



Munich Personal RePEc Archive

Standard methodology of vector autoregression (VAR) and error correction (VEC)

John Michael, Riveros-Gavilanes

Veeduría Estudios y Evaluación de la Gestión Pública Colombiana
-EEGPC-

9 March 2025

Online at <https://mpra.ub.uni-muenchen.de/124015/>
MPRA Paper No. 124015, posted 19 Mar 2025 16:53 UTC

Metodología estándar de vectores autoregresivos (VAR) y de corrección del error (VEC)

John Michael Riveros-Gavilanes¹

Master of Science Economics (M. Sc)

Ludwig-Maximilians-Universität München, Deutschland

Volkswirtschaftliche Fakultät

Resumen: Este documento presenta una introducción práctica a la metodología estándar de la estimación de vectores auto-regresivos (VAR) y su aproximación de vectores con corrección del error (VEC) en el contexto de la cointegración. El documento presenta unas nociones básicas sobre el concepto de estacionariedad y raíz unitaria, la estimación de pruebas de raíces unitarias, la revisión de cointegración y el esquema general de estimación bajo el programa Stata. El texto no ahonda con la profundización matemática de los modelos sino más que nada aspira a ser una guía aplicada de estimación para los estudiantes de pregrado.

Palabras Clave: Vectores auto-regresivos, cointegración, estacionariedad, series de tiempo.

Clasificación JEL: C10; C13; C32.

Abstract: This document provides a practical introduction to the standard methodology for estimating Vector Autoregression (VAR) models and their Vector Error Correction (VEC) form in the context of cointegration. It covers basic concepts such as stationarity and unit roots, unit root testing, cointegration analysis, and the general estimation framework using Stata. The text does not delve into the mathematical formalization of the models but rather aims to serve as an applied estimation guide for undergraduate students.

Keywords: Vector autoregression, cointegration, stationarity, time series.

9.03.2025

¹ Filiación Institucional: Veeduría Estudios y Evaluación de la Gestión Pública Colombiana - EEGPC, Corporación Centro de Interés de Público y Justicia – CIPJUS. M&S Research Hub. Enlaces: <https://scholar.google.com/citations?user=XnwgW88AAAAJ&hl=es> Contacto: empirical.analysis.und.app@outlook.com

Tabla de contenido

1.	Introducción.....	4
2.	Estructura de la metodología estándar para VAR y VEC	5
3.	Nociones previas	7
3.1.	¿Qué es una serie de tiempo?.....	7
3.2.	¿Qué es una serie estacionaria?.....	8
3.3.	¿Tendencia y no estacionariedad?	10
3.4.	Primeras diferencias.....	11
3.5.	Raíces unitarias y no estacionariedad.....	12
3.6.	Orden de integración	15
3.7.	¿Qué es la cointegración?.....	16
3.8.	Ecuación de cointegración a corrección del error	17
4.	Aplicación: Estimación VAR.....	20
I.	Análisis grafico	21
II.	Análisis de raíces unitarias y ordenes de integración.....	22
III.	Selección de mejor modelo	23
V.	Estimación VAR.....	24
VI.	Estimacion modelo VEC (omitido por no corresponder).....	26
VII.	Verificación de supuestos econométricos.....	27
	No autocorrelación	27
	Normalidad	27
	Estabilidad de parámetros.....	28
VIII.	Interpretación.	28
IX.	Función de impulso respuesta	29
X.	Conclusiones.....	32
5.	Aplicación: Estimación VEC.....	34
I.	Análisis gráfico de las series temporales.....	36
II.	Análisis de raíces unitarias y definición del orden de integración	37
III.	Selección de modelos con base en la definición de ordenes de integración.....	41
IV.	Implementación de pruebas de cointegración (solo para series I(1) únicamente).42	
VI.	Estimación del modelo VEC si las series son I(1) y cointegradas	44
VII.	Verificación de supuestos esenciales (no autocorrelación, estabilidad, normalidad)	
	49	
VIII.	Interpretación de resultados.....	50
IX.	Análisis de impulso respuesta	51
X.	Conclusión.....	53
6.	Referencias.....	55

1. Introducción

Este texto se encuentra enfocado para los alumnos del pregrado (Bachelor) en ciencias económicas, concebido como una guía práctica para ejecutar la metodología de series de tiempo asociada a los Vectores Auto-Regresivos (Vector Auto-Regressive - VAR), y también la aproximación de la vertiente de estos con cointegración, conocido como Vectores con Corrección del Error (Vector Error Correction - VEC).

Esta metodología es una oportunidad para que los alumnos puedan ejecutar sus análisis e investigaciones econométricas en el contexto de series de tiempo y multiplicidad de variables con cierta naturaleza “endógena” en el sentido económico. Este término a diferencia de su concepto econométrico (donde existe correlación entre el error y los regresores), se utiliza en este contexto para modelos en los cuales existen impactos e influencias entre las variables de análisis que se encuentran determinadas dentro del mismo sistema.

En este aspecto la definición de una “metodología estándar” para la estimación de los modelos VAR y VEC, tiene como referencia los postulados teóricos de Sims (1980), Sims (1986), Hamilton (1994), Lütkepohl (2005), y Stock & Watson (2020). La metodología estándar tiene una serie de puntos comunes que implican el análisis de las series de tiempo por variable. Estos análisis involucran la graficación o visualización de las series, identificación de posibles componentes estacionales (y su posible eliminación), el análisis de raíces unitarias, la determinación de los ordenes de integración, y en el caso particular de variables que sean integradas de primer orden $I(1)$, la exploración de relaciones de cointegración.

Particularmente como lo expresa Lütkepohl (2005), la cointegración es una revolución a las series de tiempo, donde se analiza las relaciones de equilibrio de largo plazo entre las variables, también denominadas ecuaciones de cointegración. Esta revolución, permitió comprender mejor el concepto de relaciones no espurias, las cuales, eran el común denominador de las regresiones de series de tiempo. Los autores de dicha revolución son Engle & Granger (1987) los cuales introducen el contexto de relaciones de equilibrio y corrección del error en las combinaciones lineales de series integradas de primer orden.

Bajo este contexto, la presente de la metodología estándar no tiene el objetivo de presentar al detalle la matemática detrás de los temas de los modelos VAR y VEC, sino que, por el contrario, busca proveer a los estudiantes de pregrado de herramientas de análisis econométrico aplicado, expresado en los rasgos generales de los modelos, y un ejemplo de aplicación en el software Stata.

Es necesario destacar también, que esta introducción no ahondara en los modelos estructurales de tipo VAR, dado que estos requieren de un “paso más” en el análisis económico. En este contexto, las relaciones estructurales utilizan la literatura y teoría económica para establecer potenciales relaciones contemporáneas. Dado que este es un

modelo más avanzado, el enfoque de esta introducción se enfocará en las versiones reducidas de los modelos, sus pasos esenciales de estimación, y consideraciones generales.

2. Estructura de la metodología estándar para VAR y VEC

La metodología estándar de los modelos VAR y VEC se puede expresar en una serie de pasos que tienen como propósito analizar la posibilidad, estimación e interpretación de las series temporales y sus relaciones dinámicas. Naturalmente, la teoría económica es nuestro punto de partida para analizar las cuestiones empíricas. Sin embargo, la realidad suele ser compleja y heterogénea, y, por consiguiente, debemos también reconocer que tipo de características empíricas tienen nuestras variables y muestras de análisis.

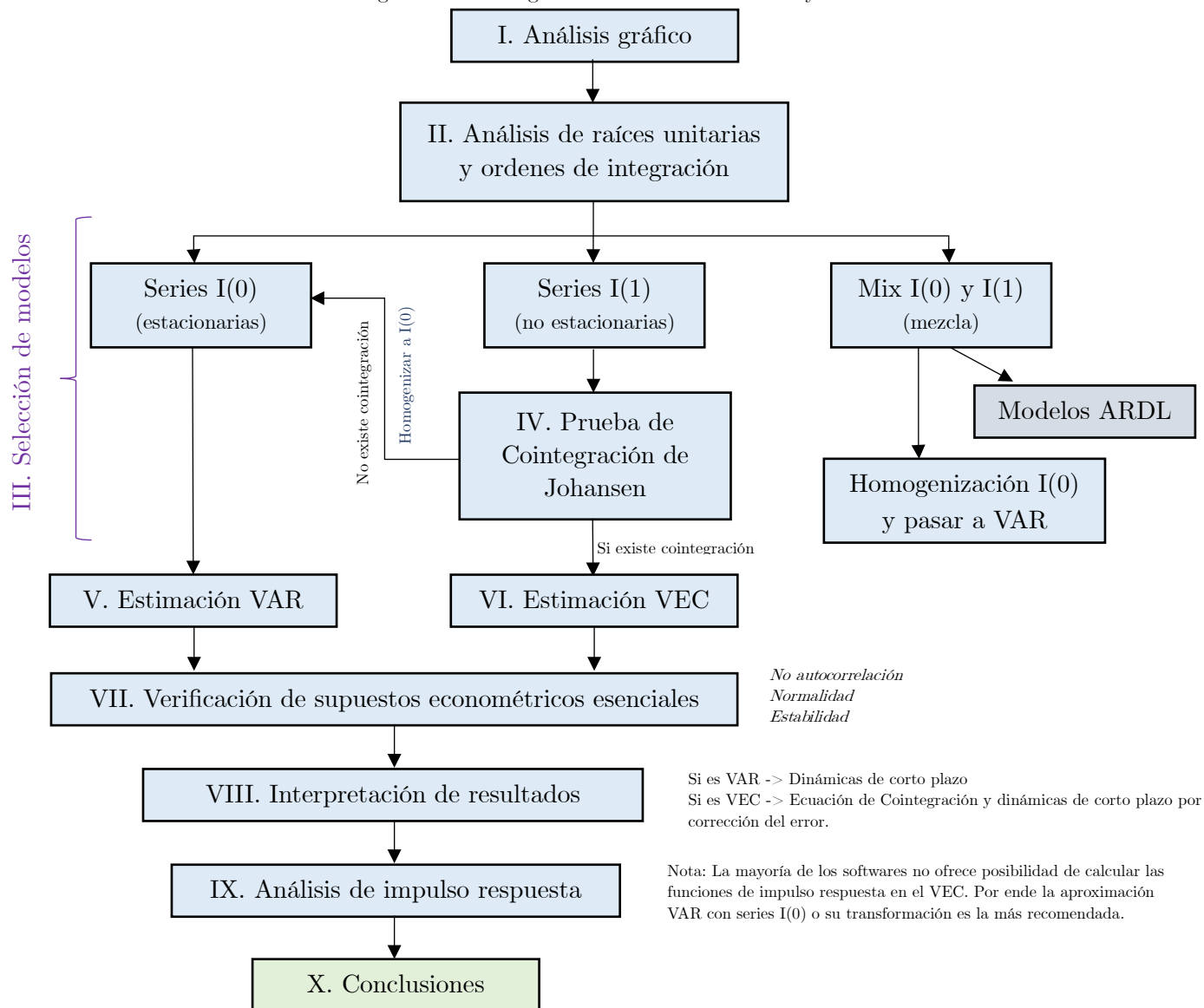
En este contexto, la cointegración entre dos variables, por ejemplo, el “desempleo” y la “inflación” puede no ser la misma para todas las economías del mundo, en todos los periodos de tiempo habidos y por haber, sino porque el contrario, la naturaleza de las series socioeconómicas suelen ser dinámica y cambiante, por lo cual puede variar de economía a economía y de muestra a muestra la posibilidad de encontrar relaciones de largo plazo el cual es el objetivo de la cointegración. Dependiendo de su existencia se procede a estimar un VAR o un VEC. Los pasos esenciales de la metodología estándar para los VAR y VEC son generalmente:

- I. Análisis gráfico de las series temporales
- II. Análisis de raíces unitarias y definición del orden de integración
- III. Selección de modelos con base en la definición de ordenes de integración
- IV. Implementación de pruebas de cointegración (solo para series $I(1)$ únicamente)
- V. Estimación del modelo VAR si las series son estacionarias $I(0)$
- VI. Estimación del modelo VEC si las series son $I(1)$ y cointegradas
- VII. Verificación de supuestos esenciales (no autocorrelación, estabilidad, normalidad)
- VIII. Interpretación de resultados
- IX. Análisis de impulso respuesta
- X. Conclusión

Esta serie de pasos son en realidad una descripción abstracta de los aspectos a considerar en el análisis y estimación de modelos VAR y VEC. En particular, la existencia de series $I(1)$ o $I(0)$ es el aspecto que definirá que tipo de modelo escogemos. Si escogemos un modelo VAR, este esencialmente debe ser estimado para series estacionarias $I(0)$. Por el contrario, si nos encontramos en el caso de que todas las series temporales son $I(1)$ y cointegradas, procederemos a la estimación del modelo VEC. Debido a que el análisis bajo esta serie de

pasos es complejo, la siguiente figura pretende dilucidar un poco el proceso general de la metodología de series de tiempo.

Figura 1 Metodología estándar de modelos VAR y VEC



Fuente: Elaboración propia.

En general en esta metodología estándar se revisarán los aspectos esenciales de cada uno de los pasos descritos en el esquema anterior. Esto debido a que la estimación de los modelos VAR y VEC depende en gran medida de las características del conjunto de series temporales de los cuales dispongamos de información. En este contexto, una serie temporal es una variable x_t de la cual tenemos información no nula en el tiempo (sean días, semanas, meses, trimestres, semestres, u años).

Es siempre importante graficar el conjunto de variables de series de tiempo (las variables x_t) que implementemos en nuestra investigación, esto con el fin de observar si tienen patrones estacionales, de tendencia, y no estacionariedad. En este sentido, en esta introducción de la metodología estándar, se proveerá también una revisión de las nociones básicas de estos conceptos, para comprender sus fundamentos en la identificación de las propiedades de raíces unitarias de cada variable x_t que incluyamos en la investigación.

Las características en particular que nos interesan de x_t para determinar si debemos escoger un modelo VAR o un modelo VEC, vendrá a partir de su orden de integración y la existencia de raíces unitarias, tradicionalmente denotado como $x_t \sim I(d)$ donde x_t es la variable o serie temporal, $\sim$ es un símbolo que hace referencia al orden de integración de orden d . En particular, se puede leer como “ x_t es una variable integrada de orden d ”. El concepto de d hace referencia al número de veces que tenemos que diferenciar x_t para que se vuelva estacionaria. Una variable $I(0)$ es estacionaria en niveles, una variable $I(1)$ es estacionaria en la primera diferencia, mientras una variable $I(2)$ es estacionaria en segunda diferencia², y así sucesivamente.

Debido a que estas nociones pueden no ser familiares, esta introducción tendrá un apartado de nociones previas que introducirá estos temas de forma general. Luego se procederá a la aplicación de modelos VAR y VEC en Stata.

3. Nociones previas

Esta sección tiene como propósito presentar las nociones previas de los conceptos de serie de tiempo, el concepto de estacionariedad, el concepto de tendencia, su visualización, el concepto de orden de integración, y el concepto de primeras diferencias. Finalmente, y de manera general, se revisará el concepto de cointegración.

3.1. ¿Qué es una serie de tiempo?

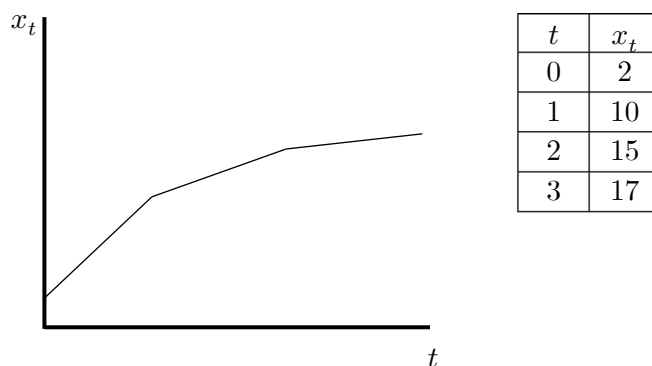
Una serie de tiempo (o variable de serie de tiempo), se puede definir como una serie de valores observados en el tiempo t . Es importante empezar por definir que el tamaño muestral viene definido por una ventana temporal T que es el conjunto total de observaciones disponibles en el tiempo. Por consiguiente, podemos escribir la serie de tiempo x_t de forma matemática como $\{x_t\}_{t \in T}$ lo cual significa que la serie x_t es la variable de tiempo observada de análisis, t es un periodo dentro de la ventana de tiempo T disponible. En general el subíndice t establece un periodo específico de análisis debido a que t puede referenciar

² Estas variables $I(2)$, particularmente aparecen cuando tenemos series macroeconómicas nominales o series afectadas fuertemente por componentes estacionales. En este caso, se debe hacer el proceso de “des-estacionalización” para eliminar el componente estacional de la serie.

diferentes periodos de tiempo. En este contexto $t = 1, 2, 3, \dots, T$ siendo T la ventana de tiempo máxima que disponemos para observar la serie x_t .

La serie de tiempo x_t por consiguiente es un proceso realizado que es observable en diferentes periodos de tiempo, otra forma de expresarlo es como $x(t)$ que tiene el mismo significado. Sin embargo, muchas veces deseamos observar siempre la evolución de la variable x_t en función de sus puntos temporales observables. Es por esta razón, que al graficar la variable x usualmente se hace en función del tiempo t .

Figura 2 Ejemplo de serie de tiempo



Fuente: Elaboración propia.

En este aspecto, la graficación es un instrumento útil para comprender las dinámicas de la serie de tiempo x_t en diferentes puntos temporales. Podemos observar cambios de pendiente, incrementos, y también desaceleración. Una ventaja adicional es que, dependiendo de la unidad de medida, se pueden graficar varias series de tiempo en un mismo gráfico, comparar su evolución conjunta y observar si existe tendencia.

Las series de tiempo vienen definidas por una serie de componentes entre los que se destacan existen tendencia, estacionalidad, ciclo y ruido. La descomposición de estos componentes se puede hacer por métodos aditivos o multiplicativos. Sin embargo, en este apartado no ahondaremos en la descomposición, sino simplemente se presenta el concepto de serie de tiempo. Es importante también no confundir el concepto de estacionalidad con estacionariedad, debido a que el primero, es decir, estacionalidad, hace parte de fenómenos estacionales, que corresponden a comportamientos.

3.2. ¿Qué es una serie estacionaria?

Una serie de tiempo estacionaria (débil) es una serie definida por tres características importantes. Para empezar, la realización de la serie de tiempo x_t proviene de una variable aleatoria X_t esto significa simplemente que lo observado y registrado es x_t mientras que la variable origen verdadera (teórica) es X_t . En este contexto, las series estacionarias son:

1. Media constante de la forma $E[X_t] = \mu$ lo que quiere decir que el valor esperado de X_t es igual a una constante denominada μ . En otras palabras, quiere decir que la serie de tiene una media numérica (dependiente del rezago) y que no tiende a cambiar drásticamente durante la ventana de tiempo.
2. Varianza constante de la forma $Var(X_t) = \sigma^2$ donde la varianza de X_t es igual a una constante positiva definida como σ , también se interpreta que esta varianza, aunque puede depender del rezago, no debe fluctuar significativamente en la ventana temporal. La varianza es una medida de variabilidad estadística, y en este contexto significa que las fluctuaciones (independientemente del tiempo) tienden a tener un comportamiento constante.
3. La dependencia entre diferentes valores de la serie de tiempo no depende del tiempo t sino del rezago de análisis. Esta propiedad es compleja de comprender, y se define formalmente como $Cov(X_t, X_{t-h}) = \gamma(h)$ que significa que la covarianza entre X_t y un rezago h de la misma variable definida como X_{t-h} no se encuentra en función del tiempo t sino de la distancia relativa al rezago. Es decir, la relación entre X_t y el rezago X_{t-h} no es una función del tiempo t sino de la distancia entre rezagos h . En términos abstractos, quiere decir que la estructura de autocorrelación (dependencia temporal) no depende del tiempo, sino que depende de la cercanía entre los rezagos dada su distancia en el tiempo.

Aquí es donde es importante empezar a dimensionar como luce una serie estacionaria, y utilizaremos el software Stata, para simular y visualizar el comportamiento de una serie temporal. El siguiente código debe ser copiado a un script (Do file) para ser ejecutado.

```
clear
set obs 200 // Definir el número de observaciones
gen t = _n // Crear una variable de tiempo
gen e = rnormal(0,1) // Generar un ruido blanco con media 0 y varianza 1
gen x = . // Inicializar la serie de tiempo
replace x = e in 1 // Condición inicial
local phi 0.6 // Parámetro AR(1) (debe ser |phi| < 1 para estacionariedad)

forvalues i = 2/200 {
    replace x = `phi' * x[_n-1] + e in `i'
}

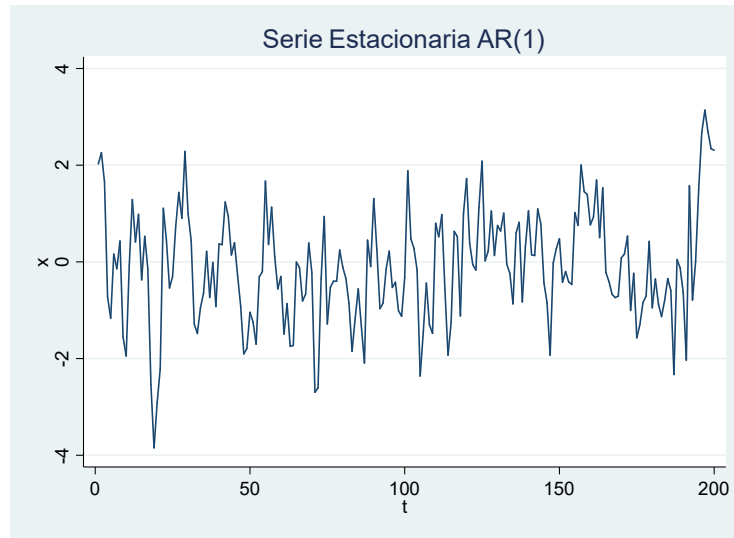
tsset t // Declarar la serie como datos de tiempo
tsline x, title("Serie Estacionaria AR(1)")
```

Al ejecutar el anterior código, se define una ventana T de 200 observaciones, donde generamos la variable índice de tiempo como $t = 1, 2, \dots, 200$. Luego creamos un ruido blanco basado en una distribución normal de media 0 y varianza 1, se inicializa la serie de tiempo con un valor inicial. Y luego, definimos la estructura de la serie temporal x_t a partir de la ecuación:

$$x_t = \phi x_{t-1} + e_t$$

Este se le conoce como proceso autorregresivo, y en el código imponemos una condición especial sobre el coeficiente ϕ tal que, el valor absoluto de phi sea menor que 1 para garantizar el comportamiento estacionario. La condición formalmente se expresa como $|\phi| < 1$. Lo que garantiza “estabilidad” del proceso frente a sus fluctuaciones y media. Nótese que e_t es un componente estocástico para la estructura de x_t .

Figura 3 Proceso estacionario



Fuente: Elaboración propia.

Como se observa, este proceso es estacionario con la interpretación visual dado que sigue una media constante (media de 0), con fluctuaciones que parecen relativamente constantes (desviaciones relativas a la media de 0). Así mismo, no parecen existir componentes tendenciales que influyan el proceso en función del tiempo t .

La definición anterior es el concepto débil de estacionariedad. Sin embargo, existe una definición de estacionariedad fuerte en la cual todas las distribuciones de la serie de tiempo x_t son invariantes. Esto incluye la varianza, autocovarianza, para cualquier subconjunto de la serie en relación con su distribución.

3.3. ¿Tendencia y no estacionariedad?

La tendencia y la no estacionariedad se encuentran interrelacionadas. Cuando existe tendencia, bien sea positiva o negativa o fluctuante totalmente, la serie puede no ser estacionaria (existe un caso particular donde las fluctuaciones se dan en torno a la tendencia sin embargo no suele ser la generalidad). En este contexto, es más fácil comprender un

proceso no estacionario con tendencia visualmente. A partir del siguiente código, podemos simular una serie con tendencia.

```
clear
set obs 50 // Definir 200 observaciones
gen t = _n // Crear una variable de tiempo
gen e = rnormal(0,1) // Generar ruido blanco ~ N(0,1)
gen X = 0 // Inicializar la serie
replace X = 0.02*t + X[_n-1] + e if _n > 1 // Proceso I(1) con tendencia positiva
tsset t // Declarar la serie como datos de tiempo
tsline X, title("Serie I(1) con tendencia positiva") lcolor(blue)
```

Esto da como resultado una serie x_t con tendencia incremental de 0.02 lineal, sujeta a un proceso autorregresivo de la forma $x_t = 0.02t + \phi x_{t-1} + e_t$. El hecho de que $\phi = 1$ inmediatamente causa que las fluctuaciones en torno a la media y la tendencia dejen de ser estacionarias, y su comportamiento sea de naturaleza volátil.

Figura 4 Serie no estacionaria con tendencia positiva



Fuente: Elaboración propia.

Los aspectos clave a considerar en esta serie x_t es que existe una clara tendencia positiva, donde además de todo hay fluctuaciones irregulares ocasionadas por el parámetro $\phi = 1$. Que se traducen en procesos aleatorios donde el componente estocástico e_t también influye en su comportamiento. Esta serie es visiblemente diferente que el de una serie estacionaria, tanto por las rupturas de media, como por la volatilidad, en este contexto, es prudente realizar el análisis de primeras diferencias, el cual se abordará a continuación.

3.4. Primeras diferencias

Las primeras diferencias se definen a partir del cálculo:

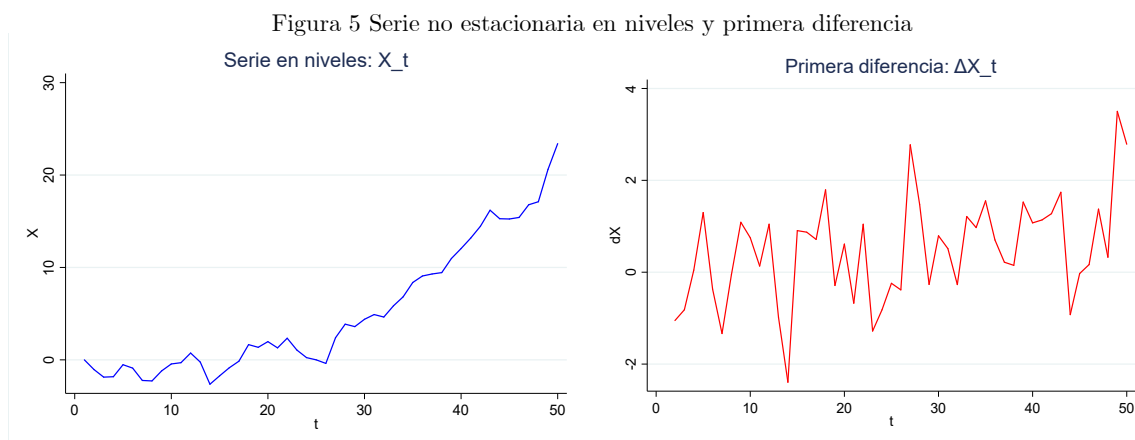
$$\Delta x_t = x_t - x_{t-1}$$

Es decir, que las primeras diferencias se definen como la variación relativa al periodo inmediatamente anterior. Esto tiene serias implicaciones en el comportamiento de la serie. Por ejemplo, retomando el ejemplo de la serie estimada con tendencia, introduzcamos ahora la primera diferencia en el calculo

```
clear
set obs 50 // Definir 200 observaciones
gen t = _n // Crear una variable de tiempo
gen e = rnormal(0,1) // Generar ruido blanco ~ N(0,1)
gen X = 0 // Inicializar la serie
replace X = 0.02*t + X[_n-1] + e if _n > 1 // Proceso I(1) con tendencia positiva
gen dX = X - X[_n-1] // Calcular la primera diferencia  $\Delta X_t$ 
tsset t // Declarar la serie como datos de tiempo

// Graficar en un solo panel
twoway (tsline X, lcolor(blue) title("Serie en niveles: X_t") graphregion(color(white)))
twoway (tsline dX, lcolor(red) title("Primera diferencia:  $\Delta X_t$ ") graphregion(color(white)))
```

Lo cual deriva en las siguientes graficas:



Fuente: Elaboración propia.

Como se observa, el comportamiento de la serie x_t cambia drásticamente en comparación que su primera diferencia $\Delta x_t = x_t - x_{t-1}$. Y parece tener un comportamiento más estacionario debido a que parece fluctuar sobre una media común y sus desviaciones parecen ser relativamente constantes. Sin embargo, dado que este es una interpretación subjetiva y visual, es necesario elaborar pruebas formales para comprobar si la serie es estacionaria o no. Cuestión que se abordará en el siguiente apartado con el mismo código generado.

3.5. Raíces unitarias y no estacionariedad

Cuando una serie tiene raíces unitarias, estamos hablando de un proceso generador de información de la forma:

$$x_t = \phi x_{t-1} + e_t \quad \text{donde} \quad \phi = 1$$

Esto automáticamente convierte a x_t en un proceso que viola una media común, una varianza constante, y si tuviese tendencia, la fluctuación de x_t también estaría en función del tiempo. En este contexto, con la serie debemos proceder además del análisis gráfico, con pruebas de raíces unitarias, para determinar el orden de integración de x_t . Cuando hablamos de orden de integración, nos referimos al número de diferencias que debemos realizar para que x_t se vuelva estacionaria.

Empezaremos con la simulación de la serie x_t que posee raíz unitaria y que graficamos en el anterior apartado.

```
clear
set seed 123 // semilla de replicación
set obs 50 // Definir 200 observaciones
gen t = _n // Crear una variable de tiempo
gen e = rnormal(0,1) // Generar ruido blanco ~ N(0,1)
gen X = 0 // Inicializar la serie
replace X = 0.02*t + X[_n-1] + e if _n > 1 // Proceso I(1) con tendencia positiva
tsset t // Declarar la serie como datos de tiempo
```

E introduciremos una selección de rezagos óptimos para escoger dentro de la prueba Aumentada de Dickey-Fuller los rezagos ideales de prueba de raíces unitarias sobre x_t . Por consiguiente, la línea de comandos a utilizar después del anterior código es:

```
varsoc X
dfuller X, lag(1)
```

Lo cual da como resultado una tabla con diferentes criterios de selección del error incluyendo: LL y LR, Final Prediction Error -FPE, Akaike Information Criterio -AIC, Hannan Quinn Information Criterio -HQIC, Bayesiano -SBIC.

Selection-order criteria

Sample: 5 - 50 Number of obs = 46

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-159.006				61.4922	6.95678	6.97167	6.99653
1	-59.9731	198.07*	1	0.000	.866459	2.69448	2.72426*	2.77399*
2	-58.8696	2.2069	1	0.137	.862683*	2.68998*	2.73466	2.80924
3	-58.1303	1.4786	1	0.224	.872738	2.70132	2.76088	2.86033
4	-57.4687	1.3232	1	0.250	.886048	2.71603	2.79049	2.9148

Endogenous: X

Exogenous: _cons

Nótese que el rezago (lag) número 1, tiene mayores * en la tabla desplegada por Stata. Esto significa que dentro del análisis de criterios de información a diferentes rezagos,

estos representan los menores valores posibles. En conclusión, el rezago no. 1 tiene tres * criterios a favor (LR, HQIC, SBIC), mientras que el rezago no. 2 tiene dos * criterios a favor (FPE, AIC).

En este caso se seleccionan los rezagos que tienen el mayor número de criterios a favor, que en este caso sería 1 rezago (1 lag).

Y procedemos a ejecutar la prueba Dickey Fuller, donde en Stata se especifica su rezago con el argumento lag(1). Si fuese otro rezago, por ejemplo 3, sería lag(3). Por ahora la prueba indica que lag(1) es el rezago ideal de acuerdo con la mayoría de criterios de información, y los resultados de la prueba Aumentada de Dickey Fuller se visualizan a continuación.

Augmented Dickey-Fuller test for unit root		Number of obs		=	48
		Interpolated Dickey-Fuller			
Test		1% Critical	5% Critical	10% Critical	
Statistic		Value	Value	Value	
Z(t)	1.694	-3.594	-2.936	-2.602	
MacKinnon approximate p-value for Z(t) = 0.9981					

Para entender a qué conclusión llegamos con esta prueba, es importante primero entender cuál es la hipótesis nula (H_0), en general, siempre debemos investigar cual es la hipótesis nula de cualquier prueba que realicemos para saber como interpretar el p-value. Para este caso, la hipótesis nula de prueba aumentada de Dickey-Fuller es la existencia de raíces unitarias de las variables. Lo que podríamos sintetizar como H_0 = Existencia de raíz Unitaria, y la hipótesis alterna como H_A = Ausencia de Raíz Unitaria.

Dado que el p-value ($p\text{ value} = 0.9981$) es mayor que un nivel de significancia del 5% (0.05), no se puede rechazar la hipótesis nula de que existen raíces unitarias en la serie x_t , por consiguiente, la serie x_t no es estacionaria³.

³ La mayoría de las veces los estudiantes tienen problemas para comprender como funciona el rechazo o aceptación de la hipótesis nula con el p-valor y el de significancia. Para empezar, como investigadores, trabajamos sobre probabilidad, y estamos dispuestos a aceptar un cierto nivel de error (o significancia del error) que se denomina α y a este valor se le asigna valores entre $\alpha = 0.1$, $\alpha = 0.05$ o $\alpha = 0.01$ (que también pueden ser leídos en porcentaje como 10%, 5%, 1%). Este valor α es nuestro nivel de significancia que es complementario al nivel de confianza. Y las reglas de interpretación aplican de la siguiente manera:

Si	$p\text{ valor} < \alpha$	entonces se rechaza la hipotesis nula
Si	$p\text{ valor} > \alpha$	entonces se acepta la hipotesis nula

3.6. Orden de integración

Hemos visto que la serie x_t no es estacionaria en niveles de acuerdo con los resultados anteriores de la prueba Aumentada de Dickey-Fuller, nótese que si obtenemos la primera diferencia de dicha serie y ejecutamos la prueba de nuevo con un rezago menos (dado que aplicamos primeras diferencias, los resultados ahora cambiarán como se muestra a continuación:

```
gen dX = X - X[_n-1] // Omitir si está repetido
dfuller dX, lag(0)
```

Obtendremos el siguiente resultado:

Dickey-Fuller test for unit root			Number of obs	=	48
		Interpolated Dickey-Fuller			
Test Statistic		1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-7.423	-3.594	-2.936	-2.602	
MacKinnon approximate p-value for Z(t) = 0.0000					

Donde dado que $p\text{ value} = 0.0000$ significa que rechazamos la hipótesis nula H_0 de existencia de raíces unitarias, y aceptamos la hipótesis alterna H_A de Ausencia de raíces unitarias, esto con un nivel de significancia α del 5%.

Por consiguiente, la conclusión es que la primera diferencia de x_t es estacionaria. Dado que x_t en niveles posee raíces unitarias, pero en primeras diferencias es estacionaria, significa que requerimos de $d = 1$ diferencias para volverla estacionaria. Esto define la propiedad más importante de la serie x_t y es que se encuentra integrada de primer orden, escrito formalmente como $x_t \sim I(1)$.

¿Por qué las diferencias de la forma $\Delta x_t = x_t - x_{t-1}$ tiende a volver una serie estacionaria? La respuesta radica en el principio de Pantula (1989), en el cual, a medida que vamos ejecutando diferencias de la serie x_t sobre sus valores pasados, la serie elimina componentes tendenciales, y las razones de cambio (diferencias) comienzan a comportarse a una media común. Una variable $I(1)$ es no-estacionaria en niveles, pero se vuelve estacionaria al utilizar la primera diferencia.

Los órdenes de integración también son sensibles a la especificación de la serie x_t . En este sentido, es común encontrar que las series macroeconómicas reales como el producto interno bruto real, la inflación, el desempleo, el stock de capital real, exportaciones o importaciones, sean integradas de primer orden $I(1)$, las cuales solo requieren una primera diferencia para volverse estacionarias. Por el contrario, las series nominales (no deflactadas), son generalmente series de orden de integración de segundo orden $I(2)$ lo cual es problemático. La sensibilidad de los órdenes de integración también depende de las

fluctuaciones o volatilidades de la serie de tiempo. Es posible, por ejemplo, que al aplicar una transformación de la forma $x_t^* = \ln x_t$ la serie pueda cambiar sus propiedades relativas al orden de integración ($x_t^* \sim I(1)$ cuando $x_t \sim I(2)$ ejemplificando la situación). Incluso, el calculo de valores per cápita puede afectar los ordenes de integración y la existencia de cointegración.

Finalmente, es de destacar que la prueba Dickey-Fuller tiene diferentes especificaciones. Las cuales pueden influir en la existencia o detección de raíces unitarias. Entre estas especificaciones encontramos: Sin constante, con tendencia, cambios/salto. Es conveniente en este punto, revisar en la literatura económica que tipo de especificaciones se utilizan para variables específicas, pero un buen punto de partida siempre es la especificación estándar de la prueba Aumentada Dickey-Fuller.

3.7. ¿Qué es la cointegración?

La cointegración es un fenómeno especial de dos o más series temporales no estacionarias y generalmente integradas de primer orden $I(1)$, las cuales convergen a un equilibrio o relación de largo plazo. También, es un buen remedio para corroborar la existencia de regresiones espurias (Engle & Granger, 1987). Es importante destacar que la cointegración es un análisis bi-variado o multi-variado, en contraste con el análisis uni-variado que hemos llevado a cabo. Para empezar, asumamos que tenemos:

$$y_t = \beta x_t + e_t \quad (1)$$

Donde $y_t \sim I(1)$ y $x_t \sim I(1)$, lo que quiere decir que tanto y como x son integradas de primer orden, no estacionarias en niveles, y que requieren una primera diferencia para volverse estacionarias.

Enfoquemos la atención sobre el orden de integración del termino e_t . Si $e_t \sim I(1)$ (es no estacionario y requiere una diferencia para ser estacionario) automáticamente la relación de (1) es espuria. Por el contrario, si $e_t \sim I(0)$, quiere decir que los residuales son estacionarios, y la interpretación cambia radicalmente porque implica que la relación (1) es una ecuación de largo plazo tendiente a un equilibrio no espurio. En este orden cuando e_t es estacionario, quiere decir que tanto x_t como y_t convergen en el largo plazo en equilibrio. Y el termino β es el coeficiente que mide la razón de cambio en el largo plazo. El termino e_t en este contexto, se interpreta como un residual de la ecuación de cointegración. Este residual, representa una combinación lineal de x_t y y_t que es estacionaria, y se representa de la forma $e_t = y_t - \beta x_t$ donde $e_t \sim I(0)$. Es importante destacar que, aunque e_t sea estacionario, no implica que no existan desviaciones entre x_t y y_t , de hecho, existen dichas desviaciones, pero en el largo plazo conforme t aumenta, las variables siempre convergen a un equilibrio.

3.8. Ecuación de cointegración a corrección del error

Debido a que a veces en la literatura econométrica no se encuentra el proceso paso a paso para llegar de la ecuación de cointegración, en este espacio se ofrece su desarrollo matemático para llegar al modelo de corrección del error. Para empezar, retomemos la ecuación de cointegración de la forma:

$$y_t = \beta x_t + e_t$$

Nosotros podemos rezagar la ecuación anterior en un rezago. Y obtenemos

$$y_{t-1} = \beta x_{t-1} + e_{t-1}$$

Ahora, si restamos las dos ecuaciones obtendremos que:

$$y_t - y_{t-1} = \beta x_t - \beta x_{t-1} + e_t - e_{t-1}$$

El cual puede ser expresado en primeras diferencias de la forma:

$$\Delta y_t = \beta(x_t - x_{t-1}) + \Delta e_t$$

$$\Delta y_t = \beta(\Delta x_t) + \Delta e_t$$

Aquí es donde es importante el supuesto de $e_t \sim I(0)$ lo que implica que la ecuación de los residuales de largo plazo es estacionaria, con un proceso de e_t con reversión a la media. Por consiguiente $e_t = \phi e_{t-1} + u_t$ donde $\phi < 1$. Si despejamos esta ecuación para dejar expresado todo en términos de Δe_t tenemos que:

$$e_t = \phi e_{t-1} + u_t$$

$$e_t - \phi e_{t-1} = u_t$$

$$\text{pero sabemos } \Delta e_t = e_t - e_{t-1} \text{ por ende } \Delta e_t + e_{t-1} = e_t$$

$$\Delta e_t + e_{t-1} - \phi e_{t-1} = u_t$$

$$\Delta e_t = e_{t-1} - \phi e_{t-1} + u_t$$

$$\Delta e_t = (1 - \phi)e_{t-1} + u_t$$

Con esta ecuación y definiendo el parámetro $(1 - \phi) = \rho$, la ecuación queda como $\Delta e_t = \rho e_{t-1} + u_t$ la cual puede ser reemplazada en $\Delta y_t = \beta(\Delta x_t) + \Delta e_t$ y por consiguiente nos da:

$$\Delta y_t = \beta(\Delta x_t) + \rho e_{t-1} + u_t$$

$$\Delta y_t = \beta \Delta x_t + \rho e_{t-1} + u_t$$

Nótese que si el coeficiente ρ es positivo, y mayor que 1, los residuales e_t no tienen un proceso de reversión a la media, por consiguiente, no se cumple la condición de que x_t y

y_t siguen un proceso estable en el largo plazo. En particular, ρ se define como el coeficiente de ajuste al equilibrio de largo plazo, el cual debe ser negativo y su valor absoluto debe ser menor que 1 para representar un proceso estable hacia el equilibrio.

Si t fueran años, la interpretación de ρ sería equivalente a una corrección cuasi-porcentual de forma anual entre y_t y x_t para retornar al equilibrio de largo plazo. Ejemplo: si $\rho = -0.2$ implica que la interpretación del término de ajuste equilibrio ρ demuestra que existe un ajuste del 20% de forma anual hacia el equilibrio de largo plazo entre x_t y y_t .

Finalmente, para observar cómo luce una ecuación de cointegración, el siguiente código simular la existencia de dos series cointegradas.

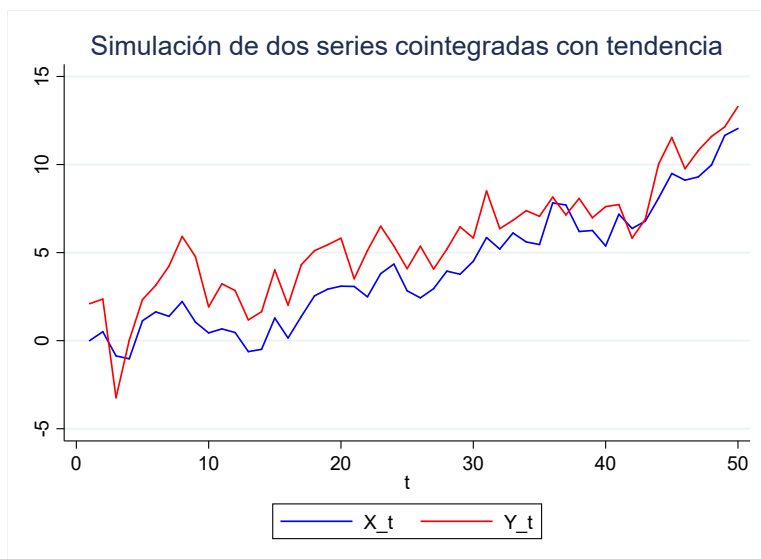
```
clear
set seed 123
set obs 50
gen t = _n
* 1. Simular X_t como un paseo aleatorio con tendencia positiva
gen X = 0
replace X = X[_n-1] + 0.01*t + rnormal(0,1) if _n > 1

* 2. Crear un término de error estacionario
gen epsilon = rnormal(0,1)

* 3. Generar Y_t con cointegración (relación de largo plazo) y tendencia
gen Y = 2 + 0.02*t + 0.8 * X + epsilon

* 4. Graficar ambas series
twoway (line X t, lcolor(blue)) (line Y t, lcolor(red)), ///
      legend(label(1 "X_t") label(2 "Y_t")) ///
      title("Simulación de dos series cointegradas con tendencia") ///
      graphregion(color(white))
```

Nótese que esta simulación define que $y_t = 2 + 0.02t + 0.8x_t + e$, donde también x_t es un proceso autorregresivo con tendencia positiva de la forma $x_t = x_{t-1} + 0.01t + u_t$. Por esta especificación de que y_t viene definido por el comportamiento de x_t , y que x_t tiene tendencia positiva, es que encontramos una relación que tiene la misma dirección.



4. Aplicación: Estimación VAR

En este caso, ejecutaremos la estimación de un modelo VAR basado en variables estacionarias en niveles de la forma $y_t, x_t \sim I(0)$ y ejecutaremos todos los pasos de la metodología estándar descritos en la sección 2. Simularemos entonces dos series temporales estacionarias y relacionadas para el análisis.

Para empezar, definiremos el sistema general del modelo VAR⁴ de la forma:

$$y_t = c_1 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=1}^p b_i x_{t-i} + e_{1,t}$$

$$x_t = c_2 + \sum_{i=1}^p l_i x_{t-i} + \sum_{i=1}^p d_i y_{t-i} + e_{2,t}$$

Donde y_t y x_t son nuestras variables de análisis, estacionarias, los términos c, a, b, l, d son coeficientes para estimar en el modelo. Y e son los componentes estocásticos que deben seguir una distribución normal de media 0 y varianza constante. En este sistema, claramente se observa que y_t y x_t hacen parte de procesos autorregresivos conjuntos de orden p rezagos. Este es un VAR de tipo restringido⁵.

Este modelo VAR estándar lo podemos representar matricialmente de la forma:

$$Y_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} \quad C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad \epsilon_t = \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \quad A_i = \begin{bmatrix} a_i & b_i \\ l_i & d_i \end{bmatrix}$$

Por lo que podemos reescribir el sistema como:

$$Y_t = C + \sum_{i=1}^p A_i Y_{t-i} + \epsilon_t$$

El cual consolida nuestro sistema de ecuaciones para la regresión con un conjunto Y de variables estacionarias que son y_t y x_t . En este contexto, simularemos a continuación la estructura de los procesos generadores de datos para estimar el VAR con la metodología estándar.

```
clear
set seed 12345
set obs 200
gen t = _n

* 1. Generate stationary AR(1) processes for X_t and Y_t
gen X = rnormal()
```

⁴ Es importante destacar que este es un modelo VAR de forma reducida, así mismo que existe una versión adicional que es el VAR estructural el cual se caracteriza por implementar relaciones contemporáneas. Sin embargo, este no será revisado en esta introducción a la metodología estándar.

⁵ La versión estructural implica definir coeficientes de efectos contemporáneos, sin embargo, esto no será abarcado en esta introducción.

```
gen Y = rnormal()

forvalues i = 2/200 {
    replace X = 0.5*X[_n-1] + rnormal(0,1) if _n == `i'
    replace Y = 0.3*Y[_n-1] + 0.2*X[_n-1] + rnormal(0,1) if _n == `i'
}

tsset t
```

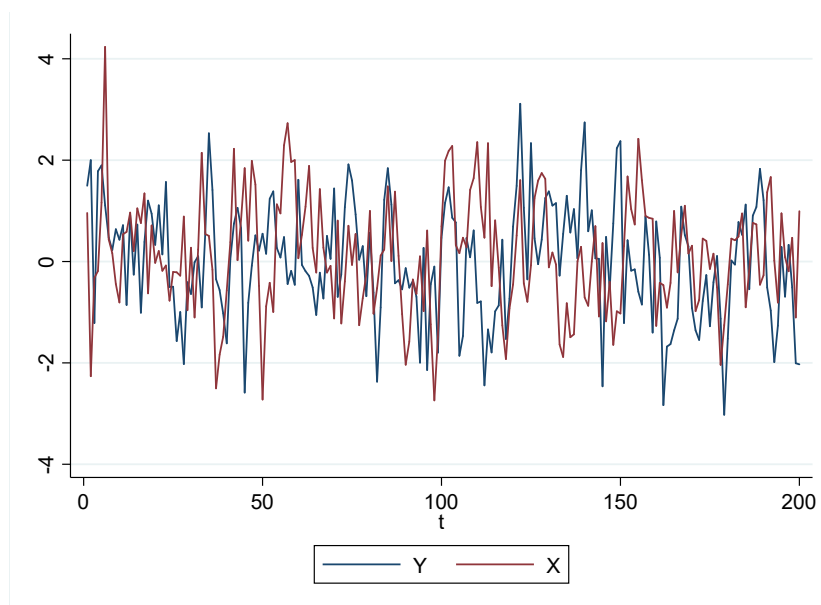
Bajo esta premisa bajos a empezar con la serie de pasos I al X en la estimación del VAR.

I. Análisis grafico

En este apartado procedemos a graficar las series del modelo VAR.

```
tsline Y
tsline X
tsline Y X
```

Del cual obtenemos un comportamiento estacionario:



Se destaca que efectivamente las series X, Y parecen seguir un proceso estacionario relacionado a una media de 0, la varianza parece constante de forma general. Así mismo existen 200 observaciones temporales de análisis para ambas series.

II. Análisis de raíces unitarias y ordenes de integración

Se destaca que efectivamente las series x y y tienen una interpretación visual de estacionarias sin tendencia, pero procederemos al análisis de raíces unitarias con la prueba formal de Dickey-Fuller. Para empezar, seleccionaremos la variable y con el número de rezagos ideales, y los resultados de la prueba:

```
varsoc X
dfuller X, lag(4)
```

```
Selection-order criteria
Sample: 5 - 200                                Number of obs   =      196
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-300.782				1.2732	3.07941	3.08618	3.09613
1	-277.507	46.55	1	0.000	1.01434	2.85211	2.86566*	2.88556*
2	-277.492	.03033	1	0.862	1.02458	2.86216	2.88248	2.91234
3	-276.924	1.1359	1	0.287	1.02912	2.86657	2.89366	2.93347
4	-274.425	4.9988*	1	0.025	1.0135*	2.85127*	2.88513	2.9349

```
Endogenous:  X
Exogenous:  _cons
```

Se observa que para la variable x se prefieren 4 rezagos bajo tres criterios (LR, FPE, AIC) y se ejecuta la prueba Dickey-Fuller con 4 rezagos.

```
Augmented Dickey-Fuller test for unit root                                Number of obs   =      195
```

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-6.107	-3.478	-2.884	-2.574

MacKinnon approximate p-value for Z(t) = 0.0000

Los resultados de la prueba rechazan la hipótesis nula de que la serie contiene raíces unitarias, y se acepta con un nivel de significancia del 5% la hipótesis alterna de que la serie es estacionaria. [Por consiguiente \$x_t\$ es estacionaria en niveles.](#)

Ahora con el análisis de la serie y procedemos de la misma manera:

```
varsoc Y
dfuller Y, lag(1)
```

```
. varsoc Y
```

Selection-order criteria

Sample: 5 - 200

Number of obs = 196

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-302.903				1.30105	3.10105	3.10782	3.11777
1	-282.2	41.405*	1	0.000	1.06409*	2.9*	2.91354*	2.93345*
2	-282.187	.02662	1	0.870	1.07486	2.91007	2.93038	2.96024
3	-281.974	.42481	1	0.515	1.08354	2.9181	2.94519	2.985
4	-281.948	.05206	1	0.820	1.09437	2.92804	2.9619	3.01167

Endogenous: Y

Exogenous: _cons

Y se destaca que, para esta variable y , el criterio indica la preferencia por 1 rezago de acuerdo con 5 criterios (LR, FPE, AIC, HQIC, SBIC). Por esta razón, estimamos la prueba de Dickey Fuller con un rezago. Y su resultado se muestra a continuación

```
. dfuller Y, lag(1)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 198

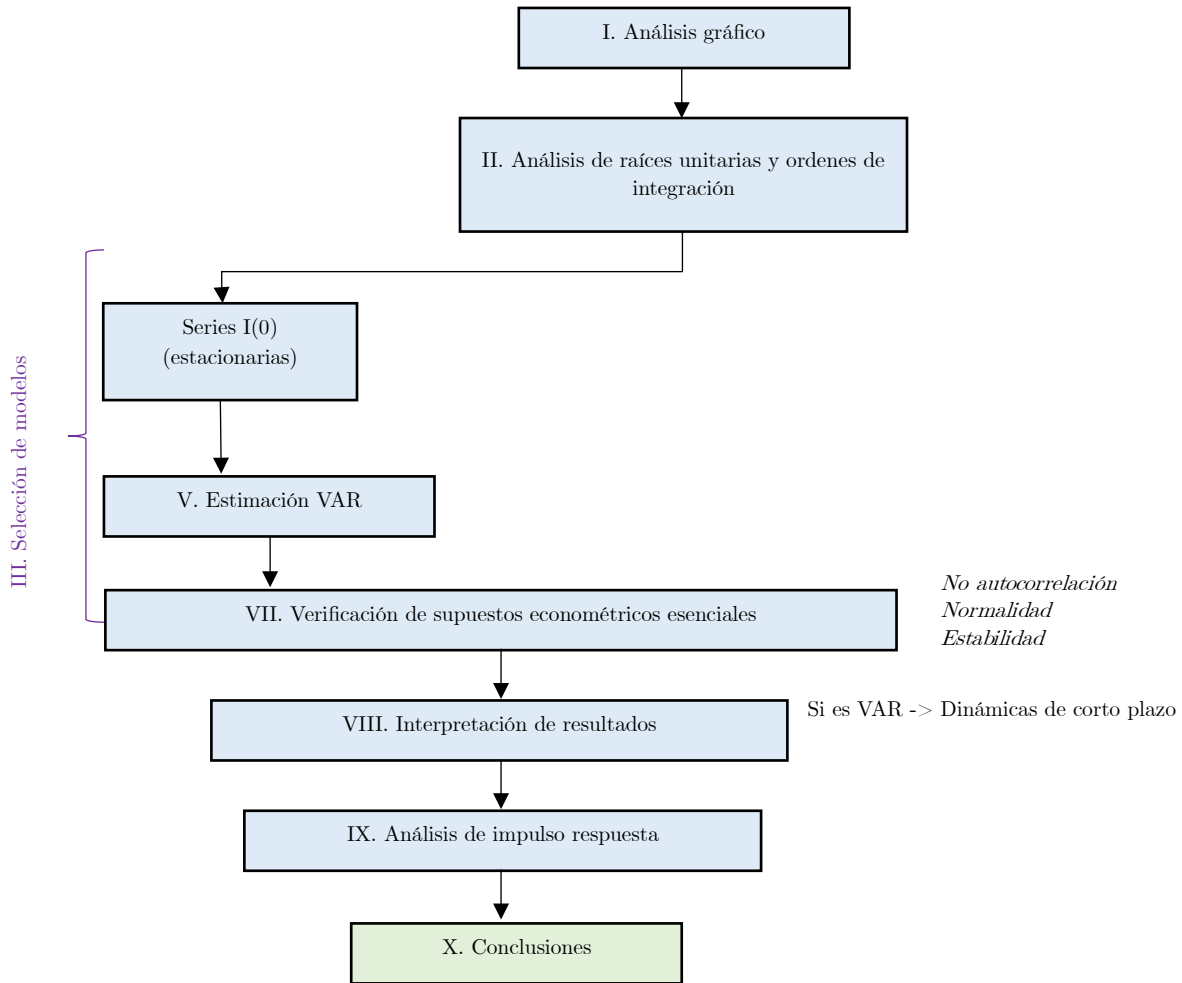
		Interpolated Dickey-Fuller		
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z (t)	-7.321	-3.477	-2.883	-2.573

MacKinnon approximate p-value for Z(t) = 0.0000

Los resultados indican que para la variable y se rechaza la hipótesis nula de que existen raíces unitarias, y se acepta la hipótesis nula con un nivel de significancia del 5% de que la variable es estacionaria. [Por consiguiente, la variable \$x_t\$ es estacionaria en niveles.](#)

III. Selección de mejor modelo

Considerando la figura 1 del apartado inicial, nos encontramos ante un conjunto de series (x y y) que son estacionarias en niveles, por lo que nuestro camino de acción se ve representado por la siguiente ruta a partir de la figura 1:



Por consiguiente, ante la presencia y confirmación de que todas las series son estacionarias escogeremos un Modelo VAR. El paso IV no aplica por no existir series integradas de primer orden en este ejercicio, por lo que el modelo VEC no procede.

V. Estimación VAR

Dado que hemos determinado que un modelo VAR es la aproximación econométrica que mejor se ajusta a las características de nuestras series temporales x y y (implicando que son estacionarias en niveles). Procedemos a estimar el VAR para esto también debemos ejecutar una de selección de rezagos para determinar el rezago optimo del VAR como guía.

* 3. Selección optima de rezagos
varsoc X Y, maxlag(5)

*4 Estimacion var
var X Y, lags(1)

Esto nos da como resultado una tabla con criterios de información, similar a los de Dickey Fuller donde seleccionamos por mayoría.

```
. varsoc X Y, maxlag(5)
```

Selection-order criteria
Sample: 6 - 200 Number of obs = 195

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-599.346				1.63541	6.16765	6.18124	6.20122
1	-555.232	88.228*	4	0.000	1.0838*	5.75622*	5.797*	5.85693*
2	-553.184	4.0954	4	0.393	1.10574	5.77624	5.8442	5.94409
3	-550.835	4.6981	4	0.320	1.12466	5.79318	5.88832	6.02816
4	-547.508	6.6539	4	0.155	1.13253	5.80008	5.92241	6.1022
5	-547.026	.96439	4	0.915	1.17427	5.83616	5.98567	6.20542

Endogenous: X Y
Exogenous: _cons

Y se destaca que los criterios LR, FPE, AIC, HQIC, SBIC favorecen la escogencia de un VAR de orden 1. Por lo que se define que para x y y es mejor un VAR(1). Los resultados de esta estimación se muestran a continuación:

```
. var X Y, lags(1)
```

Vector autoregression

Sample: 2 - 200
Log likelihood = -573.0049
FPE = 1.154186
Det(Sigma_ml) = 1.086639

Equation	Parms	RMSE	R-sq	chi2	P>chi2
X	3	1.01652	0.1999	49.71811	0.0000
Y	3	1.04164	0.1855	45.31152	0.0000

Number of obs = 199
AIC = 5.819145
HQIC = 5.859332
SBIC = 5.91844

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
X	Resultados ecuación para X						
	X						
	L1.	.4430044	.0635469	6.97	0.000	.3184548	.5675541
	Y						
	L1.	.0381754	.0628007	0.61	0.543	-.0849116	.1612624
	_cons	.0745215	.0720285	1.03	0.301	-.0666517	.2156948
Y	Resultados ecuación para Y						
	X						
	L1.	.1117769	.065117	1.72	0.086	-.01585	.2394039
	Y						
	L1.	.4108975	.0643523	6.39	0.000	.2847693	.5370257
	_cons	-.0202386	.0738081	-0.27	0.784	-.1648998	.1244227

La salida de la regresión indica que tenemos 199 observaciones de la cual se estimó el modelo junto con sus criterios de información del error (FPI, AIC, HQIC, SBIC), Tenemos también por ecuación los R^2 por ecuación, y el nivel de significancia global $p < chi2$. El modelo indica que para la primera ecuación existe un $R^2 = 0.1999$ lo que sugiere que para la

variable x el modelo tiene una bondad de ajuste lineal del 19.99%, o que el modelo explica linealmente a la variable x en un 19.99%. Aquí, sin embargo, dado que el objeto del modelo VAR no es la predicción para nuestro caso, nos concentramos en la siguiente parte que es asociada la inferencia estadística.

Nótese que después de esta tabla inicial, viene una tabla que tiene los coeficientes del modelo (en morado) y los p-valores (rojo, verde). Así mismo a la izquierda está el nombre de la variable sobre la cual se está estimando el coeficiente. Por ejemplo, para el caso de x existe un coeficiente de .4430044 asociado $L1.x$ que hace referencia a $0.4430044x_{t-1}$ indicando que el primer rezago de x (es decir x_{t-1}) tiene un impacto sobre el valor actual de x de 0.4430 unidades ceteris paribus, estadísticamente significativo al 5% dado que el p-valor es 0.000. Esto es característico de un proceso autorregresivo.

El siguiente coeficiente es de .0381754 que pertenece a y_{t-1} sobre x_t . Sin embargo, dado su pvalor de 0.543, no es estadísticamente significativo al 5%. Por lo que se determina que la variable y_{t-1} (o primer rezago de y_t) no tiene influencia o impactos sobre x_t .

Finalmente, el termino constante de 0.74 no es estadísticamente significativo porque su pvalor es de 0.301.

Ahora la tabla nos muestra la siguiente ecuación que es la ecuación de y_t . En este caso, observamos que el coeficiente de x_{t-1} (primer rezago de x) es de .1117769, y su p valor es de 0.086. Razón de la cual interpretamos que con un nivel de significancia del 10%, la variable x_{t-1} (o primer rezago de x) es estadísticamente significativo para explicar impactos en y_t , y su impacto es de 0.1177 sobre el valor actual de y_t . Nótese que aquí trabajamos el pvalor es de 0.086 que es menor de un nivel de significancia $\alpha = 0.1$ del 10%, por consiguiente, pvalor < nivel de significancia. Y rechazamos la hipótesis nula de que el coeficiente sea igual a 0, y aceptamos la hipótesis alterna de que es diferente de 0.

El siguiente coeficiente de la ecuación de y_t pertenece a y_{t-1} (o primer rezago de y) el cuales de .41089, y es estadísticamente significativo porque el p valor (0.000) es menor que un nivel de significancia del 5%. Por lo que se interpreta que el impacto de y_{t-1} sobre y_t es de 0.41089 unidades, estadísticamente significativo al 5%.

El paso VI, no es necesario puesto que estamos trabajando sobre un VAR y el modelo VEC fue descartado por no cumplir con las características necesarias.

VI. Estimacion modelo VEC (omitido por no corresponder).

VII. Verificación de supuestos econométricos

Los supuestos que debe cumplir el modelo VAR son: 1) No autocorrelación, 2) Normalidad, y 3) Estabilidad de parámetros. Estos se pueden ejecutar directamente después de la estimación del VAR con los siguientes códigos:

```
var X Y, lags(1)
* LM test for autocorrelation in residuals
varlmar
* test for normally distributed residuals
varnorm
* check stability condition of estimates
varstable, graph
```

Los cuales dan como resultado:

No autocorrelación

```
. varlmar
```

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	7.3190	4	0.11996
2	2.9989	4	0.55800

H0: no autocorrelation at lag order

Donde la hipótesis nula es la ausencia de correlación en el rezago de orden (1 o 2). En este caso como pvalor es 0.11 el cual es mayor que un nivel de significancia del 5%, no se puede rechazar la hipótesis de ausencia de correlación serial en el rezago, similar ocurre en el rezago porque el pvalor es de 0.558 el cual es mayor que un nivel de significancia del 5%. Indicando que no se puede rechazar la hipótesis nula de no autocorrelación serial en el rezago 2.

Normalidad

```
. varnorm
```

Jarque-Bera test

Equation	chi2	df	Prob > chi2
X	0.291	2	0.86460
Y	0.741	2	0.69047
ALL	1.032	4	0.90494

Las pruebas de normalidad también son satisfactorias, aquí la hipótesis nula es que los residuales de la ecuación de x, o de y, o en general del modelo, siguen una distribución normal. Dado que los p-valores son 0.86, 0.96, y 0.90 no se puede rechazar la hipótesis nula

de que los residuales (sea de x , o y o general) siguen una distribución normal. Por consiguiente, los residuales tienen un comportamiento normal.

Estabilidad de parámetros

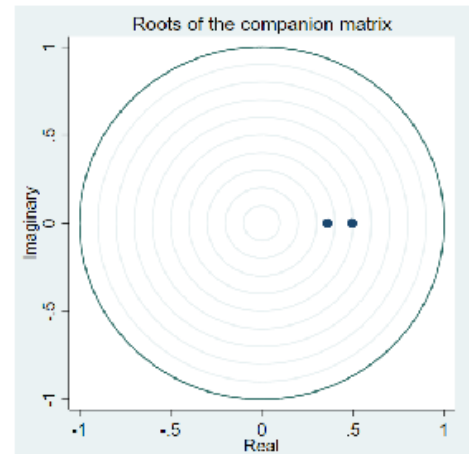
Finalmente verificaremos el supuesto de estabilidad de parámetros, el cual tiene un resultado estable dado que los valores caen dentro del círculo de la unidad.

```
. varstable, graph
```

Eigenvalue stability condition

Eigenvalue	Modulus
.4942179	.494218
.359684	.359684

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.



En este caso, el círculo de la unidad representa un espacio o área de estabilidad de los coeficientes autorregresivos del VAR. Si los coeficientes fueran mayores que 1, esto implicaría que el proceso inestable y explosivo. Sin embargo, para este caso, se cumplen que todos los coeficientes son estables al estar dentro del círculo de la unidad (y los módulos son menores que 1).

VIII. Interpretación.

Los resultados de los coeficientes fueron interpretados en el apartado de estimación de la sección V.

Interpretaciones adicionales sugieren que el modelo cumple satisfactoriamente con los supuestos econométricos estándar, así mismo, que las relaciones no son del todo bidireccionales. Debido a que para la ecuación de x , el coeficiente de y_{t-1} no es estadísticamente significativo. Pero en el caso de y_t si lo es el coeficiente de x_{t-1} lo que indica que en el sentido de Granger, x_t si causa a y_t pero no al revés. Para observar esta causalidad en el sentido de Granger tenemos:

```
var X Y, lags(1)
vargranger
```

```
. vargranger
```

```
Granger causality Wald tests
```

Equation	Excluded	chi2	df	Prob > chi2
X	Y	.36952	1	0.543
X	ALL	.36952	1	0.543
Y	X	2.9466	1	0.086
Y	ALL	2.9466	1	0.086

En este caso, la causalidad en el sentido de Granger se da de la siguiente manera. La primera columna indica la ecuación sobre la cual se hace el análisis, la segunda es la ecuación sobre la cual se sospecha que puede causar un efecto en la ecuación. Es decir, la primera línea indica que se está probando la hipótesis de que $Y \rightarrow X$ causa en el sentido de granger. Sin embargo, dado que el p-valor es de 0.543, no se puede rechazar de que dicha sea nula. Por consiguiente, se concluye que, Y no causa a X en el sentido de Granger, con un nivel de significancia del 5%.

Para la segunda ecuación, Y, aquí el orden de causalidad se invierte y decimos que $X \rightarrow Y$, con un p valor de 0.086. Por consiguiente, con un nivel de significancia del 10%, rechazamos la hipótesis nula de X no causa a Y y aceptamos la hipótesis alterna de que X si causa a Y en el sentido de Granger.

Es de destacar que hoy en día, la literatura econométrica no utiliza la causalidad como tal cuando hablamos de la causalidad en el sentido de Granger, debido a que es muy difícil probar la exogeneidad de los regresores del modelo. En este sentido, aunque si es un tipo de causalidad en el sentido de Granger, realmente no es la causalidad que esperamos en econometría (donde los regresores son enteramente exógenos a los componentes estocásticos).

IX. Función de impulso respuesta

Finalmente, el último apartado del VAR, es la posibilidad de estimar funciones de impulso respuesta, estas funciones permiten entender como responden las innovaciones de una variable sobre otra, dada la naturaleza endógena de la modelización econométrica de los VAR. En este contexto, con el VAR es posible desarrollar una función de impulso respuesta (IRF) para comprender como funcionan los choques de una variable sobre otras, esto es fácilmente estimable en Stata para los modelos VAR, sin embargo, los modelos VEC no tienen esta posibilidad en este software.

```
* 4. Estimate the VAR model with the selected lag length
var X Y, lags(1)
```

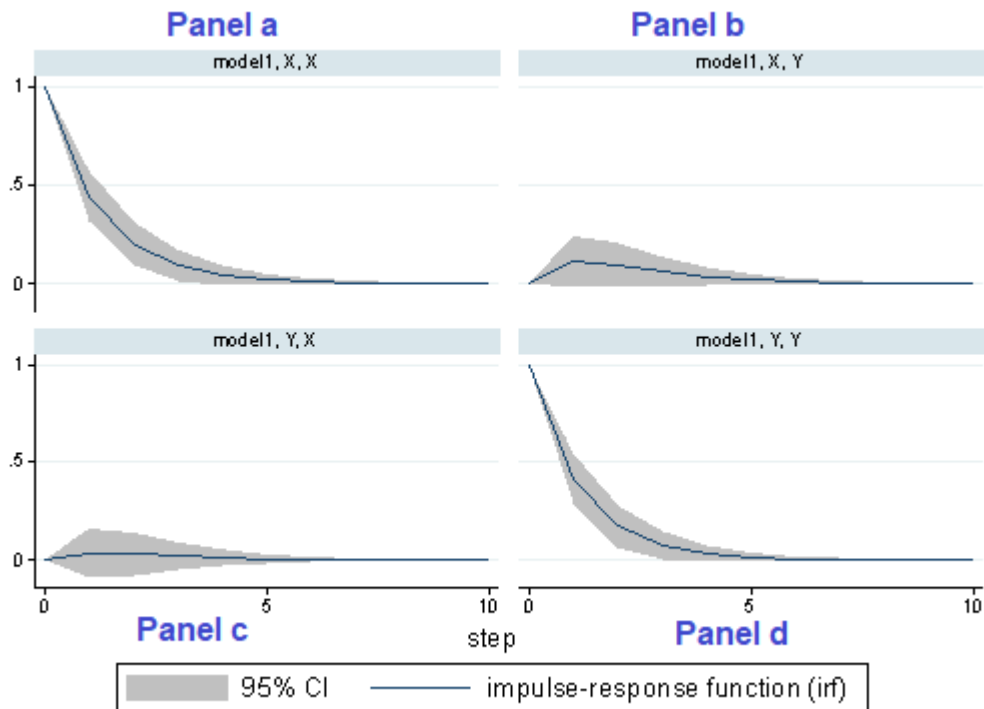
```
* 5. Create and store the IRF file
irf set results_irf, replace
```

```
irf create model1, step(10)
```

* 6. Plot the IRF with White Background

```
irf graph irf, graphregion(color(white)) plotregion(color(white)) bgcolor(white)
irf table irf, step(10)
```

El resultado del código anterior, permite utilizar el modelo VAR(1) estimado, y luego se genera la función de impulso respuesta, así mismo se define un horizonte temporal de 10 años (o unidades de tiempo que se manejen) para observar cómo afectan los impulsos en términos de respuesta de las variables de análisis.



Graphs by irfname, impulse variable, and response variable

En este caso, la IRF permite observar las respuestas por un impulso ejecutado en alguna de las variables del modelo. El área gris representa los intervalos de confianza por donde la respuesta puede llegar a oscilar con un 95% de confianza. En la parte superior se encuentran títulos del estilo “model1, X, X” que significan que nuestro model1 (var estimado), presenta el impulso en la variable X y la respuesta en la variable X. En el Panel b, por ejemplo, esta situación se refiere a “model1, X, Y”, que significa que ejecutó un impulso en X para observar la respuesta de Y.

Para la interpretación de paneles, empecemos por el panel a, que contiene el impulso de x sobre si misma (impulso de x hacia x). En este caso, se observa una curva con pendiente decreciente que deja de ser estadísticamente significativa en el periodo 4. Este comportamiento de decaimiento es clásico de un proceso autorregresivo estacionario, dado

que x va perdiendo influencia sobre si misma a través del tiempo dado que el impacto se diluye en el horizonte temporal.

Para el panel b, en cambio ahora se pasa a analizar un impulso sobre la variable x y la respuesta de la variable y . En este caso, existe un leve efecto positivo, pero casi que es estadísticamente no significativo al 5% (seguramente al 10% lo es por la tabla de regresión vista anteriormente). Sin embargo, este efecto positivo de x se desvanece y se diluye en el periodo 2 y 3.

Para el panel c, en este caso se analiza un impulso en y para observar la respuesta de x . Sin embargo, consistente con la salida de la regresión, pese a que hay un leve efecto positivo en el primer periodo, este no es estadísticamente significativo, así como en el futuro.

Finalmente, el panel d, muestra el impacto de y sobre si misma en el tiempo. Se destaca que este se va diluyendo en el tiempo y deja de ser estadísticamente significativo en el periodo 4. Claro ejemplo de un proceso autorregresivo cuya inercialidad tiende a la estabilidad por ser estacionario.

Otra interpretación, para ser específicos en la magnitud, se da en termino de desviaciones estándar. Por ejemplo, Si la respuesta en y_t ante un shock de x_t es de 0.5 en el primer periodo, el incremento en una desviación estándar expresada como σ_x genera una respuesta de 0.5 unidades en y_t . Para nuestro caso aplicado:

➔ Al introducirse un impulso de una desviación estándar en x_t se produce una respuesta de 0.011 unidades en promedio en y_t

Esta interpretación se puede extender a todos los otros paneles. La variable de impulso usualmente se mide en términos del incremento de una desviación estándar. Pero la respuesta se da en unidades promedio de la variable de respuesta.

El código genera la tabla de impactos de shocks sobre la variable respuesta, donde se puede cuantificar dichos cambios ejemplificados en el grafico anterior, sin embargo, esos no serán discutidos.

Ejemplo adicional: Para un VAR(1) de forma general:

$$x_t = \alpha_1 + \phi_{11}x_{t-1} + \phi_{12}y_{t-1} + e_{xt}$$

$$y_t = \alpha_2 + \phi_{21}x_{t-1} + \phi_{22}y_{t-1} + e_{yt}$$

Los errores o innovaciones del sistema son e_{xt}, e_{yt} y por consiguiente un shock en x_t (primera ecuación) significa que aumenta e_{xt} en una desviación estándar σ_x

Por el contrario, si existe un shock en y_t (segunda ecuación) significa aumentar e_{yt} en una desviación estándar σ_y .

Por esta razón pese a que los impulsos se midan en desviaciones estándar σ_x o σ_y la respuesta de dicho impulso se da en términos de las variables x, y y de forma promedio.

X. Conclusiones

De acuerdo con lo visto en este apartado. Las series de tiempo simuladas x_t y y_t para este ejercicio son estacionarias en niveles. Sin embargo, el orden de causalidad en el sentido de Granger refleja solamente un sentido de causalidad-Granger de x_t hacia y_t .

Así mismo, el efecto de x_t hacia y_t de una desviación estándar de x_t solamente se ve reflejado en un incremento de 0.11 unidades en promedio sobre y_t .

La variable x_t es estadísticamente significativa al 10% de significancia para explicar el comportamiento de y_t . Sin embargo, y_t no parece explicar al proceso x_t .

Finalmente, x_t y y son procesos estacionarios que parecen seguir un patrón autorregresivo, dado el comportamiento de los impulsos y respuestas sobre si mismas individualmente y los resultados de las raíces unitarias. La interpretación de coeficientes del VAR se puede encontrar en el apartado de estimación (V), y también interpretaciones adicionales VIII y IX.

Ahora, en el siguiente apartado del texto, pasaremos a la versión cointegrada con series no estacionarias integradas de primer orden. A continuación, se presenta todos los comandos utilizados en el actual apartado.

```
clear
set seed 12345
set obs 200
gen t = _n

* 1. Generate stationary AR(1) processes for X_t and Y_t
gen X = rnormal()
gen Y = rnormal()

forvalues i = 2/200 {
    replace X = 0.5*X[_n-1] + rnormal(0,1) if _n == `i'
    replace Y = 0.3*Y[_n-1] + 0.2*X[_n-1] + rnormal(0,1) if _n == `i'
}

tsset t

tsline Y
tsline X
tsline Y X

varsoc X
dfuller X, lag(4)

varsoc Y
dfuller Y, lag(1)
```

* 3. Selección optima de rezagos

varsoc X Y, maxlag(5)

*4 Estimacion var

var X Y, lags(1)

* LM test for autocorrelation in residuals

varlmar

* test for normally distributed residuals

varnorm

* check stability condition of estimates

varstable, graph

var X Y, lags(1)

vargranger

var X Y, lags(1)

* 5. Create and store the IRF file

irf set results_irf, replace

irf create model1, step(10)

* 6. Plot the IRF with White Background

irf graph irf, graphregion(color(white)) plotregion(color(white)) bgcolor(white)

irf table irf, step(10)

5. Aplicación: Estimación VEC

En este apartado, a diferencia del anterior, simularemos dos series que no son estacionarias en niveles y que se encuentran cointegradas en el largo plazo. Por consiguiente, x_t y y_t se encuentran integradas de primer orden $I(1)$. Por lo que abarcaremos todos los pasos descritos en la metodología estándar de estimación de VEC, descritos en la figura 1, para este tipo de modelos.

Antes de iniciar la simulación, es prudente y útil entender como un modelo VAR puede derivar en un modelo de Vectores con Corrección del Error (VEC de Vector Error Correction). Empecemos con un VAR(p) que es un VAR de p rezagos, o de orden p:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + e$$

Donde Y_t es un vector de variables endógenas de la forma $Y_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$, A_i son matrices de coeficientes que acompañan estas variables. p es el orden del VAR. Y e es el vector de residuales o errores de cada ecuación.

Si restamos Y_{t-1} a ambas partes de la ecuación anterior obtendríamos:

$$Y_t - Y_{t-1} = -Y_{t-1} + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + e$$

El cual dado que $\Delta Y_t = Y_t - Y_{t-1}$ el resultado seria:

$$\Delta Y_t = (A_1 - I)Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + e$$

Donde I es la matriz de identidad. En este punto es importante jugar un poco con el algebra, y generalicemos que la primera diferencia puede ser reescrita como rezago dado que:

$$\Delta Y_t = Y_t - Y_{t-1} \quad \rightarrow \quad Y_{t-1} = Y_t - \Delta Y_t$$

$$\Delta Y_{t-1} = Y_{t-1} - Y_{t-2} \quad \rightarrow \quad Y_{t-2} = Y_{t-1} - \Delta Y_{t-1}$$

$$\Delta Y_{t-2} = Y_{t-2} - Y_{t-3} \quad \rightarrow \quad Y_{t-3} = Y_{t-2} - \Delta Y_{t-2} \rightarrow Y_{t-3} = Y_{t-1} - \Delta Y_{t-1} - \Delta Y_{t-2}$$

$$\Delta Y_{t-i+1} = Y_{t-i} - Y_{t-i+1} \quad \rightarrow \quad Y_{t-i} = Y_{t-i+1} + \Delta Y_{t-i+1}$$

En este contexto, Y_{t-1} siempre estará presente en cada diferencia que se exprese, dado que está contenida dentro de las diferencias históricas en el cálculo.

Y aquí sustituimos los términos Y_{t-2} en delante de la ecuación $\Delta Y_t = (A_1 - I)Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + e$ sabiendo que Y_{t-1} va a estar implícito en cada uno de ellos de acuerdo con lo visto anteriormente, por lo que tendríamos:

$$\Delta Y_t = (A_1 - I)Y_{t-1} + A_2(Y_{t-1} + \Delta Y_{t-1}) + A_3(Y_{t-2} + \Delta Y_{t-2}) + \dots + A_p(Y_{t-p+1} - \Delta Y_{t-p+1}) + e$$

Al distribuir coeficientes A_p en la ruptura de paréntesis desde A_2 en adelante, es posible factorizar todos los coeficientes A en el termino inicial Y_{t-1} .

$$\Delta Y_t = (A_1 + A_2 + A_3 + \dots + A_p - I)Y_{t-1} + A_2\Delta Y_{t-1} + A_3\Delta Y_{t-2} + \dots + A_p\Delta Y_{t-p+1} + e$$

Por lo que podemos generalizar todos los términos Y_{t-p} de la anterior ecuación al definir la siguientes sumatorias para asegurar una representación más sencilla:

$$\Pi = \sum_{i=1}^p (A_i - I)$$

$$\Gamma_i = - \sum_{j=i+1}^p j$$

El cual al agruparlos juntos coinciden en la expresión estándar de los vectores con corrección del error. El cual, es un modelo de corto plazo donde existen coeficientes de ajuste a largo plazo en Π y coeficientes de corto plazo de las diferencias rezagadas Γ_i .

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + e$$

Este es el modelo más sencillo y estándar de Vectores con Corrección del Error o VEC(p-1), y es la especificación común más aplicada. El termino ΔY_{t-i} hace referencia a las primeras diferencias rezagadas hasta p-1 rezagos del vector de variables endógenas.

La serie simulada para $x_t, y_t \sim I(1)$ y será entonces:

```
clear
set seed 123
set obs 100

gen t = _n // Generar un índice de tiempo
// Generate a common stochastic trend
gen s = 0
gen e_s = rnormal()

forvalues i = 2/100 {
    replace s = s[_n-1] + e_s in `i' // Stochastic trend
}

// Generate two cointegrated series around the trend
gen y = 0.1*s + 0.1*t + rnormal()
gen x = 0.1*s + 0.1*t + 0.01*y + rnormal() // Cointegration relationship

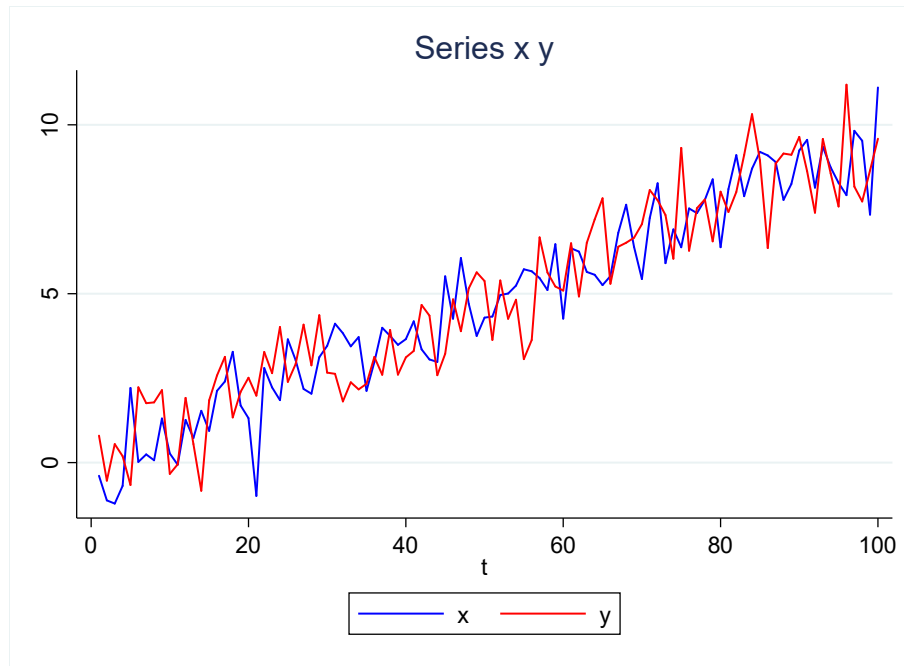
tsset t
```

I. Análisis gráfico de las series temporales

En esta sección se procede con el primer paso de la metodología estándar (Sección 2, Figura 1) el cual es el análisis gráfico, considerando nuestras series x_t y y_t .

```
// Graficar las series
twoway (line x t, lcolor(blue)) ///
      (line y t, lcolor(red)), ///
      title("Series x y") ///
      legend(label(1 "x") label(2 "y"))
```

El cual da como resultado:

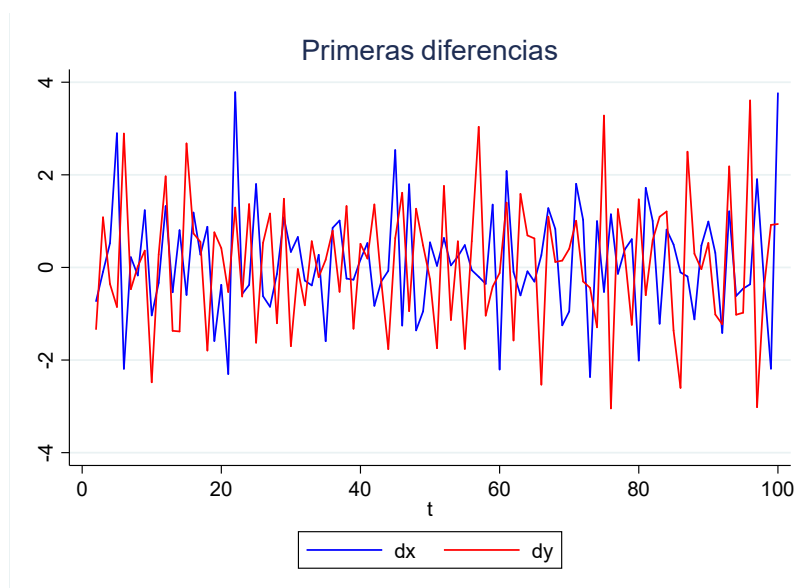


Donde se observa que tanto x como y no parecen ser estacionarias debido a los quiebres constantes en la media, creciente tendencia, así mismo, existe cierto patrón entre x y y pero con diferencias significativas en algunos periodos de tiempo. La tendencia, sin embargo, parece dirigir a los dos procesos en el transcurso del tiempo.

Ahora observemos las primeras diferencias y grafiquemos:

```
gen dx = x - L.x
gen dy = y - L.y

twoway (line dx t, lcolor(blue)) ///
      (line dy t, lcolor(red)), ///
      title("Primeras diferencias") ///
      legend(label(1 "dx") label(2 "dy"))
```



Es destacable visualmente que es difícil observar movimientos comunes en primeras diferencias, sin embargo, la conclusión primordial de este gráfico es que las series sí parecen seguir un comportamiento estacionario al usar la primera diferencia de ambas. Por lo que se presume y se interpreta que son series integradas de primer orden. Sin embargo, pasaremos a las pruebas de raíces unitarias formales para verificar esto.

II. Análisis de raíces unitarias y definición del orden de integración

En este apartado ahora comprobaremos el orden de integración de las series, recuerde que debe utilizar un criterio de información para seleccionar los rezagos ideales de la prueba de Dickey-Fuller como guía. Así que utilizaremos los criterios de información estándar primero sobre x :

```
varsoc x
dfuller x, lag(4)
```

El cual da como resultado:

```
. varsoc x
```

Selection-order criteria

Sample: 5 - 100

Number of obs = 96

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-234.995				7.99402	4.91657	4.92737	4.94328
1	-151.858	166.27	1	0.000	1.44413	3.20538	3.22697	3.2588
2	-142.669	18.379	1	0.000	1.21763	3.03477	3.06716	3.1149
3	-136.874	11.59	1	0.001	1.1019	2.93487	2.97806	3.04172
4	-133.182	7.3829*	1	0.007	1.04187*	2.8788*	2.93278*	3.01236*

Endogenous: x

Exogenous: _cons

Donde se define que para la prueba Dickey-Fuller aumentada, debemos utilizar idealmente 4 rezago. Al ejecutar la prueba tenemos que:

```
. dfuller x, lag(4)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 95

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z (t)	-0.628	-3.517	-2.894	-2.582

MacKinnon approximate p-value for Z(t) = 0.8644

Donde con un nivel de significancia del 5%, no podemos rechazar la hipótesis nula de existencia de raíces unitarias en x . Esto dado el p valor de 0.8644 que es mayor que el nivel de significancia del 5%. [Es decir que la variable no es estacionaria y posee raíces unitarias.](#)

Por consiguiente, ahora procedemos a diferenciar x y analizar el comportamiento de su primera diferencia de la forma $\Delta x_t = x_t - x_{t-1}$ con $n - 1$ rezagos

* opcional varsoc, pero se puede ejecutar la prueba con n-1 rezagos
varsoc dx
dfuller dx, lag(3)

Y obtenemos que:

```
. dfuller dx, lag(3)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 95

		Interpolated Dickey-Fuller		
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z (t)	-8.257	-3.517	-2.894	-2.582

MacKinnon approximate p-value for Z(t) = 0.0000

Donde ejecutamos la prueba con n-1 rezagos sobre la primera diferencia de x . En este caso, se rechaza la hipótesis nula de que la primera diferencia de x tenga raíces unitarias y

aceptamos la hipótesis alterna de que la primera diferencia de x es estacionaria con un nivel de significancia del 5%. Esto debido a que el p valor = 0.0000 es menor que 0.05 del nivel de significancia.

Por ende, definimos que x_t es un proceso integrado de primer orden definido como $x_t \sim I(1)$, dado que solo requiere una diferencia para volverse estacionaria.

Proseguimos ahora con la variable y_t .

```
varsoc y
dfuller y, lag(4)
```

Donde se observa que:

```
. varsoc y

Selection-order criteria
Sample: 5 - 100                Number of obs   =          96
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-235.085				8.00897	4.91844	4.92924	4.94515
1	-164.271	141.63	1	0.000	1.87033	3.46398	3.48558	3.51741
2	-155.927	16.688	1	0.000	1.605	3.31098	3.34337	3.39112
3	-147.859	16.136	1	0.000	1.38527	3.16373	3.20692	3.27057*
4	-145.716	4.2866*	1	0.038	1.35273*	3.13991*	3.19389*	3.27347

```
Endogenous:  y
Exogenous:  _cons
```

El cual, tenemos un rezago de 4 niveles como preferencia de acuerdo con todos los criterios. En particular esto indica que ejecutaremos la prueba Dickey Fuller con cuatro rezagos.

```
. dfuller y, lag(4)

Augmented Dickey-Fuller test for unit root                Number of obs   =          95
```

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-1.141	-3.517	-2.894	-2.582

```
MacKinnon approximate p-value for Z(t) = 0.6986
```

Al observar que el p valor es de 0.6986, que es mayor que un nivel de significancia de 0.05, no se puede rechazar la hipótesis nula de que la serie y tiene raíces unitarias y por consiguiente no es estacionaria.

```
dfuller dy, lag(3)
```

Al analizar la primera diferencia de y con $n-1$ rezagos, tenemos los siguientes resultados:

Metodología estándar de vectores autoregresivos (VAR) y de corrección del error (VEC)

```
. dfuller dy, lag(3)
```

Augmented Dickey-Fuller test for unit root Number of obs = 95

Test Statistic	Interpolated Dickey-Fuller		
	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-8.038	-3.517	-2.894

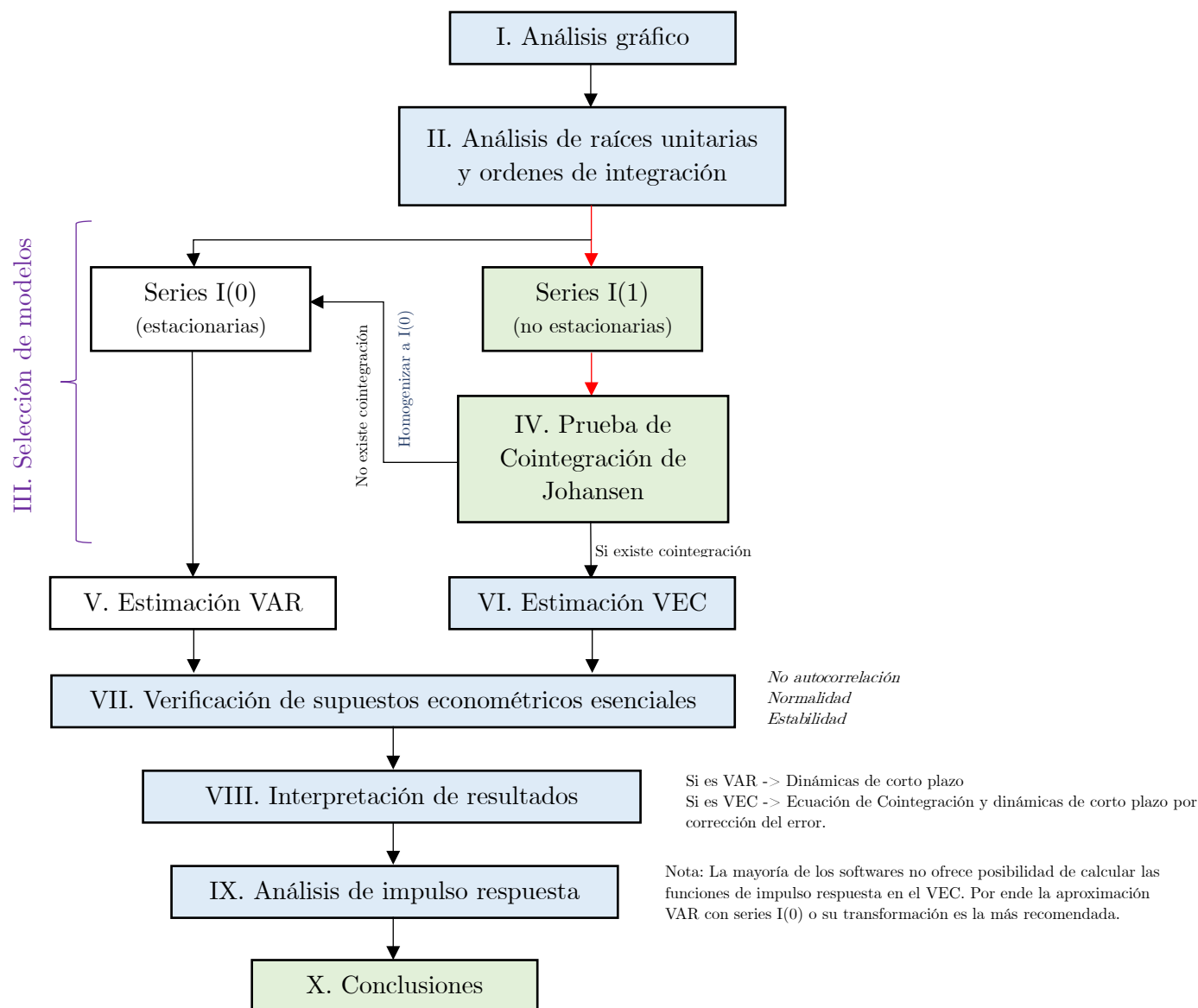
MacKinnon approximate p-value for Z(t) = 0.0000

Con este resultado, se rechaza la hipótesis nula de que existen raíces unitarias en la primera diferencia de y y se acepta la hipótesis alterna de que la serie Δy_t es estacionaria. Esto debido a que el p valor es menor que un 5% de significancia.

Esto confirma que el orden de integración de la serie y_t es de primer orden, dado que solo requiere una primera diferencia para volverse estacionaria. Por ende $y_t \sim I(1)$

III. Selección de modelos con base en la definición de ordenes de integración

Dado que nos encontramos ante un conjunto de variables que poseen raíces unitarias pero las cuales al ser diferencias una vez se vuelven estacionarias, sabemos que el conjunto $x_t, y_t \sim I(1)$ es integrado de primer orden. Por consiguiente, de acuerdo con la figura 1 de la metodología estándar, debemos de proceder como se muestra a continuación:



Por lo cual el siguiente paso, es ejecutar la prueba de cointegración. Y si existe cointegración, ejecutaremos el modelo VEC. De lo contrario, transformaríamos las series en primeras diferencias para que sean estacionarias e iríamos por el modelo VAR.

IV. Implementación de pruebas de cointegración (solo para series $I(1)$ únicamente)

Dado que cumplimos el requisito de que tenemos series $I(1)$ en nuestro sistema, procederemos a la prueba de cointegración de Johansen, esta prueba tuvo una trayectoria desde Johansen (1988), Johansen & Juselius (1990), hasta Johansen (1991). En esencia, la prueba realiza un test de prueba sobre la existencia de relaciones de cointegración en los modelos vectoriales. De esta manera, al existir una ecuación de cointegración, el estadístico nos indica si podemos rechazar o no su existencia.

La prueba de Cointegración de Johansen también necesita una selección de rezagos para su ejecución. Algunos autores sugieren utilizar $n - 1$ rezagos ideales, otros sugieren dejar los n rezagos. Otros autores argumentan que n rezagos óptimos corresponden al VAR, y por consiguiente el número de rezagos óptimos para el VEC debe ser $n - 1$. Esto debido a que a un VAR(p) es equivalente a un VEC(p-1).

Finalmente, la mayoría tiene un criterio flexible al escoger n rezagos tal que, muestren la existencia de cointegración y eventualmente, corregir problemas de correlación serial.

Así mismo, la prueba también es sensible a la especificación, y existen cuatro especificaciones generales que involucran la estimación con: constante, constante restringida, tendencia lineal, y tendencia restringida. Estas especificaciones, dependiendo del software, incluyen términos lineales, o constantes tanto en la ecuación de cointegración como en la ecuación de corrección del error. Los términos restringidos hacen referencia a cuando se incluyen en el corto plazo, mientras que los no restringidos se incluyen en la relación de largo plazo.

Es de destacar que la existencia o no de cointegración puede depender de la especificación, lo importante y sugerido, es que, si se detecta cointegración con una determinada especificación, esta se mantenga en el VEC a estimar, e igual el número de rezagos donde exista la cointegración.

Esto debido a que debemos tener cierta flexibilidad con la especificación del modelo, y la teoría económica, para detectar la presencia de cointegración, y también la resolución de problemas en los supuestos econométricos. Por ejemplo, si tenemos un problema de autocorrelación, tenemos que aumentar el número de rezagos en el VEC pese a que no sea lo indicado por los criterios de información del error, aquí debemos también volver a comprobar la cointegración y ver si se solucionó el problema o si persiste.

Por ahora procederemos con la aproximación más simple, que es la prueba de Johansen estándar.

```
varsoc x y
vecrank x y, lag(3)
```

Nótese que aquí el sugerido de óptimos de rezagos es 3.

```
. varsoc x y
```

Selection-order criteria
Sample: 5 - 100 Number of obs = 96

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-398.02				14.2677	8.33375	8.35534	8.38717
1	-285.601	224.84	4	0.000	1.49078	6.07501	6.1398	6.23529
2	-274.772	21.657	4	0.000	1.29329	5.93275	6.04073	6.19987*
3	-267.878	13.789*	4	0.008	1.21801*	5.87245*	6.02362*	6.24642
4	-264.303	7.1504	4	0.128	1.22956	5.8813	6.07566	6.36212

Endogenous: x y
Exogenous: _cons

Pese a que x y y sean no estacionarias, integradas de primer orden. No se puede ejecutar la prueba sin rezagos, esto daría como resultado:

```
. vecrank x y, lag(0)
lags() invalid -- invalid number, outside of allowed range
r(125);
```

El investigador debe basarse (pero no regirse exclusivamente) del conjunto de rezagos. Dado que esto es una guía, por lo que es importante observar la sensibilidad de la prueba de acuerdo con la teoría económica y la literatura econométrica.

En este caso, ejecutaremos la prueba con 3 rezagos ideales como lo dictan los criterios de información. (en el caso de tener problemas de autocorrelación, podríamos aumentar el número de rezagos).

```
. vecrank x y, lag(3)
```

Johansen tests for cointegration

Trend: constant Number of obs = 97
Sample: 4 - 100 Lags = 3

maximum				5%	
rank	parms	LL	eigenvalue	trace statistic	critical value
0	10	-286.93543	.	34.0302	15.41
1	13	-270.49485	0.28750	1.1490*	3.76
2	14	-269.92034	0.01178		

Aquí es importante entender que significa la prueba. La columna maximum rank, se refiere a la hipótesis nula del número de relaciones o ecuaciones de cointegración. El estadístico de prueba de Johansen, empieza inspeccionando la relación entre variables con 0 relaciones de cointegración, de ahí el 0 en la primera fila. Luego tenemos los eigenvalues, el trace statistic (que es el estadístico calculado) y el 5% critical value (que es el estadístico de tablas).

En una primera aproximación, siempre comparamos el trace statistic contra el estadístico del 5% de valor crítico. Si el trace statistic es mayor que el estadístico de valor crítico, rechazamos la hipótesis nula. Por ejemplo, para la hipótesis de 0 relaciones de cointegración, el trace statistic es de 34.03 y el valor crítico es de 15.41. Dado que el trace statistic es mayor que el valor crítico, rechazamos la hipótesis nula de 0 ecuaciones de cointegración y saltamos a la siguiente fila, donde probamos la siguiente hipótesis, la cual es que existe al menos una ecuación de cointegración. En este caso ($\text{rank}=1$) indica la hipótesis de que existe al menos una ecuación de cointegración, el trace statistic es de 1.14 mientras que el valor crítico es de 3.76. Dado que el trace statistic es menor que el valor crítico, no podemos rechazar la hipótesis de que existe al menos una ecuación de cointegración.

En este punto se invita al lector a observar como cambian los resultados de la cointegración, a través de las otras tres especificaciones de la prueba de cointegración.

Por ahora, la conclusión es que existe evidencia de cointegración de acuerdo con la prueba de Johansen entre las variables x_t y y_t .

V. Estimación del modelo VAR si las series son estacionarias I(0) No APLICA.

Dado que las series no son estacionarias, poseen cointegración, el VAR no aplica y se debe proseguir con VEC.

VI. Estimación del modelo VEC si las series son I(1) y cointegradas

Estamos en el contexto de series cointegradas, las cuales muestran propiedades de cointegración de acuerdo con la evidencia de la prueba de cointegración de Johansen en su especificación más sencilla. Por consiguiente, dado que usamos 3 rezagos para estimar la prueba, y encontramos la cointegración, en este contexto, también se utiliza tres rezagos para estimar el modelo VEC.

<code>vec x y, lags(3)</code>

```
. vec x y, lags(3) rank(1)
```

Vector error-correction model

Sample: 4 - 100

Number of obs = 97

AIC = 5.845255

HQIC = 5.984782

Det(Sigma_ml) = .9061589

SBIC = 6.190319

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_x	6	.960926	0.4024	61.28431	0.0000
D_y	6	1.05828	0.4481	73.89131	0.0000

Esta parte inicial, contiene la información del número de observaciones, criterios de información del error (AIC, HQIC, SBIC). Y también encontramos las pruebas de significancia sobre las ecuaciones de corto plazo (D_x como Δx_t y D_y como Δy_t). Observamos que el $R^2 = 0.4024$ pertenece a la ecuación D_x e implica una explicación lineal del modelo del 40.2% para explicar el comportamiento de la primera diferencia de x, el p valor es de 0.0000 lo cual significa que el conjunto de variables explicativas para esta ecuación es estadísticamente significativo en conjunto para explicar la primera diferencia de x.

En el caso de D_y se destaca que tiene una explicación del 44.81% en términos lineales dado el $R^2 = 0.4481$ para explicar el comportamiento de la primera diferencia de y y que el p valor es de 0.000, lo cual es menor que un nivel de significancia del 5%, y se rechaza la hipótesis nula de que los regresores en conjunto sean iguales a 0 para explicar el comportamiento de la primera diferencia de y.

La siguiente parte de la salida de la regresión es:

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
D_x	_ce1						
	L1.	-.5013295	.1285545	-3.90	0.000	-.7532918	-.2493673
	x						
	LD.	-.284212	.1233069	-2.30	0.021	-.5258892	-.0425349
	L2D.	-.1791009	.1056425	-1.70	0.090	-.3861564	.0279545
	y						
	LD.	-.2739854	.1196543	-2.29	0.022	-.5085036	-.0394672
	L2D.	-.1133959	.0927385	-1.22	0.221	-.2951601	.0683682
	_cons	.219534	.0997432	2.20	0.028	.0240409	.4150271
D_y	_ce1						
	L1.	.6185822	.1415787	4.37	0.000	.3410932	.8960713
	x						
	LD.	-.4832685	.1357994	-3.56	0.000	-.7494304	-.2171065
	L2D.	-.2865174	.1163453	-2.46	0.014	-.5145501	-.0584847
	y						
	LD.	-.2045153	.1317768	-1.55	0.121	-.4627931	.0537625
	L2D.	-.1805721	.1021341	-1.77	0.077	-.3807512	.0196069
	_cons	.1779212	.1098484	1.62	0.105	-.0373778	.3932201

La cual contiene las ecuaciones de corto plazo (separadas por un D_x, D_y). La ecuación tiene varios términos asociados, el primero al término de corrección del error (_ce1) con un coeficiente de -0.501, estadísticamente significativo al 5% dado su p valor de 0.000. Este significa para Δx_t , que existe una corrección anual del 50.1% en promedio para llegar a la relación de equilibrio entre x y y . El coeficiente, es negativo, su valor absoluto es menor que 1, y es estadísticamente significativo, por lo que existe una relación tendiente al equilibrio de largo plazo por parte del ajuste parcial de x_t .

En esta misma ecuación, tenemos también las diferencias rezagadas para x_t y para y_t , que son respectivamente los términos Δx_{t-p} y Δy_{t-p} en el VEC. Y representa coeficientes de ajuste parcial sobre los que se ejecuta la prueba de causalidad en el sentido de Granger. Finalmente, existe una constante de 0.21 que es estadísticamente significativa al 5%, y su significado es que existe un valor autónomo e independiente del ajuste parcial de la primera diferencia de x de 0.21.

Para los coeficientes de ajuste parcial, se interpretan de la siguiente forma: x (LD) significa lagged difference o diferencia rezagada, con un coeficiente de -.2842 con p valor 0.021. Quiere decir que cuando aumenta en una unidad la primera diferencia rezagada de x_t existe un decrecimiento de -0.28 unidades en el valor presente de la primera diferencia de x_t en promedio, ceteris paribus estadísticamente significativo al 5%. El resto de las diferencias rezagadas se interpretan bajo la misma lógica. Sin embargo, el aporte más valioso del modelo, se encuentran en las relaciones de cointegración.

Si quisiéramos probar la causalidad en el sentido de Granger en el corto plazo, de Δy_{t-i} hacia Δx_t se puede utilizar el siguiente comando:

```
test ([D_x]:LD.y L2D.y)
```

```
. test ([D_x]:LD.y L2D.y)

( 1)  [D_x]LD.y = 0
( 2)  [D_x]L2D.y = 0

      chi2( 2) =      5.34
Prob > chi2 =      0.0694
```

Este comando se hace dentro de un paréntesis () y se especifica cual ecuación queremos testear [D_x] o bien [D_y]. Nótese que el operador para seleccionar la ecuación esta dado por || con el respectivo nombre de corto plazo D_. Luego seleccionamos los coeficientes de testeo. En este caso los asociados a todas las variables de y que son la primera y segunda diferencia rezagada. La hipótesis es que el conjunto de coeficientes es igual a 0. La aceptación de la hipótesis nula implica que no hay causalidad en el sentido de Granger en conjunto de y hacia x en el corto plazo. Esto dado que el p-valor es mayor que 0.05 y no se puede rechazar la hipótesis nula de que los coeficientes en conjunto de y en sus primeras diferencias rezagadas puedan explicar el comportamiento de la primera diferencia de x .

Pasando a la siguiente ecuación denotada por D_y que representa Δy_t encontramos que existe un término de ajuste al equilibrio (_ce1) que es positivo y estadísticamente significativo al 5%. Este coeficiente tiene un valor de 0.61 e implica que para el proceso y_t **no** existe un ajuste parcial tendiente a un equilibrio, sino que, al contrario, parece un proceso divergente. Esto tiene interpretaciones muy importantes, y es que las relaciones de equilibrio no son totalmente estables en el sistema, y la razón de esto, es el proceso interno de la simulación de y_t dado que está definido (en nuestro apartado inicial) por una tendencia estocástica s y por un termino de error distribuido normalmente e . Contrario a x_t que depende conjunta de y_t , mientras que para y_t no existe tal dependencia, y por consiguiente al estimar un VEC, la relación de equilibrio es inestable.

Finalmente, se observa que la ecuación también contiene variables rezagadas tanto de x como y donde se ejecuta la prueba Granger.

En un sistema estable, todos los términos de ajuste parcial (_ce) deben ser negativos, y estadísticamente significativos menores en su valor absoluto que 1. Sin embargo, esto no se cumple para Δy_t . Indicando problemas de estabilidad en el largo plazo.

Finalmente la ultima parte de la salida de la regresión indica que:

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	1	495.4144	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_ce1					
x	1	.	.	.	.
y	-1.023337	.0459764	-22.26	0.000	-1.113449 -.9332255
_cons	.2158919	.	.	.	.

Tenemos la prueba de significancia estadística sobre la ecuación de cointegración estimada, el cual el p valor es 0.000 (chi2=495.41) por lo que se rechaza la hipótesis nula de que sea igual 0 la relación de cointegración.

En este punto tenemos lo más importante, la ecuación de cointegración normalizada. Expresada de la forma:

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_ce1					
x	1	.	.	.	.
y	-1.023337	.0459764	-22.26	0.000	-1.113449 -.9332255
_cons	.2158919	.	.	.	.

Esta ecuación está escrita de la forma:

$$x_t + 0.21 - 1.02y_t = e_t$$

Por lo que la interpretación de los signos debe invertirse de la forma:

$$x_t = -0.21 + 1.02y_t + e_t$$

Lo cual significa que, en el equilibrio de largo plazo, un incremento de una unidad en y_t se encuentra asociado a un incremento de 1.02 unidades en promedio sobre x_t , ceteris paribus estadísticamente significativo, por el $p>|z| = 0.00$ que es menor que un 5% de significancia. Varios autores eliminan el termino e_t al momento de escribir esta ecuación.

Algunos autores, interpretan el hecho de que los términos de ajuste al equilibrio de largo plazo (_ce) representan también la causalidad al largo plazo cuando cumplen las condiciones de ser negativos, estadísticamente significativos, y menores en valor absoluto que la unidad. Aunque hay mucha discusión de lo que es la causalidad en el análisis econométrico últimamente, se puede interpretar también como estabilidad en relaciones y co-movimientos de largo plazo.

Si quisiéramos estudiar la causalidad en el sentido de Granger, utilizaríamos pruebas de hipótesis sobre los coeficientes. Sin embargo, dado que esta causalidad es discutida, y se interpreta como propensiones, no ahondaremos en este apartado sobre esta discusión.

VII. Verificación de supuestos esenciales (no autocorrelación, estabilidad, normalidad)

Así como en el modelo VAR tenemos supuestos econométricos, el VEC también sigue los mismos. Se presentan primero los comandos y luego la interpretación de supuestos uno a uno.

```
vec x y, lags(3) rank(1)
```

```
veclmar
vecnorm
vecstable
```

```
. veclmar
```

```
Lagrange-multiplier test
```

lag	chi2	df	Prob > chi2
1	6.1808	4	0.18605
2	5.3994	4	0.24872

```
H0: no autocorrelation at lag order
```

La prueba de autocorrelación indica como hipótesis nula, la no autocorrelación en el rezago de orden p . En este caso, la prueba tiene dos rezagos (aunque el número de rezagos de prueba se pueden modificar). En este orden de ideas, dado que ambos p valores (0.18 y 0.24) son mayores que un 5% de significancia estadística, no se rechaza la hipótesis nula de que no hay autocorrelación en los ordenes de rezagos revisados.

Por consiguiente, el modelo cumple satisfactoriamente el supuesto de no autocorrelación serial.

Pasando a la prueba Jarque-Bera para inspeccionar la normalidad tenemos:

```
. vecnorm
```

```
Jarque-Bera test
```

Equation	chi2	df	Prob > chi2
D_x	3.526	2	0.17156
D_y	0.672	2	0.71459
ALL	4.198	4	0.37990

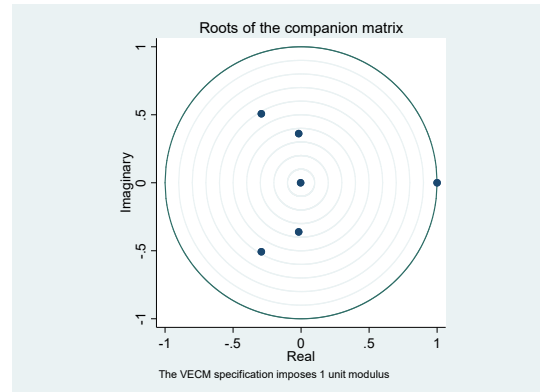
En este caso, se muestran las pruebas de normalidad para las ecuaciones y en general para el modelo. Nótese que todos los p valores son mayores que 5%, y la hipótesis nula es el comportamiento normal en los residuales de cada ecuación y en general del modelo. Esto significa que con un nivel de significancia del 5% no se puede rechazar la hipótesis de que los residuales siguen un comportamiento normal, el modelo cumple con el supuesto de normalidad.

```
. vecstable, graph
```

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
-.2922394 + .5073993i	.585541
-.2922394 - .5073993i	.585541
-.0176363 + .3614862i	.361916
-.0176363 - .3614862i	.361916
-.00332385	.003324

The VECM specification imposes a unit modulus.



En este caso, la estabilidad de parámetros se encuentra dentro del círculo unitario. Por lo que podemos asumir que en general, el modelo cumple con la estabilidad requerida. (Nótese que hay una notificación de que existe una unidad en el módulo, sin embargo, esto teóricamente no se ha logrado comprobar como aspecto de inestabilidad de modelos).

VIII. Interpretación de resultados

Podemos segmentar la interpretación de resultados de dos formas, una de corto y otra de largo plazo. Por ejemplo, en las ecuaciones de corto plazo (D_y , D_x) observamos que para cada proceso de ajuste parcial (denominado así por que la dependiente es una primera diferencia). Existen para Δx_t influencias tanto de sí misma, como de la primera diferencia rezagada de y_t para encaminar el ajuste parcial hacia una relación estable de equilibrio en el largo plazo (dado que se cumplieron las condiciones sobre el coeficiente de ajuste en el largo plazo $_ce$ que implican que su valor absoluto es menor que uno, es estadísticamente significativo, y negativo). Para el caso de la ecuación (D_y) que representa el ajuste parcial Δy encontramos que no existe una estabilidad de las relaciones hacia el largo plazo dado que el coeficiente ($_ce$) es positivo. Por consiguiente, existe un problema en la dinámica de convergencia de ambos procesos. Si interpretáramos los coeficientes de corto plazo de la ecuación (D_y) podríamos caer en un error, dado que el proceso no tiene convergencia, por lo que su interpretación se omite.

Ahora para el largo plazo, podemos destacar que existe un incremento de 1.02 unidades en x_t por cada unidad que aumenta en promedio en el largo plazo y_t , lo cual es estadísticamente significativo al 5%. Aquí nuevamente, aunque la ecuación (D_x) es estable

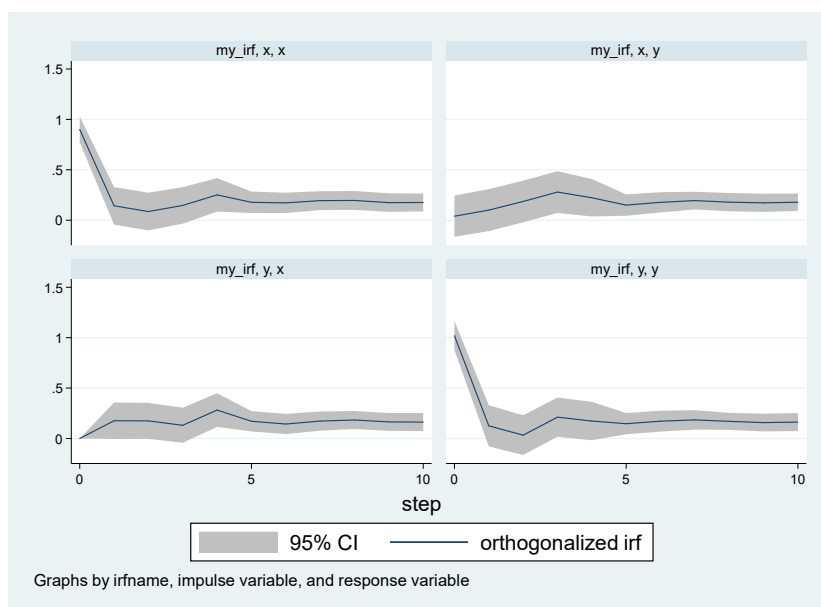
en el corto plazo, el sistema en general debe ser interpretado con cautela, puesto que la relación de (D_y) no tiene un coeficiente de ajuste a largo plazo ($_{ce}$) que cumpla con las condiciones ideales de estabilidad (dado que no es negativo). Por consiguiente, el sistema anterior es defectuoso y debe ser interpretado con cautela, incluso, pese a que se cumplan los supuestos econométricos.

IX. Análisis de impulso respuesta

Considerando que uno de los instrumentos de interpretación econométrica es la función de impulso respuesta, nos encontramos ante una dificultad de su estimación, dado que originalmente (y casi ningún software) viene diseñado para estimar directamente del VEC, la función de impulso respuesta. Esto genera una limitación para comprender la estructura de los impulsos en el sistema endógeno.

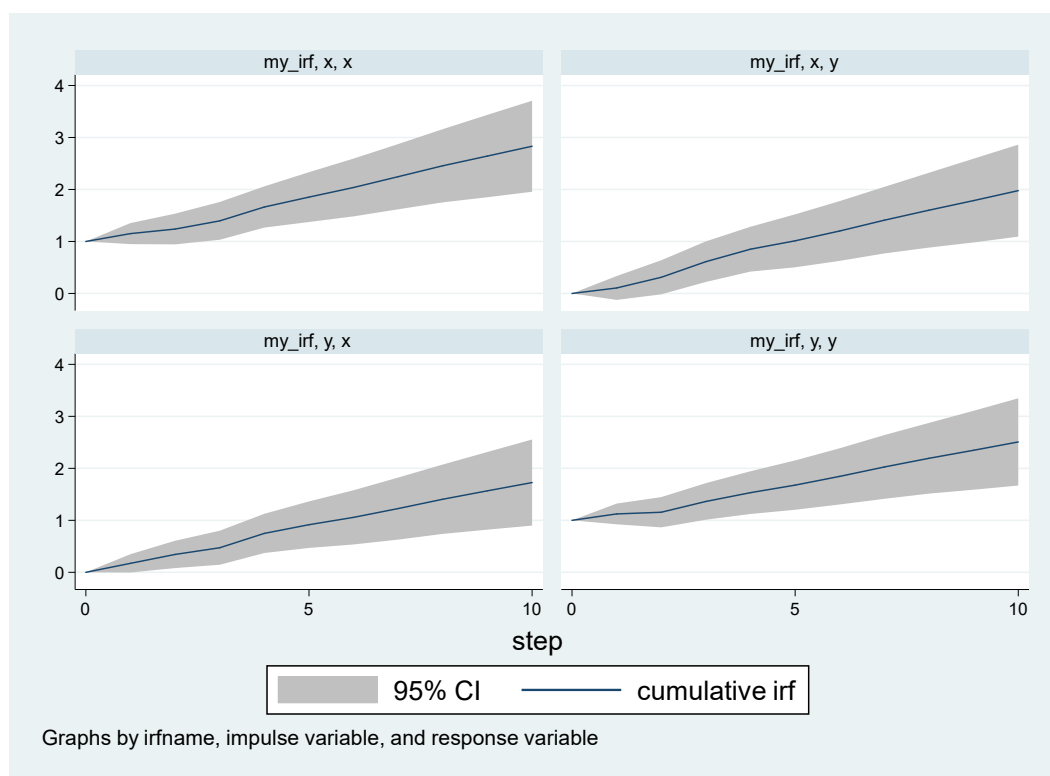
Dado que es posible demostrar que un VAR(p) es equivalente a un VEC(p-1). Seguiremos el enfoque de Mohr (2019) donde se pueden estimar las funciones de impulso respuesta agregando un rezago más a la estimación del VEC como transformación a VAR. En este contexto, solo basta con incluir la especificación del modelo y proseguir como se ha ejecutado el VAR en su estimación de la IRF con (p+1) rezagos.

```
var x y, lags(1/4)
irf set my_irf
irf create my_irf, step(10)
irf graph oirf
irf graph cirf
```



Se destaca que se ha creado la función de impulso respuesta entre x_t y y_t donde los impulsos de (x sobre x) y de (y sobre y) claramente evidencia un proceso autorregresivo, con decaimiento secuencial en el tiempo. Algo interesante al observar las IRF es que tanto los impulsos de x a y y de y a x aunque no son significativos al inicio del shock, parecen volverse significativos entre el periodo 4 y 5 en adelante, dado que los intervalos de confianza dejan de tocar la línea horizontal del 0. Esto sucede porque el sistema no trabaja sobre variables estacionarias, sino integradas de primer orden, sin embargo, recuérdese que en este caso es resultado de un sistema que no cumple condiciones de estabilidad en el largo plazo (por el coeficiente de ajuste de largo plazo positivo). Para mayor ejemplo de interpretación, revisar el apartado del VAR de las IRF.

Finalmente, como análisis adicional de esta aparente significancia estadística después del periodo 4 o 5. Conviene revisar si el sistema es incremental, acumulado o tendiente al crecimiento de forma acumulada. Esto se visualiza a continuación.



Efectivamente existen procesos incrementales que no parecen disminuir en este sistema de relaciones económicas. Lo anterior es producto de que trabajamos sobre procesos autorregresivos integrados de primer orden, sin embargo, es crucial destacar que con los resultados actuales tenemos un problema de inestabilidad del largo plazo reflejado en el coeficiente de ajuste a largo plazo ($_ce$) en $(D_y) \Delta y_t$ puesto que no tiene un coeficiente de ajuste al largo plazo que cumpla con las condiciones tradicionales (negativo, menor que 1 y significativo), lo que provoca que el sistema pueda ser erróneo en las interpretaciones de largo plazo.

Un grafico de impulsos acumulados, en un VEC estable, igualmente se ve con curvas incrementales debido a que trabajamos sobre un proceso integrado de primer orden.

X. Conclusión

El modelo estimado, posee problemas de estabilidad en el largo plazo dado que el coeficiente de ajuste de largo plazo ($_ce$) de la ecuación Δy_t es positivo. Por otro lado, aunque el proceso de corto plazo Δx_t cumple con las condiciones de estabilidad, el hecho de que la otra variable no los cumpla indica que hay problemas de estabilidad hacia el largo plazo entre las variables de análisis.

En este mismo contexto, al observar las funciones de impulsos respuesta (IRF) se destaca que existe un comportamiento que no se diluye a través del tiempo, entre las interacciones de las variables x y y .

Finalmente, la relación de largo plazo indica que por cada una unidad que aumente y en el largo plazo, existe un incremento de 1.02 unidades en x_t estadísticamente significativo ceteris paribus. Sin embargo, este resultado no es confiable y debe ser interpretado con cautela dado que el sistema es inestable por las razones antes mencionadas.

La totalidad del código anterior se presenta a continuación:

```
clear
set seed 123
set obs 100

gen t = _n // Generar un índice de tiempo

// Generate a common stochastic trend
gen s = 0
gen e_s = rnormal()

forvalues i = 2/100 {
    replace s = s[_n-1] + e_s in `i' // Stochastic trend
}

// Generate two cointegrated series around the trend
gen y = 0.1*s + 0.1*t + rnormal()
gen x = 0.1*s + 0.1*t + 0.01*y + rnormal() // Cointegration relationship
tsset t

// Graficar las series
twoway (line x t, lcolor(blue)) ///
      (line y t, lcolor(red)), ///
      title("Series x y") ///
      legend(label(1 "x") label(2 "y"))

// Graficar las series
twoway (line x t, lcolor(blue)) ///
      (line y t, lcolor(red)), ///
      title("Series Cointegradas") ///
      legend(label(1 "x") label(2 "y"))

gen dx = x - L.x
gen dy = y - L.y
```

Metodología estándar de vectores autoregresivos (VAR) y de corrección del error (VEC)

```
twoway (line dx t, lcolor(blue)) ///  
      (line dy t, lcolor(red)), ///  
      title("Primeras diferencias") ///  
      legend(label(1 "dx") label(2 "dy"))
```

* x

```
varsoc x  
dfuller x, lag(1)  
varsoc dx  
dfuller dx, lag(0)
```

* y

```
varsoc y  
dfuller y, lag(4)  
varsoc dy  
dfuller dy, lag(3)
```

```
varsoc x y  
vecrank x y, lag(3)
```

```
vec x y, lags(3) rank(1)  
test ([D_x]:LD.y L2D.y)
```

```
vechmar  
vecnorm  
vecstable, graph
```

* IRF

```
var x y, lags(1/4)  
irf set my_irf  
irf create my_irf, step(10)  
irf graph oirf  
irf graph cirf
```

6. Referencias

- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251–276.
<https://doi.org/10.2307/1913236>
- Hamilton, J. D. (1994). *Time Series Analysis / Princeton University Press* (1st ed.).
Princeton New Jersey.
<https://press.princeton.edu/books/hardcover/9780691042893/time-series-analysis>
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2), 231–254. [https://doi.org/10.1016/0165-1889\(88\)90041-3](https://doi.org/10.1016/0165-1889(88)90041-3)
- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59(6), 1551–1580.
<https://doi.org/10.2307/2938278>
- Johansen, S., & Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics*, 52(2), 169–210. <https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x>
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
<https://doi.org/10.1007/978-3-540-27752-1>
- Mohr, F. (2019). *An Introduction to Vector Error Correction Models (VECMs)*. *r-econometrics*. <https://www.r-econometrics.com/timeseries/vecintro/>

Pantula, S. G. (1989). Testing for Unit Roots in Time Series Data. *Econometric Theory*, 5(2), 256–271.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1), 1–48.

<https://doi.org/10.2307/1912017>

Sims, C. A. (1986). *Are Forecasting Models Usable for Policy Analysis?* / *Federal Reserve Bank of Minneapolis*. <https://www.minneapolisfed.org/research/quarterly-review/are-forecasting-models-usable-for-policy-analysis>

Stock, J. H., & Watson, M. W. (2020). *Introduction to econometrics* (Fourth, global edition.). Pearson.