

# Var PLS

Graciela González Farias

Centro de Investigación en Matemáticas.  
Monterrey, NL.  
Mayo de 2012

- Motivación
- PLSAR
- Revisión breve de modelos VAR y PLS
- Definición de VAR-PLS
- Bootstrap para VAR-PLS
- Aplicaciones
- Conclusiones

# Motivación

- PLS es una técnica que ha mostrado su utilidad en muchas áreas de aplicación, tales como el control de procesos en la industria química, producción por *batches*, imágenes médicas, donde se introducen modelos espacio-temporales PLS, Path Modeling, Clasificación, microarreglos, solo por mencionar algunos (ver los trabajos de McGregor, Nomikos, MacIntosh, Esposito Vinz, Paul Garthwaite, entre otros).
- El método puede aplicarse a datos univariados y multivariados.
- PLS ha mostrado mejor capacidad predictiva que otros métodos, incluso cuando no se cumplen totalmente los supuestos estándar. (SI ES ESTO LO QUE QUERIAS DECIR???)
- Philip Hans Franses (2006), propuso una metodología para realizar pronósticos  $h$  pasos adelante de manera óptima a través de una representación autorregresiva de orden  $p$ . El método es llamado Autorregresive Partial Least Squares, y lo denotamos mediante  $PLSAR(h, p)$

# Nuestro caso de interés

- Desarrollar un modelo para predecir la inflación en México.
- El modelo considerará el crecimiento o variación de las condiciones monetarias del país como fuente principal de la dinámica inflacionaria.
- Existe una gran discusión, incluso hoy en día, sobre si existe una relación de largo plazo entre el fenómeno monetario y el traspaso inflacionario que tiene
- No obstante, hay un consenso en que la inflación, en el largo plazo, es un fenómeno netamente monetario.
- En este trabajo no abordaremos tal discusión, sino que mostraremos las propiedades empíricas del modelo desarrollado en términos del error de predicción fuera de muestra.

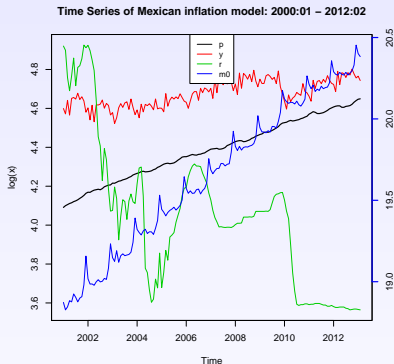
# Nuestro caso de interés

Utilizaremos las siguientes variables en el periodo comprendido entre enero de 2000 a febrero de 2012:

- $p$ : el índice nacional de precios al consumidor
- $m0$ : billetes y monedas en circulación
- $r$ : la tasa de interés interbancaria a 28 días
- $y$ : el indicador global de la actividad económica

El objetivo es relacionar empíricamente la variable de precios, que a su vez es una función de la tasa de inflación (mensual, interanual, acumulada, etcétera) con el resto de las variables, permitiendo relaciones multivariadas, generando así un  $VAR(p)$  de rango completo y/o un VECM para el caso cointegrado.

# Nuestro caso de interés



Podemos apreciar una tendencia creciente en niveles. El índice monetario presenta una estacionalidad característica en todo el periodo de tiempo, la producción económica con tendencia de menor pronunciación que la serie de precios, que a partir del 2010 muestra cierta recuperación respecto a los niveles observados en la primera mitad de la gráfica. La tasa de interés claramente ha tenido un periodo de estabilidad a partir del segundo

- Nosotros generalizamos la propuesta de Franses para lo que llamaremos  $VAR - PLS(h, p)$ , y se aplicará al modelo en cuestión, aunque solo se mostrarán resultados comparativos. Esta generalización comprende lo siguiente:
  - 1 Extensión al caso multivariado aprovechando la flexibilidad de los modelos de Vectores Autoregresivos (VAR), generando un modelo llamado VAR-PLS.
  - 2 Introducir variables determinísticas (dummies, tendencias, etcétera) y exógenas.
  - 3 Construcción de intervalos de predicción via Bootstrap.
  - 4 Construcción de un modelo VAR con capacidad predictiva y compararlo con los pronósticos realizados con el modelo VAR-PLS.

$$PLSAR(h, p)$$



Franses plantea la comparación entre tres formas de hacer pronósticos bajo un  $AR(p)$ :

- 1- Un modelo único para todos los horizontes, usando un procedimiento iterativo

$$AR(p) : y_{T+h} = \mu + \rho_1 y_{T+h-1} + \rho_2 y_{T+h-2} + \cdots + \rho_p y_{T+h-p} + \epsilon_T$$

El modelo  $AR(p)$  es la forma clásica de realizar los  $h$  pasos hacia adelante, cuyos parámetros son estimados generalmente mediante Mínimos Cuadrados Ordinarios (OLS).

- 2- Un modelo para cada horizonte, donde la varianza cambia con cada horizonte

$$AR_h(p) : y_{t+h} = \mu + \rho_{1,h}y_t + \rho_{2,h}y_{t-1} + \cdots + \rho_p y_{t-p} + \epsilon_t$$

- El  $AR_h(p)$  es una alternativa al caso anterior, ya que OLS minimiza la suma cuadrada de  $\epsilon_t$ , pero no garantiza que sea mínima para  $h$  errores futuros.
- Una opción es contar con diferentes modelos para cada horizonte de pronóstico.
- Para series de tiempo estacionarias, recordemos que el pronóstico de un  $AR(p)$  converge rápidamente a la media incondicional (obviamente, la rapidez depende directamente de  $h \geq p$ ).
- Para más detalles sobre este tipo de modelos, puede consultarse Pesaran & Pick (2010), Marcellino, Stock & Watson (2004), Carreiro, Kapeterios & Marcellino (2010), Tiao & Xu (1993), entre otros.

- 3- Algo intermedio: *PLSAR*. Este modelo se sitúa entre un  $AR(p)$  que pronostica todos los  $h$  pasos adelante y diferentes modelos  $AR$  para cada horizonte.

$$PLSAR(h, p) : \hat{Y} = XB_{PLS}$$

- Es claro que existen correlaciones entre las series de tiempo que los dos modelos anteriores no explotan. Dicho en otras palabras: sabemos que  $(y_t, t_{t-j})$  están correlacionadas, mas aún  $(y_{T+h}, t_{T+h-j})$  también lo están, entonces, una alternativa viable es predecir conjuntamente  $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$  con la ayuda de  $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$ , y PLS es una técnica atractiva para ello.

- 3- Algo intermedio: *PLSAR*. Este modelo se sitúa entre un *AR*( $p$ ) que pronostica todos los  $h$  pasos adelante y diferentes modelos *AR* para cada horizonte.

$$PLSAR(h, p) : \hat{Y} = XB_{PLS}$$

- Franses propone reorganizar la información de la siguiente manera:

$(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$  en una matriz de predictores  $X$

$(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$  en una matriz de respuestas  $Y$

y realizar la regresión con PLS, de tal manera que el proceso de construcción de variables latentes y cargas asociadas contengan la información relevante que tiene  $X$  en  $Y$ .

- En su trabajo muestra que un modelo PLS planteado como un fenómeno autorregresivo es competitivo comparado con modelos clásicos existentes en la literatura.

- El método PLS puede abordarse desde diferentes ópticas. En la perspectiva de este trabajo, podemos relacionar la expresión del vector autorregresivo con la forma de un modelo lineal en la forma clásica

$$Y = XB + U,$$

donde  $Y$  es de  $N \times k$ ,  $X$  es de  $n \times N$ ,  $B$  es una matriz de  $(N + 1) \times k$  y  $U$  es de  $N \times k$ .

- Bajo la concepción de Franses que explota las correlaciones existentes entre  $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$  y  $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$ , la diferencia esencial entre *VAR* y *PLS* es que, mientras el primero usa directamente OLS, el segundo usa proyecciones sobre variables latentes, es decir, una descomposición que maximiza la covarianza existente entre la matriz  $X = Y'_{t-1}$  con  $Y = Y'_t$ .

*PLS*

- El procedimiento básico resuelve

$$\max \text{cov}(X\alpha, Y\beta)^2$$

con las restricciones

$$\alpha'(S_{xx}^* + \lambda_x)\alpha = 1 \quad \text{y} \quad \beta'(S_{yy}^* + \lambda_y)\beta = 1$$

donde  $S_{xx}^* = (1 - \lambda_x)S_{xx}$  y  $S_{yy}^* = (1 - \lambda_y)S_{yy}$ .

- $(X\alpha, Y\beta)$  representa una combinación lineal de las variables que maximizan la covarianza o covarianza al cuadrado (no interesa el signo, sino la maximización).
- $S_{xx}$  y  $S_{yy}$  son las matrices de varianza y covarianzas,  $\beta'\beta = 1$  y  $a'S_{xx}a = 1$ .

- Para maximizar nuestra función objetivo obtenemos:

$$\mathcal{L} = (\alpha' S_{xy} \beta)^2 - \gamma (\alpha' (S_{xx} + \lambda_x) \alpha - 1) - \mu (\beta' (S_{yy} + \lambda_y) \beta - 1).$$

- Luego de un poco de álgebra, obtenemos los scores para  $X$  y  $Y$ ,  
 $t = Xw = Ew$  y  $u = Yq = Fq$ .
- Normalizando los scores  $t = t/\sqrt{t't}$ , y continuando con más simplificaciones, se obtienen los loadings para  $X$  y  $Y$ :  $p = E't$  y  $q = F'u$ .



- Agrupamos cada  $w$ ,  $t$ ,  $p$  y  $q$  en matrices  $R = W(P'W)^{-1}$  y finalmente se reconstruye

$$Y = XB + U, \text{ entonces } \hat{Y} = XB_{PLS},$$

donde  $B_{PLS} = R(T'T)^{-1}T'Y = RQ'$ .

- Una buena introducción puede encontrarse en

P.H. Garthwaite (1994). An Interpretation of Partial Least Squares. JASA Vol 89, No 425, pp 122-127.

Agnar Hoskuldsson (1988). PLS Regression Methods. Journal of Chemometrics, Vol 2, pp 221-228.

# Vectores Autorregresivos y Mínimos Cuadrados Parciales

Un proceso  $VAR(p)$  se define como

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + CD_t + u_t,$$

donde

- $A_i$ ,  $i = 1, 2, \dots, p$ , son matrices de coeficientes
- $u_t$  es un proceso de ruido blanco con matriz de covarianzas  $\Sigma_u = E(u_t, u_t')$
- $C$  es una matriz de regresores determinísticos
- $D_t$  es un vector de regresores determinísticos apropiados

Notemos que el  $Var(p)$  puede definirse como un  $Var(1)$  mediante

$$Y_t = AY_{t-1} + V_t,$$

con

$$Y_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}, \quad A = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{pmatrix}, \quad V_t = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Si los valores propios de  $A$  son menores que 1, entonces el proceso  $Var(p)$  es estable.

- Un procedimiento para encontrar el orden  $p$  del modelo consiste en ordenar los  $p = 0, \dots, p_{\max}$  y elegir el valor  $p$  que minimice cierto criterio de información de la forma

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

donde

- $\hat{\Sigma}(p) = T^{-1} \sum_{i=1}^T \hat{u}_t' \hat{u}_t$ ,
- $C_T$  es una secuencia indexada por el número de realizaciones de  $T$
- $\varphi(K, p)$  es una función que penaliza la complejidad del modelo

- Los cuatro criterios de información más utilizados son
  - Akaike:  $AIC(p) = |\Sigma(\hat{p})| + \frac{2}{t}pK^2$
  - Schwartz-Bayesiano:  $BIC(p) = |\Sigma(\hat{p})| + \frac{\log T}{t}pK^2$
  - Hannan-Quinn:  $HQ(p) = |\Sigma(\hat{p})| + \frac{2 \log T}{t}pK^2$
  - Error final de predicción:  $FPE(p) = \left(\frac{T+p^*}{T-p^*}\right)^K \det(\Sigma(\hat{p}))$
- El criterio AIC sobreestima asintóticamente el orden con probabilidad positiva, mientras que BIC y HQ estima consistentemente el orden cuando el verdadero valor de  $p \leq p_{max}$ .

Notemos que al igual que el caso univariado, podemos pronosticar recursivamente mediante

$$y_{T+h\backslash T} = A_1 y_{T+h-1} + \cdots + A_p y_{T+h-p} + CD_{T+h}$$

- La estimación de  $A_i$  se realiza generalmente mediante OLS

$$\text{vec}(\hat{A}) = \begin{pmatrix} \hat{A}_1 \\ \vdots \\ \hat{A}_p \end{pmatrix}.$$

- Bajo ciertas condiciones de generalidad del comportamiento estacionario y ergodicidad en los modelos VAR (Hamilton, 1994, Lutkepohl, 1991),  $\text{vec}(\hat{A})$  es consistente y distribuido asintóticamente con matriz de covarianzas

$$\widehat{\text{var}} \left( \text{vec}(\hat{A}) \right) = \hat{\Sigma} \otimes (Z'Z)^{-1},$$

donde

$$\hat{\Sigma} = \frac{\sum_{t=1}^T \hat{\epsilon}_t' \hat{\epsilon}_t}{T - K}$$

y

$$\hat{\epsilon}_t = Y_t - \hat{A}'Z_t = Y_t - \hat{A}'Y_t$$

es el residual de mínimos cuadrados al tiempo  $t$ .



- El  $i$ -ésimo elemento de  $\text{vec}(\hat{A})$  es asintóticamente normal (para un VAR estable) con errores estándar dados por la raíz cuadrada de los elementos de la diagonal de  $\hat{\Sigma} \otimes (Z'Z)^{-1}$ .
- Las pruebas  $t$  son válidas asintóticamente para los coeficientes estimados.

Una situación de interés sucede cuando existen una o más raíces unitarias en las  $y_j$ 's. Esto ha originado una teoría económica que se basa en modelar el comportamiento de largo plazo y analizar la dinámica temporal de un conjunto de series.

- **Cointegración:** Las componentes del vector  $y_t$  se dice que son cointegradas de orden  $d, b$ , lo cual denotamos con  $y_t \sim CI(d, b)$ , si
  - 1 todos los componentes de  $y_t \sim$  son  $I(d)$
  - 2 existe un vector  $\beta \neq 0$  tal que  $z_t = \beta' y_t \sim I(d - b)$ ,  $b > 0$ . El vector  $\beta$  es llamado vector de cointegración.

- **Corrección de error:** Consideremos el vector bivariado  $y_t = (y_{1t}, y_{2t})'$  con vector de cointegración  $\beta = (1 - \beta_2)'$ , entonces  $\beta' y_t = y_{1t} - \beta_2 y_{2t} \sim I(0)$  y existe una representación de corrección de error de

- 1  $\Delta y_{1t} = \alpha_1 + \gamma_1(y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^K \psi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^K \psi_{2,i} \Delta y_{2t-i} + u_{1t}$
- 2  $\Delta y_{2t} = \alpha_2 + \gamma_2(y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^L \xi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^L \xi_{2,i} \Delta y_{2t-i} + u_{2t}$

Notemos que el modelo  $VAR(p)$  puede expresarse como

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t,$$

donde  $\Gamma_i = -(A_{i+1} + \cdots + A_p)$ , para  $i = 1, \dots, p-1$  y

$\Pi = -(I - A_1 - A_2 - \cdots - A_p)$ .

Lo anterior es llamado Modelo de Vector de Corrección de Error (VECM) transitorio.

O bien como

$$\Delta y_t = \Pi y_{t-p} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t,$$

donde  $\Gamma_i = -(I - A_1 - A_2 - \cdots - A_i)$ , para  $i = 1, \dots, p-1$  y

$\Pi = -(I - A_1 - A_2 - \cdots - A_p)$ . Este modelo es llamado VECM de largo plazo.

La matriz  $\Pi$  tiene las siguientes características:

- 1  $rk(\Pi) = n$ , todas las  $n$  combinaciones deben ser estacionarias; el VECM es un modelo VAR en niveles
- 2  $rk(\Pi) = 0$ , no existe una combinación lineal estacionaria tal que  $\Pi y_{t-1}$  sea estacionaria, excepto la solución trivial, i.e, es un modelo  $VAR(p - 1)$  en primeras diferencias.
- 3  $0 < rk(\Pi) < n$ , en este caso  $\Pi = \alpha\beta'$  con dimensión  $n \times r$  y  $\beta'y_{t-1}$  es estacionaria. Cada columna de  $\beta$  representa una relación de largo plazo.

**Si el objetivo es pronosticar, aún en el caso de variables integradas y cointegradas, hacerlo mediante la representación VAR es muy apropiado (ver Lutkepohl, 2006).**

Del ejemplo, notamos que (VER ESTA PARTE!1!1)

- Se especificó el orden  $p$  del VAR teórica a través del criterio de Error Final de Predicción, el cual fue de 12. También se realizó la prueba de Johansen para verificar la presencia de relaciones de largo plazo.
- Los resultados obtenidos mostraron que al 1%, 5% y 10% existe una ecuación de cointegración, la cual está dada por la siguiente expresión:

$$p + 9.68 - 0.43m - 0.89y + 0.1r = 0$$

- Lo anterior es congruente con la realidad, ya que establece que el traspaso inflacionario está impulsado por el crecimiento monetario, exceso de demanda y la reducción del costo del dinero.

# VAR-PLS

- Se propone utilizar la representación de un modelo VAR para el proceso generador de datos
- El modelo VAR determinará el orden del proceso autorregresivo que será utilizado en la regresión PLS
- Utilizaremos matrices en estructura de rezagos según cada una de las variables



- Agrupamos en la matriz  $X$  la información observada en el pasado:

$$X = Y_{t-1} = \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{pmatrix},$$

- En  $Y$  lo observado en el tiempo  $t$ :

$$Y = Y_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix},$$

estimando el modelo mediante el proceso de construcción de variables latentes.

- Es fácil introducir variables exógenas o determinísticas mediante una matriz  $C$ , modificando la composición de

$$X = Y_{t-1}^* = \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \\ D_t \end{pmatrix},$$

y por ende, de la matriz de coeficientes

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

- De esta forma obtenemos un modelo  $VAR - PLSX$  con la única finalidad de pronosticar  $h$  pasos adelante.

Es importante señalar que, para el VAR-PLS podemos obtener hasta  $pK$  componentes. Para este ejercicio se estiman cada una de ellas solo con fines comparativos, y para observar la capacidad predictiva respecto a métodos recursivos de pronóstico contruidos a partir de un modelo VAR.

Como el objetivo es construir un modelo de pronóstico robusto que a su vez se compare con la técnica VAR-PLS, realizaremos el siguiente procedimiento:

- Se guardan las 24 observaciones finales con el fin de obtener una ventana de tiempo de horizonte largo que sirva para comparar cada uno de los modelos.
- Para determinar el orden del VAR-PLS se determina la selección del rezago óptimo a través de algún criterio de información.

- Para verificar que el proceso generador de datos sea consistente con la teoría económica, se estima el  $VAR(p)$  denotando su característica estocástica, es decir, si es cointegrado o no, obteniendo en su caso los coeficientes de largo plazo.
- Para  $p$  óptimo con una especificación constante, y sin variables determinísticas, se procede a estimar el VAR-PLS con  $h = 24$ , obteniendo el error fuera de muestra para cada una de las componentes existentes  $pK = 48$
- Para el modelo  $VAR(p)$ , se tienen 4 variables, 4 posibles especificaciones del VAR (ninguna, constante, tendencia, constante y tendencia), y 11 posibles variables dummies estacionales (una para cada mes). Combinando todas las variables con las especificaciones y variables determinísticas posibles nos da un total de  $VAR_j(p)$ ,  $j = 1, \dots, 484$  posibles a estimar.

- Se estiman cada uno de los modelos y se seleccionaron los que minimizan 7 criterios de error fuera de muestra, es decir, se eliminaron otras  $h = 24$  observaciones con el fin de calcular los estadísticos de error (Hyndman & Koehler, 2006).
  - MAPE: Mean Absolute Percentage Error
  - MdAPE: Median Absolute Percentage Error
  - RMSPE: Root Mean Square Percentage Error
  - RMdSPE: Root Median Square Percentage Error
- Adicionalmente se trabajó con un modelo benchmark (AR de orden 1) donde se calcularon los  $i = 1, \dots, 24$  pronósticos, generando el siguiente estadístico:

$$test = \frac{Y_{t-i} - Y_{t+i, VAR_j(p)}^f}{Y_{t+i} - Y_{t+i, AR(1)}^f}$$

- Posteriormente se obtienen los siguientes 3 estadísticos:
  - MRAE: Mean Relative Absolute Error
  - MdRAE: Median Relative Absolute Error
  - GMRAE: Geometric Mean Absolute Error
- Con este procedimiento se obtienen 7 modelos, los cuales se integran en un solo número obtenido del cuantil 50% para cada uno de los horizontes de pronóstico.
- Se estima el  $VAR(12) - pls(24, j)$  y se realizan los  $pK = 48$  modelos, obteniendo para cada serie el MAPE para fines comparativos y para un horizonte de pronóstico de 24 datos.

## Intervalo de predicción: VAR-PLS

Uno de los objetivos de este trabajo es la construcción del intervalo de predicción para el modelo VAR-PLS.

El procedimiento es similar a lo propuesto por Pascual y colaboradores (2011), que a su vez, se basan en el trabajo de Kim (2001) y de Pascual y colaboradores (2004)

El procedimiento, cuando se realizar para el VAR consiste en

- 1 Introdurcir la incertidumbre debida a la estimación de parámetros y corregir las regiones de confianza asociadas (válidas asintóticamente bajola distribución Gaussiana).
- 2 Simplificar computacionalmente el cálculo con VAR que tienen muchos rezagos y regularmente implican que la representación backwards resulte muy compleja.



El procedimiento es el siguiente:

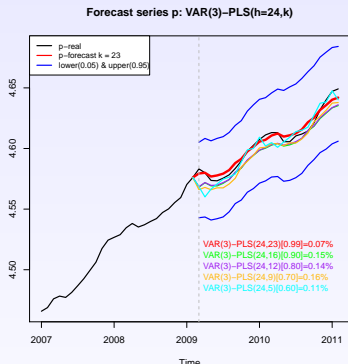
- 1 Estimar  $Y_t = X_t \hat{B}_{PLS} + \hat{U}_t$ .
- 2 Obtenemos  $\hat{U}_t^* = \hat{U}_t - \tilde{\bar{U}}$  para obtener una muestra ordenada de los residuos centrados y reescalados.
- 3 Con los  $p$  valores iniciales  $Y_0 = \{Y_p, \dots, Y_1\}$  y con lo obtenido en los pasos 1 y 2, generamos  $Y_t^*$  a través de

$$Y_t^* = X_t \hat{B}_{PLS} + \hat{U}_t^*, \quad t = 1, \dots, n - p$$

- 4 Obtenemos  $\hat{Y}_{T+h}^*$  repitiendo los pasos 2 a 4, para  $n = 1, \dots, N$  repeticiones
- 5 Finalmente, para cada una de las  $n$ th variables y los  $N$  conjuntos de pronósticos obtenemos

$$CI_{T+h} = \{y_{n,T+k} | y_{n,T+k} \in [q_B^*(\tau), q_B^*(\tau - 1)]\},$$

donde  $q_B^*(\tau)$  es el  $\tau$ -ésimo percentil de  $G_{n,B}^*(x) = \# \left( y_{n,T+k}^{*(b)} \leq x \right) / N$



- Podemos observar que el valor real prácticamente es el mismo que el pronóstico obtenido con un error absoluto promedio de 0.11%. Se muestran también los intervalos de predicción via Bootstrap que “atrapan” el valor real de la serie.
- En términos económicos, esta aproximación es competitiva.

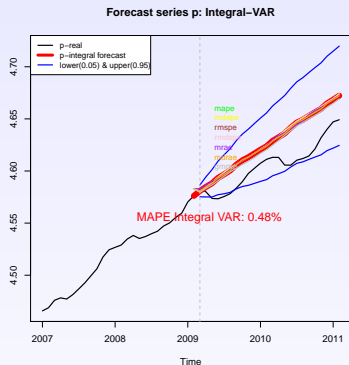
Notemos que, al ser un modelo multivariado se obtienen el resto de los pronósticos de las series, sin embargo, dado que el objetivo es predecir la serie de precios, no se considera relevante para este trabajo observar a detalle el comportamiento. No obstante, podemos comentar que el pronóstico con esta misma componente tuvo un error de 0.23% para  $m_0$ , para y 0.83% y finalmente con un 5.84% la tasa de interés.  
(NO LE ENTIENDO A ESTA REDACCIÓN!!!!!!!!!!)

Para el Integral-VAR, los  $VAR_J(P)$  óptimos fueron los siguientes:

Criterio	MAPE	MdAPE	RMSPE	RMdSPE
Estadístico	0.29	0.30	0.33	0.30
Rezagos	2	2	2	2
Estacionalidad	10	10	10	10
Especificación	Constante y tendencia	Constante y tendencia	Constante y tendencia	Constante y tendencia

Criterio	MRAE	MdRAE	GMRAE
Estadístico	0.24	0.22	0.18
Rezagos	2	2	3
Estacionalidad	4	11	7
Especificación	Constante y tendencia	Constante y tendencia	Constante y tendencia

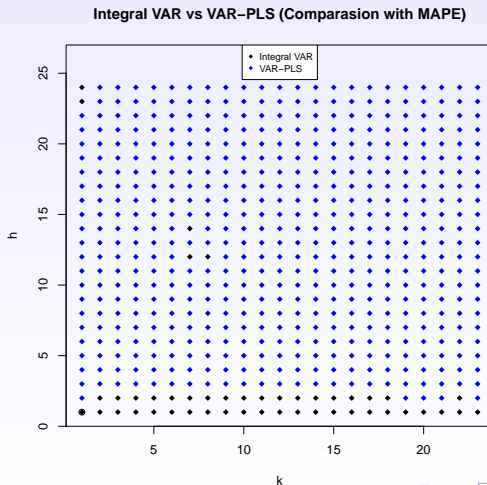
Integrando los pronósticos y el intervalo de predicción con su respectivo cuantil del 50% para cada modelo y horizonte de pronóstico, obtenemos la siguiente figura:



Aunque en términos numéricos el error fue de 0.32%, se observan grandes diferencias respecto al real observado, sobreestimando la inflación en cada uno de los casos. No obstante se decidió realizarlo de esta manera para comparar las variables que intervienen en los pronósticos para ambas especificaciones.

# VAR – PLS

La siguiente figura muestra cómo se comportó cada componente respecto al Integral VAR:



Resulta interesante que en promedio, un 73.52% de las veces fueron superior las componentes, y es lógico que aquellas más lejanas dejan de tener efectividad, dado que explican menos la variabilidad observada.

En otras palabras, la metodología VAR-PLS resulta atractiva respecto a su competidor inmediato, que es un VAR enfocado a predecir.

## Conclusiones



# Conclusiones

- En este trabajo se presentó una alternativa para realizar pronósticos multivariados con el enfoque de explotar un conjunto de series de tiempo  $y_{t+j}$  con  $y_t$ , aprovechando la naturaleza que tiene por construcción la técnica de PLS, específicamente al momento de plantear un modelo lineal, situación que se presenta para el caso de un VAR.
- Se planteó un modelo que considera como principales fuentes inflacionarias un agregado monetario (billetes y monedas), tasas de interés y una variable de ingreso, y se realizaron dos metodologías de pronóstico:
  - un VAR-PLS cuyo orden  $p$  está dado por el criterio de FPE de un VAR, estimando todas las  $pK$  componentes
  - Finalmente, los intervalos de predicción se hicieron mediante Bootstrap

NO CREO QUE ESTE BIEN ESTA REDACCIÓN... REVISAR!!!

De acuerdo a los resultados obtenidos, el VAR-PLS es una técnica atractiva de pronóstico multivariado.

Como trabajo futuro queda:

- Incluir PLS en el ejercicio de cointegración, considerando las implicaciones teóricas.
- Construir modelos PLS-VAR que integren cada una de las componentes
- Comparar VAR-PLS y VAR que mezclen las variables que intervienen dentro de un pronóstico.

Gracias por su atención !