**An autoregressive representation for forecast macroeconomic time series with Partial Least Squares: The case of Mexican inflation**

1. **Introducción**

En econometría y en el análisis de las series de tiempo macroeconómicas desde una perspectiva estadística, existen diversas metodologías para pronosticar los valores futuros de éstas y también, si es el objetivo del problema, dar explicaciones del porqué tomará tales valores el pronóstico realizado, sea cual sea la técnica seleccionada; más si consideramos la importancia que tiene en política económica el hacer aseveraciones sobre el futuro de la o las series analizadas.

En lo que a esto respecta, la literatura que aborda estas cuestiones es extensa, donde podemos mencionar a los modelos ARIMA (ver Box & Jenkins, 1970), los VAR (ver Sims, 1980; Johansen 1991, entre otros) y en sí, modelos de carácter estructural dictaminados por la teoría económica (ver Klein & Young, 1980; Blanchard & Quah, 1989; Lütkepohl, 2006).

Otras técnicas que podemos mencionar, y que por su facilidad en la construcción de pronósticos resulta atractiva, encontramos a toda la familia de suavizamientos exponenciales (e.g. Trigg & Leach, 1967), los cuales no haremos referencia en este trabajo **(poner porqué)**

A la par de estas técnicas y metodologías desarrolladas, existe una muy popularizada en diversas áreas del conocimiento, conocida como Mínimos Cuadrados Parciales (PLS) cuyas aplicaciones se destacan más en Biología, Química, Ingenierías, etc., (e.g. Frank & Friedman, 1993; Nomikos, McGregor, 1995; Barker & Raynes, 2003; Barceló, Vidal-Puig, Ferrer, 2010, entre muchos otros) pero que sin embargo, su utilización no se encuentra limitada a éstas áreas e incluso, en Econometría ha mostrado su potencialidad en diversas aplicaciones **(poner citas)**

Aunque son muchos los atributos de la técnica de PLS (ver Esposito-Vinzi, Chin,Henseler, Wang, 2010), dos características esenciales que hacen atractiva su utilización en la modelación de series de tiempo macroeconómicas, es que no representa problema la correlación existente entre las variables llamadas predictivas, dado que una de sus propiedades es extraer el número de componentes significativas no correlacionadas, y más aún, PLS es un método predictivo que maximiza la relación existente entre una combinación lineal de la matriz respuesta y la matriz predictora .

Para fines del modelo, las componentes significativas utilizadas para la predicción pueden ser iguales o mayores a 1 y hasta equivalente el número de variables predictoras .

En otras palabras, PLS es un ejercicio que transversalmente, genera combinaciones lineales de las variables predictoras originales, comúnmente aplicamos el concepto de reducción de dimensionalidad y asimismo éstas, están relacionadas con la matriz respuesta, pudiendo establecer una relación lineal entre ellas, componente por componente, haciendo una secuensación de regresiones ponderadas.

La motivación de este trabajo radica en explotar la relación existente entre un grupo de variables de tal manera que las comúnmente llamadas independientes en un modelo lineal, tengan capacidad predictiva para pronosticar las variables dependientes.

En lo que respecta a los modelos más utilizados en econometría, hace sentido pensar el problema de esta manera, donde un grupo de variables seleccionadas por la teoría económica tengan relación con la o las series de tiempo de interés, existiendo la posibilidad de que haya relaciones multivariadas entre éstas, uno de los motivos por los cuales la flexibilidad de los modelos VAR ha sido ampliamente popularizada propiciando su utilización en la economía.

Claro está existe una amplia literatura sobre la especificación de los modelos VAR (ver Granger, 1969, 1988, 2009; Ericsson, Hendry & Mizon 1998, entre otros), pero en este trabajo nos limitamos a especificar un modelo VAR de la forma tradicional y posteriormente, a través de PLS, especificar las matrices de la forma , estimando la proyección o pronóstico de la matriz respuesta o de variables dependientes de interés, evaluando el error de predicción fuera de muestra ambos modelos.

En la sección 2 se detalla la metodología propuesta que se utilizará en este trabajo, el perfil macroeconómico del modelo y los datos que lo conforman. En la sección 3 se presentarán los cálculos de las estimaciones realizadas. En la sección 4 se presentarán de manera puntual los resultados y en la 5 se discutirá lo obtenido respecto comparándolos con otros existentes en la literatura. Finalizamos en la sección 6 con las conclusiones y líneas futuras.

1. **Métodos y materiales**
   1. **Vectores Autorregresivos**

Los modelos VAR han sido ampliamente comentados en la literatura y no es objetivo de este trabajo abordar el tema, no obstante abordaremos los detalles más importantes que dan forma a la especificación autorregresiva, importante en el procedimiento de pronóstico sugerido en este trabajo.

Es importante recalcar que el VAR o la expresión de Vector de Corrección de Errores (VEC) depende directamente del concepto de cointegración (ver Engle & Granger, 1987; Johansen, 1988 entre otros), siendo indistinta la especificación si el objetivo es el pronóstico (ver Lütkepohl, 2006) ya que éste depende directamente de la estabilidad que posea el modelo VAR estimado.

En términos econométricos, siguiendo con la representación clásica del proceso generador de datos (DGP) (ver Ericsson, Hendry & Mizon 1998), éste se representa a través de la distribución conjunta de los datos conformado de la forma que denota el espacio de probabilidad, donde es el espacio muestral para el vector de variables en el tiempo que caracterizan a la economía, es el espacio de eventos de y es la medida de probabilidad de los eventos . Considerando la historia de , como un proceso estocástico de la forma tradicional hasta el tiempo con la expresión donde o , denotamos a como las condiciones iniciales de la economía y para .

Así el DGP puede ser factorizado como producto de las distribuciones marginales:

()

Donde es la densidad conjunta de dado , es un vector de parámetros dentro de la muestra , es un subconjunto de parámetros de que integran cada una de las distribuciones marginales y .

De esta manera se puede plantear el modelo econométrico para , el cual se expresa similarmente a () de la siguiente forma:

()

Donde representa el modelo econométrico, es la distribución marginal condicionada de , es un vector de vector de parámetros del espacio paramétrico .

En otras palabras el modelo VAR de orden , más generalizado que se puede presentar, es como una función (en nuestro caso lineal) de sus observaciones pasadas:

()

Con siendo un vector de variables , matrices de coeficientes de tamaño , es un vector de errores con media 0 (no correlacionados e independientes) con matriz de varianzas y covarianzas .

Notemos que el la expresión () podemos representarla como un modelo de la siguiente manera:

()

Donde:

**FALTAN LAS MATRICES PARA EJEMPLIFICAR**

Para fines de pronóstico, podemos utilizar () y realizar la estimación de comúnmente a través del método de Mínimos Cuadrados Ordinarios (OLS) o Máxima Verosimilitud (ML) y obtener el estimador de con:

()

En este caso, el orden autorregresivo del VAR será determinado a través del criterio Error Final de Predicción (FPE) **(porqué)**:

()

Es claro que el modelo () da cabida a una serie de conceptos interesantes en materia de política económica, tales como el de cointegración, restricciones estructurales y representaciones alternas subsecuentes (ver Vahid & Engle, 1993; Gonzalo & Granger 1995), las cuales aunque importantes, no son materia de análisis de este trabajo dado que el objetivo primordial es una realización del pronóstico con el modelo VAR planteado.

* 1. **Mínimos Cuadrados Parciales**

El concepto de PLS que comúnmente se maneja en la literatura, se basa como un problema de maximización, que difiere al de Correlación Canónica (CCA), en que maximiza la covarianza existente entre una combinación lineal de variables predictoras en función de otra combinación lineal de variables predichas Y, mientras que en CCA lo realiza en función de la correlación.

Es interesante desde la perspectiva econométrica establecer las analogías entre uno y otro ejercicio, principalmente en lo que respecta al análisis de cointegración, el cual hace uso del concepto de CCA (ver Johansen 1988; Johansen & Juselius, 1991) sin embargo, este trabajo representa una línea futura de investigación y no será abordada en este momento.

En otras palabras, en PLS el objetivo podemos expresarlo como el siguiente problema de maximización:

()

()

Nótese que representa una combinación lineal de las variables involucradas siendo más o menos importante en sentido de explicación alguna de estas, similar al caso del método de componentes principales (PCA) **(poner referencia)**, y maximizar la covarianza o la covarianza al cuadrado es indistinto pues no interesa el signo sino la maximización. Finalmente y son las matrices muestrales de varianza y covarianza.

Un caso interesante es el ejercicio canónico de Ridge (ver Vinod, 1976), donde al añadir una serie de elementos a las restricciones () podemos solucionar diferentes casos de PLS y CCA. Es decir:

& ()

Dónde:

& ()

Es fácil demostrar que la solución de () sujeto a () nos lleva a las siguientes soluciones:

()

()

Donde y son los valores propios asociados a la solución de cada matriz correspondiente a la solución característica, siendo y son los vectores propios ,i.e., las cargas asociadas que tiene cada combinación lineal.

Para expresar a en función lineal de las , de tal manera que intervengan los vectores propios producto de ejercicio de maximización, podemos realizar el denominado ejercicio de Regresión PLS (ver Wold, Sjostrom, Eriksson, 2001), el cual es un procedimiento recursivo que consiste básicamente en obtener los scores para y :

()

()

Asimismo, se normalizan los scores .

Es importante aclarar que los scores de no son normalizados ya que nos interesa para la interpretación de y , las cuales son iniciadas con y respectivamente. De esta manera obtenemos los loadings para y :

()

()

Con esto las matrices son “desinfladas” quitando el efecto relacionado a la variable latente anterior como:

()

()

Este proceso se repite hasta que y

Agrupamos cada y en matrices (las expresamos en mayúsculas) obtenemos:

()

De igual manera que en el caso tradicional de PCR **(poner cita)**, realizamos la regresión de en , usando los scores , calculando los coeficientes de regresión como un modelo lineal clásico () premultiplicando con () donde , para obtener finalmente:

()

Donde ()

De esta manera, concluye el ejercicio de regresión PLS. Existe una gran literatura al respecto, dentro de la cual podemos mencionar a…

* 1. **PLS y representación autorregresiva**

Expresar la matriz con los términos autorregresivos de ha sido utilizado en sentido univariado (Franses, 2006; Barceló, Vidal-Puig, Ferrer, 2010), principalmente para obtener los pronósticos de una variable de interés o para la estimación de funciones respuesta-impulso equivalente a la una representación ARIMA.

En nuestro trabajo proponemos utilizar la representación que da cabida a un modelo VAR, utilizando la matriz del modelo (), en lugar de la matriz del modelo lineal obtenido a través de PLS (), de tal manera que sea el DGP y la en sí, la metodología VAR, la que dictamine el orden del proceso autorregresivo que será utilizado en la regresión PLS.

Notemos que realizar el pronóstico utilizando () es equivalente a la expresión (), así, el método de comparación empírica que utilizaremos en este trabajo será la raíz media del error cuadrático medio, esto para cada una de las variables que integran a ambas representaciones:

()

Al seleccionar un orden autorregresivo , perdemos las primeras observaciones, de manera que, similar a Barceló, Vidal-Puig, Ferrer (2010), pero expresando multivariadamente las matrices según el modelo () podemos apreciar que la expresión () equivale a () de la siguiente manera:

**PONER MATRICES PARA OBSERVAR**

Empíricamente, estimaremos primero el VAR () a través de OLS y posteriormente el modelo PLS (). Así, calcularemos los pronósticos para cada modelo, computando los errores de pronóstico fuera de muestra para cada una de las variables que integren la modelación econométrica.

* 1. **Modelo empírico**

El modelo considerará, cómo fuente principal de la dinámica inflacionaria mexicana, el crecimiento o variación de las condiciones monetarias del país. Al respecto 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 éste.

Varios autores de la nueva economía Neo-Keynesiana determinan que la dinámica del sector monetario no determina la inflación de largo plazo, sin embargo otros autores critican estas conclusiones. En los trabajos de investigación de Edward Nelson y Michael Woodford **(poner año de referencia)** (ambos importantes en las áreas de monetarismo teórico y macroeconomía) se han expuesto, y mutuamente criticado, ambas teorías. Bajo los supuestos de Woodford (existencia de mercados completos, precios flexibles, sin fricciones monetarias) y basándose en la ecuación de Fisher-Euler, concluye que el impacto del Banco Central en el largo plazo se limita exclusivamente a la tasa de interés. Nelson, crítica a Woodford, argumentando fallas en el modelo Neo-Keynesiano que se basan en considerar a la política monetaria relevante para la inflación esperada a largo plazo, pero no considerar el crecimiento monetario como relevante. Nelson **(poner año de referencia)** concluye que la única forma que el Banco Central puede afectar las tasas de interés en el largo plazo es afectando la inflación.

Claro está, en este trabajo no abordaremos tal discusión, sino que se mostraremos las propiedades empíricas del modelo desarrollado en términos del error de predicción fuera de muestra.

**Falta**

**Cálculos**

**Resultados**

**Conclusiones**

**Apéndice**

**Vitae**