covarianza Σ de a_t existe. Tomando la expectativa de la ecuación descrita con anterioridad, se obtiene que $\mu = E(r_t) = \theta_0$, entonces el vector constante θ_0 es el vector medio de r_t del modelo VMA.

Sea $\bar{r_t} = r_t - \theta_0$ la media corregida del proceso $VAR(q)$. Cuando la ecuación anterior y suponiendo que a_t no tiene correlaciones, se tiene que

1. $Cov(r_t, a_t) = \Sigma$
2. $\Gamma_0 = \Sigma + \Theta_1 \Sigma \Theta_1^{'} + ... + \Theta_q \Sigma \Theta_q^{'}$
3. $\Sigma_\iota = 0 \quad \text{si} \quad \iota > q$
4. $\Sigma_\iota = \sum_{j=\iota}^{q} \Theta_j \Sigma \Theta_{j-\iota}^{'} \quad \text{si} \quad 1 \leq \iota \leq q, \quad \text{donde} \quad \Theta_0 = -I$

Como $\Gamma_\iota = 0$ para $\iota > q$, la matríz de correlaciones cruzadas (CCMs) de un proceso $VMA(q)$, r_t satisface

$$\rho_\iota = 0 \quad \iota > q \tag{4}$$

Al igual que en el caso univariado, las CCMs pueden usarse para identificar el orden de un proceso VMA.

Para hacer más entendible un modelo VMA, considere el modelo bivariado MA(1)

$$r_t = \theta_0 + a_t - \Theta a_{t-1} = \mu + a_t - \Theta a_{t-1} \tag{5}$$

por simplicidad Θ_1 se elimina, de forma que el modelo puede escribirse explicitamente como

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix} \tag{6}$$

Dice que la serie de rendimientos actuales r_t solo depende de los eventos actuales y pasados. Por lo que es posible decir que es un modelo de memoria finita.

Estimación

La estimación para los modelos VMA es compleja, de hecho existen dos métodos disponibles para el enfoque de probabilidad. El primero es el método de verosimilitud condicional que asume que $a_t = 0$ para $t \leq 0$. El segundo es el método de verosimilitud exacta que trata a_t con $t \leq 0$.

Tomando el ejemplo de un modelo VMA(1), suponga que los datos son $r_t | t = 1, ..., T$ y a_t es multivariable normal.

MLE condicional

Dicho modelo asume que $a_0 = 0$, bajo el supuesto y reescribiendo el modelo como $a_t = r_t - \theta_0 + \Theta a_{t-1}$, se calcula a_t como

$$a_1 = r_1 - \theta_0, \quad a_2 = r_2 - \theta_0 + \Theta_1 a_1, \quad ...$$

En consecuencia, la función de verosimilitud se ve de la forma

$$f(r_1, ..., r_T | \theta_0, \Theta_1, \Sigma) = \prod_{t=1}^{T} \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} exp(-\frac{1}{2} a_t^{'} \Sigma^{-1} a_t)$$

que puede evaluarse para obtener las estimaciones de los parámetros.

MLE exacto

Para el método exacto, a_0 es un vector desconocido que debe ser estimado a partir de los datos para evaluar la función de verosimilitud. Sea, $\bar{r}_t = r_t - \theta_0$ la serie media corregida, se tiene

$$a_t = \bar{r_t} + \Theta a_{t-1} \tag{7}$$

De hecho, a_0 se relaciona con toda $\bar{r_t}$ como

$$a_1 = \bar{r_1} + \Theta a_0$$

$$a_2 = \bar{r_2} + \Theta a_1 = \bar{r_2} + \Theta \bar{r_1} + \Theta^2 a_0$$

$$\dots$$

$$a_T = \bar{r_T} + \Theta \bar{r_T} + ... + \Theta^{T-1} \bar{r_1} + \Theta^T a_0 \tag{8}$$

Por lo tanto, a_0 es una función lineal de los datos si θ_0 y Θ son dadas. Los resultados permiten estimar a_0 usando los datos y estimaciones iniciales de θ_0 y Θ, con dichos datos se puede reescribir la siguiente ecuación,

$$r_t^* = \bar{r}_t + \Theta \bar{r}_{t-1} + \cdots + \Theta^{t-1} \bar{r}_1, \quad con \quad t = 1, 2, \dots, T$$

Entonces, se escribe la ecuación anterior como,

$$r_1^* = -\Theta a_0 + a_1$$

$$r_2^* = -\Theta^2 a_0 + a_2$$

$$\dots$$

$$r_T^* = -\Theta^T a_0 + a_T$$

Esta ecuación tiene la forma de una regresión lineal multiple con parámetro vector a_0, incluso aun que la matríz de covarianza Σ de a_t puede no ser una matríz diagonal.

3 Modelos de vector ARMA

Los modelos ARMA univariados pueden ser generalizados para series de tiempo vectoriales y se escriben como VARMA. Sin embargo, la generalización trae el problema de la *identificabilidad*, es decir, dichos modelos no se defnen de una única forma. Por ejemplo, un modelo VMA(1)

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

es identico a un modelo VAR(1)

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Otro tipo de problema de identificabilidad vectorial puede ser de la sigueinte manera. Considere un modelo VARMA(1,1)

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} -0.5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

Este modelo es identico a un modelo VARMA(1,1)

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & -2+\eta \\ 0 & \omega \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} -0.5 & \eta \\ 0 & \omega \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

Los efectos de los parámetros ω y η en el sistema se cancela entre las partes AR y MA del segundo modelo. Pero el problema de identificabilidad es serio porque, sin restricciones adecuadas, la función de probabilidad de un modelo ARMA(1,1) vectorial no se define de forma única, lo que resulta una situación similar a la multicolinearidad (situación en la que se presenta una fuerte correlación entre variables explicativas del modelo).

Cuando se utilizan modelos VARMA, solo se entrenan modelos de orden inferior (por ejemplo, un modelo VARMA(1,1) o VARMA(2,1)) especialmente cuando las series de tiempo involucradas no son estacionales.

Entonces, un modelo $VARMA(p,q)$ se escribe como

$$\Phi(B)r_t = \phi_0 + \Theta(B)a_t \tag{9}$$

donde $\Phi(B) = I - \Phi_1 B - ... - \Phi_p B^p$ y $\Theta(B) = I - \Theta_1 B - ... - \Theta_q B^q$ son dos matrices polinomiales k x k. Una condición necesaria y suficiente de estacionalidad débil para r_t es la misma para un modelo $VAR(p)$ con matriz polinomial $\Phi(B)$.

Para $v > 0$, los (i, j) elementos de los coeficiente de la matríz Φ_v y Θ_v miden la dependencia lineal de r_{1t} en $r_{j,t-v}$ y $a_{j,t-v}$. Si el elemento (i, j) es cero para todas las matrices de coeficiente AR y MA, r_{it} no depende de los valores retrasados de r_{jt}.

Considerar el modelo bivariado,

$$\begin{bmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{bmatrix} \begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \begin{bmatrix} \Theta_{11}(B) & \Theta_{12}(B) \\ \Theta_{21}(B) & \Theta_{22}(B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Las condiciones necesarias y suficientes para la existencia de una relación dinámica de r_{1t} a r_{2t} son

$$\Phi_{22}(B)\Theta_{12}(B) - \Phi_{12}(B)\Theta_{22}(B) = 0$$

pero

$$\Phi_{11}(B)\Theta_{21}(B) - \Phi_{21}(B)\Theta_{11}(B) \neq 0 \tag{10}$$

Estás condiciones se pueden obtener dejando,

$$\Omega(B) = |\Phi(B)| = \Phi_{11}(B)\Phi_{22}(B) - \Phi_{12}(B)\Phi_{21}(B)$$

El determinante de la matríz polinomial AR y premultiplicar el modelo por la matríz,

$$\begin{bmatrix} \Phi_{22}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{11}(B) \end{bmatrix}$$

puede escribirse el modelo bivariado como

$$\Omega(B) \begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \begin{bmatrix} \Phi_{22}(B)\Theta_{11}(B) - \Phi_{12}(B)\Theta_{21}(B)\Phi_{22}(B)\Theta_{12}(B) - \Phi_{12}(B)\Theta_{22}(B) \\ \Phi_{11}(B)\Theta_{21}(B) - \Phi_{21}(B)\Theta_{11}(B)\Phi_{11}(B)\Theta_{22}(B) - \Phi_{21}(B)\Theta_{12}(B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

La estimación del modelo VARMA puede llevarse a cabo mediante el método condicional o método exacto de máxima verosimilitud.