

Vectores Autorregresivos

Francisco Corona

fcoronav@gmail.com

franciscoj.corona@inegi.org.mx

Septiembre 29, 2018.

Objetivo

El objetivo de esta clase será denotar algunas generalidades de los modelos VAR como:

- ▶ Vectores Autorregresivos como modelos lineales: estimación.
- ▶ Características y detalles a tomar en cuenta en la especificación y estimación: selección del rezago óptimo, correlación serial y efecto ARCH multivariado.
- ▶ Utilización: causalidad de Granger, pronósticos y funciones de respuesta-impulso.
- ▶ Aplicaciones.

VAR como modelos lineales: estimación

Vectores Autorregresivos

La extensión natural de modelos ARMA desde una perspectiva multivariada son los siguientes modelos:

- ▶ Vectores Autorregresivos (VAR).
- ▶ Vector de Medias Móviles (VMA).
- ▶ Vector Autorregresivos de Medias Móviles (VARMA).

Vectores Autorregresivos

Los **Vectores Autorregresivos o modelos VAR** fueron popularizados por Chris Sims (Econometrica, 1980) y son muy utilizados en diversas aplicaciones en el área de economía dada su flexibilidad en términos estadísticos.

Vectores Autorregresivos

Sea

$$\{y_t = (y_{1t}, \dots, y_{Kt})'; t \in \mathbb{Z}\},$$

un vector de K variables **estacionarias** indexadas en el tiempo.

Vectores Autorregresivos

Decimos que el vector $\{y_t; t \in \mathbb{Z}\}$, es un proceso autorregresivo de orden p , denotado como $VAR(p)$ si

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}$$

- ▶ p es un entero positivo.
- ▶ A_i son matrices de coeficientes $(K \times K)$.
- ▶ $\nu = (\nu_1, \dots, \nu_K)'$ es el intercepto de dimensión $(K \times 1)$.
- ▶ $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Kt})'$ es un vector K -dimensional de ruidos blancos con matriz de covarianza no-singular Σ_ε .

Ejemplo: $VAR(1)$ bivariado

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

o

$$y_{1t} = \nu_1 + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t},$$

$$y_{2t} = \nu_2 + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t},$$

donde $cov(\varepsilon_{1t}, \varepsilon_{1s}) = \sigma_{12}$ para $t = s$, 0 en caso contrario.

Ventajas de los modelos VAR

- ▶ Son fáciles de estimar.
- ▶ Suelen dar pronósticos más certeros que modelos univariados.
- ▶ Los investigadores no requieren especificar qué variables son endógenas o exógenas dado que todas las variables son endógenas.
- ▶ En un sistema VAR es fácil realizar pruebas de causalidad en el sentido de Granger.

Desventajas de los modelos VAR

- ▶ Se estiman demasiados parámetros. Si existen K ecuaciones, y en cada ecuación hay p rezagos, se tienen que estimar $K(1 + p^2)$ parámetros.
- ▶ Pueden resultar de carácter no teórico, es decir, pueden ser modelos más de una naturaleza más estadística que econométrica. Este tipo de modelos suelen no ser utilizados para realizar política económica, aunque no es una generalidad.

Estimación

Dentro del contexto de regresión lineal, considere el siguiente modelo $VAR(1)$

$$y_{1t} = \nu_1 + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t},$$

$$y_{2t} = \nu_2 + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t},$$

donde $cov(\varepsilon_{1t}, \varepsilon_{1s}) = \sigma_{12}$ para $t = s$, 0 en caso contrario.

Estimación

- ▶ El modelo corresponde a realizar dos regresiones con diferente variable dependiente pero con las mismas variables explicativas.
- ▶ Podríamos utilizar OLS y estimar los parámetros por separado para cada ecuación.

Estimación

Centrémonos en la primera ecuación

$y_{1t} = \nu_1 + a_{11}y_{1t-1} + a_{12}y_{2t-1} + \varepsilon_{1t}$, para $t = 1, \dots, T$, de tal forma que

$$y_{11} = \nu_1 + a_{11}y_{10} + a_{12}y_{20} + \varepsilon_{11},$$

$$y_{12} = \nu_1 + a_{11}y_{11} + a_{12}y_{21} + \varepsilon_{12},$$

$$\vdots$$

$$y_{1T} = \nu_1 + a_{11}y_{1T-1} + a_{12}y_{2T-1} + \varepsilon_{1T}.$$

Estimación

Definimos

$$\mathbf{y}_1 = [y_{11}, \dots, y_{1T}]',$$

$$\mathbf{X}_1 = \begin{bmatrix} 1 & y_{10} & y_{20} \\ 1 & y_{11} & y_{21} \\ \vdots & \vdots & \vdots \\ 1 & y_{1T-1} & y_{2T-1} \end{bmatrix},$$

$$\boldsymbol{\pi}_1 = [\mu_1, a_{11}, a_{12}]',$$

$$\boldsymbol{\varepsilon}_1 = [\varepsilon_{11}, \dots, \varepsilon_{1T}].$$

Estimación

Por lo tanto

$$\mathbf{y}_1 = \mathbf{X}\boldsymbol{\pi}_1 + \boldsymbol{\varepsilon}_1.$$

Por lo que el estimador $\boldsymbol{\pi}_1$ está dado por

$$\hat{\boldsymbol{\pi}}_1 = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}_1.$$

Estimación

De igual forma, consideremos la segunda ecuación,

$y_{2t} = \nu_2 + a_{21}y_{1t-1} + a_{22}y_{2t-1} + \varepsilon_{2t}$, para $t = 1, \dots, T$, entonces

$$y_{21} = \nu_2 + a_{21}y_{10} + a_{22}y_{20} + \varepsilon_{21},$$

$$y_{22} = \nu_2 + a_{21}y_{11} + a_{22}y_{21} + \varepsilon_{22},$$

$$\vdots$$

$$y_{2T} = \nu_2 + a_{21}y_{1T-1} + a_{22}y_{2T-1} + \varepsilon_{2T}.$$

Estimación

En este caso

$$\mathbf{y}_2 = [y_{21}, \dots, y_{2T}]',$$

$$\mathbf{X}_2 = \begin{bmatrix} 1 & y_{10} & y_{20} \\ 1 & y_{11} & y_{21} \\ \vdots & \vdots & \vdots \\ 1 & y_{1T-1} & y_{2T-1} \end{bmatrix},$$

$$\boldsymbol{\pi}_2 = [\mu_2, a_{21}, a_{22}]',$$

$$\boldsymbol{\varepsilon}_2 = [\varepsilon_{21}, \dots, \varepsilon_{2T}].$$

Estimación

Entonces

$$\mathbf{y}_2 = \mathbf{X}\boldsymbol{\pi}_2 + \boldsymbol{\varepsilon}_2.$$

Por lo que el estimador $\boldsymbol{\pi}_2$ está dado por

$$\hat{\boldsymbol{\pi}}_2 = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}_2.$$

Estimadores OLS

Por lo tanto, los estimadores OLS son

$$\hat{\pi}_1 = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}_1 \Rightarrow \pi_1 = [\nu_1, a_{11}, a_{12}]',$$

$$\hat{\pi}_2 = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}_2 \Rightarrow \pi_2 = [\nu_2, a_{21}, a_{22}]'.$$

Estimadores OLS

¿Son los estimadores OLS de π_1 y π_2 eficientes?

Estimadores OLS

- ▶ Recordamos que cuando $Cov(\varepsilon_{1t}, \varepsilon_{2t}) = \sigma_{12} \neq 0$, es decir, cuando hay correlaciones contemporáneas, una estimación eficiente es incorporar dicha estructura de correlación.
- ▶ Es decir, una estimación apropiada sería utilizar GLS en lugar de OLS.

Estimadores OLS

Consideremos las siguientes ecuaciones

$$y_{1t} = \beta_{10} + \beta_{11}x_{1t} + \beta_{12}z_{1t} + \varepsilon_{1t},$$

$$y_{2t} = \beta_{20} + \beta_{21}x_{2t} + \beta_{22}z_{2t} + \varepsilon_{2t},$$

O en su forma matricial

$$\mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1,$$

$$\mathbf{y}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2.$$

Estimadores OLS

Donde asumimos

- ▶ $E(\varepsilon_{it}) = 0 \quad i = 1, 2 \quad t = 1, \dots, T.$
- ▶ $var(\varepsilon_{it}) = \sigma_i^2 \quad i = 1, 2. \quad t = 1, \dots, T.$
- ▶ $E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}^2 \quad i = 1, 2.$
- ▶ $E(\varepsilon_{it}\varepsilon_{js}) = 0 \quad i = 1, 2 \quad \text{para } t \neq s.$

Estimadores OLS

Note que podemos describir las ecuaciones anteriores como

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 \\ 0 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix},$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Nótese que podríamos estimar este modelo a través de OLS pero dado la estructura de covarianza, no obtendríamos una estimación eficiente.

Estimadores OLS

Notemos que la matriz de covarianza de ε viene dada por

$$\Omega = \begin{bmatrix} \sigma_1^2 \mathbf{I} & \sigma_{12}^2 \mathbf{I} \\ \sigma_{12} \mathbf{I} & \sigma_2^2 \mathbf{I} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \otimes \mathbf{I}_T = \Sigma \otimes \mathbf{I}_T.$$

En este caso, el mejor estimador lineal insesgado viene dado por

$$\beta_{GLS} = (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1} \mathbf{X}'\Omega^{-1}\mathbf{y} = [\mathbf{X}'(\Sigma^{-1} \otimes \mathbf{I}_T)\mathbf{X}]^{-1} \mathbf{X}'(\Sigma^{-1} \otimes \mathbf{I}_T)\mathbf{y}.$$

Que tiene menor varianza que el estimador OLS.

Estimadores OLS

Obviamente, lo anterior siempre será cierto al menos que:

- ▶ Que las covarianzas contemporáneas sean 0, es decir $\sigma_{12} = 0$
- ▶ Que las variables independientes sean iguales, es decir $\mathbf{X}_1 = \mathbf{X}_2$.

En este último caso, el estimador consistente de Σ viene dado por

$$\hat{\sigma}_{ij} = \frac{\hat{\varepsilon}_i' \hat{\varepsilon}_j}{T - 2p - 1} \quad \text{para } i, j = 1, 2$$

donde $\hat{\varepsilon}_i = \mathbf{y}_i - \mathbf{X}\hat{\pi}_i$.

Estimadores OLS

En general, para un $VAR(p)$ de dimensión K

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t,$$

sin pérdida de generalidad, puede ser estimado realizando la regresión de y_{it} con una constante más

$y_{1t-1}, \dots, y_{Kt-1}, \dots, y_{1t-p}, \dots, y_{Kt-p}$ para $j = 1, \dots, K$.

Adicionalmente, se puede incluir variables exógenas y/o variables determinísticas y realizar la estimación ecuación por ecuación.

Nótemos que podemos realizar la inferencia como un modelo clásico de regresión lineal. **Ojo: autocorrelación y series de tiempo NO-ESTACIONARIAS.**

Características y detalles a tomar en cuenta en la
especificación y estimación

Características

Desde la perspectiva de series de tiempo, podemos utilizar resultados de modelos AR para analizar la estabilidad del modelo, la cual es importante para concluir que el **modelo VAR sea estacionario**. Notemos que el modelo $VAR(p)$ puede ser rescrito como un modelo $VAR(1)$

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \nu \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_K & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$Y_t = \tilde{\nu} + \tilde{A}Y_{t-1} + \tilde{\varepsilon}_t.$$

Características

En este caso, y análogo a modelos univariados, Y_t será estable si

$$\det(\mathbf{I}_{Kp} - z\tilde{A}) \neq 0 \quad \text{para } |z| \leq 1.$$

Es decir, si todas las raíces del polinomio característico anterior son menores a 1. Por otra parte, el vector de medias viene dado por

$$\mu = E(Y_t) = (\mathbf{I}_{Kp} - \tilde{A})^{-1}\nu,$$

y la función de autocovarianzas

$$\Gamma(h) = \sum_{i=1}^{\infty} \tilde{A}^{i+h} \hat{\Sigma}(\tilde{A}^i)'$$

Selección del rezago óptimo

La forma funcional para decidir el número de rezagos viene dada por la manera tradicional de especificar criterios de penalización al incluir más parámetros en modelos de regresión de la forma

$$IC(p) = \ln|\hat{\Sigma}(p)| + C_T \varphi(K.p),$$

donde $\hat{\Sigma}(p) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t' \hat{\varepsilon}_t$, C_T , el factor de penalización por el número de observaciones y $\varphi(K.p)$, la penalización por introducir más regresores que minimizan la varianza residual. En la práctica se suelen utilizar los criterios AIC, SC y HQ, cuyas penalizaciones son respectivamente, $(2/T)pK^2$, $(\ln T/T)pK^2$ y $(2\ln \ln T/T)pK^2$.

Prueba de autocorrelación

Al igual que en los modelos AR, es necesario que los errores no estén autocorrelacionados, lo que nos otorgaría intervalos de confianza inválidos para interpretar los t -estadísticos asociados al modelo de regresión. En este caso, podemos utilizar la prueba Portmanteau para verificar si los residuales del modelo $VAR(p)$ están libre de correlación serial. Dicha prueba se especifica como

$$Q_h = T \sum_{j=1}^h \text{tr}(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}),$$

donde $\hat{C}_i = \frac{1}{T} \sum_{t=i+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-i}'$. Este estadístico está distribuido aproximadamente como una $\chi^2(K^2(h-p))$. Hipótesis nula denota la NO presencia de errores autocorrelacionados.

Heteroscedasticidad condicional

El equivalente a errores heteroscedásticos en el caso de corte transversal está relacionado al efecto ARCH en series de tiempo, que son varianzas cambiantes y dependientes a través del tiempo. Sus efectos son similares que al de autocorrelación y en este caso, una prueba para verificar si dicho efecto está presente viene dado por

$$VARCH_{LM}(q) = \frac{1}{2} TK(K+1)R_m^2,$$

donde $R_m^2 = 1 - \frac{2}{K(K+1)} tr(\hat{\Omega}\hat{\Omega}_0^{-1})$ y $\hat{\Omega}$ es la matriz de covarianza del modelo de regresión. Este estadístico está distribuido como $\chi^2(qK^2(K+1)^2/4)$ donde hipótesis nula NO indica la presencia del efecto ARCH.

Utilización: causalidad de Granger, pronósticos y funciones de respuesta-impulso

Causalidad en el sentido de Granger

Una prueba tradicional en modelos VAR es la causalidad en el sentido de Granger, que básicamente indica bajo hipótesis nula si el pasado de una variable, digamos, y_{jt} no tiene un impacto en otra, y_{it} . En este caso, se pueden utilizar pruebas F de significancia conjunta o pruebas Wald de restricciones múltiples en los modelos de regresión lineal.

Causalidad en el sentido de Granger

Para un modelo $VAR(p)$ bivariado, si queremos probar que y_{2t} no es causal en el sentido de Granger tendríamos entonces

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} + \begin{bmatrix} a_{11}^1 & 0 \\ a_{21}^1 & a_{22}^1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \cdots + \begin{bmatrix} a_{11}^p & 0 \\ a_{21}^p & a_{22}^p \end{bmatrix} \begin{bmatrix} y_{1t-p} \\ y_{2t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Lo anterior es equivalente a probar si los coeficientes A_i son matrices triangulares inferiores. Se puede generalizar esta idea para modelos $VAR(p)$ de dimensión K .

Pronósticos

Para pronosticar valores futuros de y_t , podemos utilizar la “cadena de pronóstico” como sigue

$$y_{T+h|T} = A_1 y_{T+h-1|T} + \cdots + A_p y_{T+h-p|T},$$

donde $y_{T+j|T} = y_{T+j}$ para $j \leq 0$. Se puede mostrar que el h error de pronóstico viene dado por

$$y_{t+h} - y_{T+h|T} = \sum_{s=0}^{h-1} \psi_s \varepsilon_{T+h-s},$$

Pronósticos

donde las matrices Ψ_s son determinadas por recursión de la siguiente manera

$$\Psi_s = \sum_{j=1}^{p-1} \Psi_{s-j} A_j,$$

con $\Psi_0 = I_K$ y $A_j = 0$ para $j > p$. Entonces, la matriz de covarianza del error de pronóstico viene dada por

$$\Sigma(h) = MSE(y_{t+h} - y_{t+h|T}) = \sum_{s=0}^{h-1} \Psi_s \Sigma \Psi_s'.$$

Pronósticos

Puede mostrarse que el mejor predictor de y_{t+h} es $y_{t+h|T}$, el cual se obtiene al estimar \hat{A}_j a través de OLS de tal forma que, la estimación de $\Sigma(h)$ viene dada por

$$\hat{\Sigma}(h) = \sum_{s=0}^{h-1} \hat{\Psi}_s \hat{\Sigma} \hat{\Psi}_s',$$

donde $\hat{\Psi}_s = \sum_{j=1}^{p-1} \hat{\Psi}_{s-j} \hat{A}_j$. Finalmente, el intervalo de confianza asintótico de confianza $(1 - \alpha)100\%$ viene dado por

$$\left[y_{iT+h|T} - c_{1-\alpha/2} \hat{\sigma}_i(h), y_{iT+h|T} + c_{1-\alpha/2} \hat{\sigma}_i(h) \right],$$

donde $c_{1-\alpha/2}$ es el $(1 - \alpha/2)$ cuantil de la distribución normal y $\hat{\sigma}_i(h)$ el i ésimo elemento de la diagonal de $\hat{\Sigma}(h)$.

Funciones de respuesta-impulso

Se sabe que cualquier proceso (estacionario) $VAR(p)$ puede escribirse en su forma $MA(\infty)$ de la siguiente manera

$$y_t = \mu + \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \dots$$

De tal forma que

$$\frac{\partial y_{it+s}}{\partial \varepsilon_{jt}} = \frac{\partial y_{it}}{\partial \varepsilon_{jt+s}} = \Psi_{ij}^s, \quad i, j = 1, \dots, K,$$

donde Ψ_{ij}^s es el i, j elemento de la matriz Ψ_s , siendo éstas las **las funciones de respuesta-impulso no ortogonalizadas**.

Funciones de respuesta-impulso

- ▶ Las funciones de respuesta-impulso son complicadas de entender, pero representan el cambio de y_{it} ante las variaciones en el tiempo de los shocks de ε_{jt} . Suelen utilizarse para el impacto temporal que tiene un shock aleatorio de alguna variable económica en otra, esperando que la estabilización se alcance en cierto s periodo.
- ▶ Los coeficientes de respuesta-impulso suelen estimarse de manera ortogonalizada.

Aplicaciones

Experimento Monte Carlo

Generamos un $VAR(2)$ bidimensional con coeficientes $U(-0.5, 0.5)$ y computamos lo siguiente a través de 500 réplicas:

- ▶ Máxima raíz del polinomio característico $\det(\mathbf{I}_{Kp} - z\tilde{A})$
- ▶ Determinación del número de rezagos a través de los criterios de AIC, SC y HQ.
- ▶ Valor p de la prueba de correlación serial y efecto ARCH.
- ▶ Valor p de pruebas de causalidad de Granger para las dos variables.

Resultados

```
remove(list = ls())

# paqueterías
library(vars)
library(portes)
library(tseries)

set.seed(1234)

# parámetros
T <- 100
C <- 100
K <- 2
p <- 2
R <- 500
alpha <- 0.05

# simulamos algunos coeficientes
Phi <- runif(K*K*p, -0.5, 0.5)

# coeficientes VAR(2)
Phi <- array(Phi, dim = c(K, K, p))

est <- c("maxroot", "AIC", "SC", "HQ", "serial", "arch",
        paste("Series", 1:p, sep = ""))
resultados <- matrix(0, R, length(est))
colnames(resultados) <- est
```

Resultados

```

for(i in 1 : R){ # print(i)
  # simulaciOn
  Y <- varima.sim(model=(list(ar = Phi)), n = T + C, k = K,
    innov.dist = "Gaussian")
  # quitamos las primeras series
  Y <- Y[-(1:C),]
  # rezago
  k <- VARselect(Y, type = "none")
  resultados[i,c("AIC", "SC", "HQ")] <- k$sselection[c("AIC(n)", "SC(n)",
    "HQ(n)")]

  # estimaciOn
  var1 <- VAR(Y, p = p, type = "none")

  # estacionariedad
  resultados[i, "maxroot"] <- max(roots(var1))

  # pruebas de residuales
  resultados[i,"serial"] <- serial.test(var1)$serial$p.value
  resultados[i,"arch"] <- arch.test(var1)$arch$p.value
  # causalidad
  for(j in 1 : p){
    resultados[i,paste("Series",j, sep = "")] <- c(causality(var1,
      cause = paste("Series.",j,
        sep = ""))$Granger$p.value)
  }
}

```

Resultados

```
> # observamos los resultados
> sum(resultados[, "maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.848 0.970 0.972
> colSums(resultados[, c("serial", "arch")] > alpha)/R
serial  arch
  0.99   0.97
> colSums(resultados[, paste("Series", 1:p, sep = "")] < alpha)/R
Series1 Series2
  0.306   0.994
```

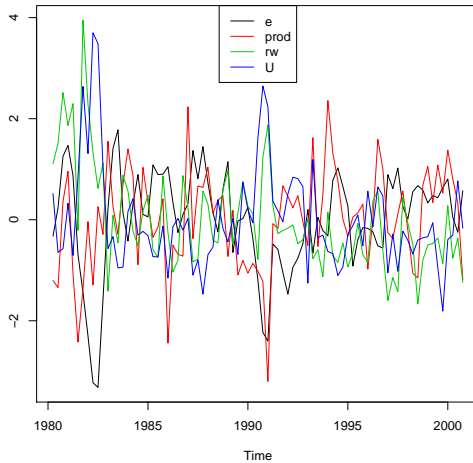
Ejercicio empírico

Con la base de datos de series trimestrales de Canadá (1980:I-2000:IV) de la librería `vars`, que incluye las series de productividad laboral (`prod`), empleo (`e`), tasa de desempleo (`u`) y salarios reales (`rw`). Especificaremos un modelo VAR y realizaremos pruebas de causalidad, pronósticos y las funciones de respuesta-impulso.

Ejercicio empírico

```
> # @ ejemplo aplicado con la base de datos de Canada
> data(Canada)
>
> # en logaritmos
> lcan <- log(Canada)
>
> # trabajamos desde el punto de vista estacionario, realizamos pruebas de raíces
> # unitarias
>
> adf.p <- matrix(0, ncol(lcan), 2)
> rownames(adf.p) <- colnames(lcan)
> colnames(adf.p) <- c("l", "fd")
>
> for(i in 1 : ncol(Canada)){
+   adf.p[i, "l"] <- adf.test(lcan[,i])$p.value
+   adf.p[i, "fd"] <- adf.test(diff(lcan[,i]))$p.value
+ }
>
> adf.p
           l           fd
e    0.5146190 0.07921231
prod 0.2062697 0.08800657
rw    0.5033446 0.08359261
U     0.3655768 0.03196395

# al menos parecen tener raiz unitaria al 10%
```

Ejercicio empírico

Determinamos en número de rezagos óptimo y obtenemos $p = 1$, estimamos el $VAR(1)$ y podemos ver que los R^2 ajustados son 0.5622 para la ecuación de empleo, 0.1471 para productividad laboral, 0.2493 para los salarios reales y 0.4645 para el desempleo. Las raíces del polinomio característico son 0.7386, 0.4451, 0.1663 y 0.1512.

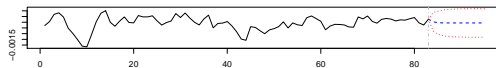
No parecen existir problemas ni de correlación serial ni de efecto ARCH, donde obtenemos valores p de 0.989 y 0.098.

Ejercicio empírico

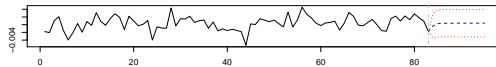
```
> # causalidad de Granger
> cau <- matrix(0, ncol(dlcan), 1)
> colnames(cau) <- "p.value"
> rownames(cau) <- colnames(dlcan)
>
> for(i in 1 : ncol(dlcan))
+   cau[i,] <- causality(varc, colnames(dlcan)[i])$Granger$p.value
>
> cau <- round(cau, 4)
> cau
      p.value
e      0.0000
prod   0.0078
rw     0.4385
U      0.0212

# sólo la no causal en el sentido de Granger parecieran ser los salarios reales
```

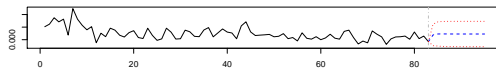
Forecast of series e



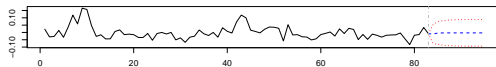
Forecast of series prod

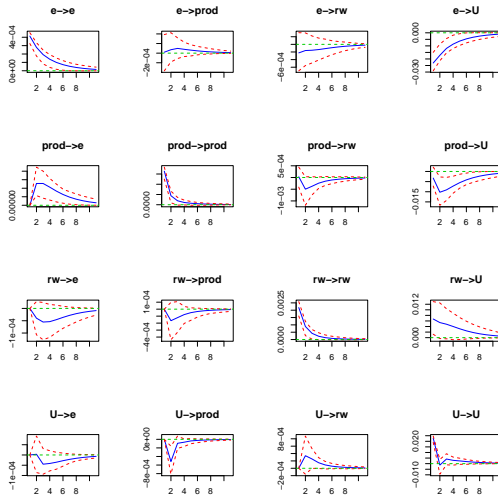


Forecast of series rw



Forecast of series U





Tarea: entregar hasta el 8 de octubre de 2018

1. (2 pts) Realizar el ejercicio 3 del capítulo 10 de Chan.
2. (1 pts) Muestra el procedimiento de cómo presentar las funciones de respuesta-impulso de manera ortogonalizada.
3. (3.5 pts) Modifica el experimento Monte Carlo de tal forma que observes el funcionamiento muestral de los diferentes estadísticos que se analizaron en clase. Modifica a tu interés los siguientes parámetros: coeficientes del modelo VAR, rezagos, estructuras de covarianza en los errores y tamaño de muestras y comenta lo hallado. Para cada uno de ellos prueba dos diferentes casos. *Idea de trabajo final, nota que estamos analizando el funcionamiento de diferentes pruebas bajo stress.*
4. (3.5 pts) Construye un modelo VAR cuyo objetivo sea analizar y pronosticar la inflación interanual de México para el siguiente periodo no disponible. Elige un periodo de muestra que consideres apropiado y realiza las pruebas que consideres necesarias para verificar que el modelo pronosticará bien a futuro. Puedes utilizar variables relacionadas endógenamente (según la teoría económica) como oferta de billetas y monedas (M_0), tipo de cambio nominal, tasa de desempleo y salarios reales. Alguna variable de tu sugerencia es bienvenida. Al final del curso, veremos quien fue el más certero. *Un problema tipo examen dado que abarca completamente el tema raíces unitarias y modelos VAR.*