

INDICADOR ECONÓMICO REGIONAL: ESTIMACIÓN MENSUAL POR COMPONENTES PRINCIPALES

Mónica Iris CALDERÓN
Fernando DELURET

INTRODUCCIÓN

El rol creciente que han asumido las cuentas regionales en los últimos años demuestra su importancia para el análisis económico de corto plazo y justifica el esfuerzo creciente dedicado a estimarlas. Dado que, en algunas provincias, las actividades económicas se compilan sobre una base anual, la obtención y difusión de indicadores mensuales ha tenido un inicio relativamente reciente. Debe reconocerse que existen regiones en las que no se dispone de ningún tipo de información histórica y sistemática. La Dirección de Estadísticas e Investigaciones Económicas (DEIE) de Mendoza publica en su página Web el Producto Bruto Geográfico (PBG) de todas las provincias argentinas en forma anual del año 2005 al 2013, siendo su fuente el Ministerio de Industria de la Nación. Los datos son desagregados según rama de actividad, lo que permite observar que, dentro de cada una de las regiones, las actividades económicas de periodicidad menor a la anual, pueden ser disímiles, así ser muy importantes para una región, y no así para otras o para la nación misma; por ello, no necesariamente sus comportamientos se verán reflejados en un indicador a nivel país como es el Estimador Mensual de Actividad Económica (EMAE) elaborado para la Argentina en su conjunto. Configurar un índice explicativo de lo que se denomina como economías regionales, aunque la misma sea únicamente una diferenciación práctica, es de suma utilidad. Este proyecto se enfocará en la economía regional que define el territorio comprendido en la provincia de Mendoza. El presente proyecto tiene dos propósitos: uno es ofrecer un paneo general para comprender el desarrollo de la investigación, circunscripta a la Argentina. Incluye una revisión de la literatura referida a estimaciones y predicciones de la actividad económica de distintas regiones. El segundo es presentar empíricamente un indicador compuesto de periodicidad mensual, para las actividades económicas mendocinas.

La provincia de Mendoza cuenta con una medición anual del Producto Bruto Geográfico (PBG), desagregada en nueve divisiones, según código CIIU versión 2; también presenta una descomposición del PBG para cada uno de los dieciocho departamentos provinciales. El PBG según lo define CEPAL, es “el cálculo del Producto para una jurisdicción que representa una porción del territorio total del país” ... “el Producto Bruto Interno de una jurisdicción determinada debería reflejar la actividad económica de las unidades productivas residentes en esa jurisdicción. La delimitación de las jurisdicciones relevante para la determinación del territorio económico se establece en función de las fronteras políticas de las provincias y los territorios autónomos, cubriendo la totalidad del territorio económico del país”. Como se aprecia, estimar el PBG implica una laboriosidad, que lo hace poco oportuno para la toma de decisiones en el corto plazo, por ello, se requieren índices de una periodicidad de menos de año, principalmente construidos en forma mensual. Se recurre a la obtención de indicadores compuesto de actividad económica, que buscan reflejar la evolución del nivel de actividad de una economía a partir de la combinación de varios indicadores sectoriales. Existen varias metodologías, las mismas dependerán tanto del criterio estadístico utilizado, como el criterio con el cuál se seleccionan las variables y, a su vez, la ponderación que se les otorga.

Esta investigación avanza en la construcción de un Índice de Actividad Económica (IAEMZA) para la provincia de Mendoza, el cual se estima de manera mensual. Esta propuesta intenta presentar un

aporte útil para satisfacer la necesidad de información actualizada de la situación económica regional. Para la construcción del índice se usa el Análisis Factorial por Componentes Principales a partir de un conjunto de variables mensualizadas, que se consideran representativas de la actividad económica de Mendoza; las cuales se encuentran disponibles en forma oportuna, esto es, el rezago para su publicación es de menos de seis meses.

El análisis de componentes principales (ACP) es un método de reducción de dimensionalidad lineal de los datos, el objetivo es poder reducir el número de variables necesarias para representar un conjunto de datos. Expresado en términos estadísticos, lo que se hace es realizar una proyección ortogonal de los datos originales en un subespacio vectorial de menor dimensión, que por ende pueda ser representado como una agrupación menor de variables, llamadas factores o componentes.

La investigación se estructura de la siguiente manera, se inicia con una presentación del estado actual tanto de la situación de este tipo de estadísticas a nivel provincial y nacional, como así también de la situación de la herramienta teórica usada que es el Análisis de Componentes Principales (ACP en sus siglas en español). En la sección que sigue se analizan una serie de conceptos previos necesarios tanto para el desarrollo de ACP. Luego, se presenta el tema desde la perspectiva del análisis factorial, lo cual permite añadir el componente fundamental de interpretabilidad de los componentes. Luego se revisa la rotación de componentes, que es la herramienta que permite relacionar las variables originales con los componentes o factores estimados por ACP. Se interpretan estas últimas como variables subyacentes observadas a través de los datos iniciales. Hasta aquí se trabaja con aspectos netamente teóricos, para dedicarse a la aplicación empírica que es construir un índice de actividad mensual para la provincia de Mendoza. Se evalúa el comportamiento del indicador obtenido, para arribar a algunas conclusiones.

ESTADO ACTUAL E IMPORTANCIA

Un antecedente insoslayable para Argentina, es el realizado por el Instituto Nacional de Estadísticas y Censos –INDEC-, el cual, ejerce la dirección superior de todas las actividades estadísticas oficiales que se realizan en el territorio de la República Argentina (INDEC, 2017); es el responsable del cálculo del Estimador Mensual de Actividad Económica (EMAE), se trata de “un indicador provisorio de la evolución del PIB a precios constantes de 2004 que es difundido con un rezago de 50 a 60 días de concluido el mes de referencia, de acuerdo al calendario de difusión del INDEC” (INDEC, 2016). Un avance importante a nivel de todas las provincias argentinas lo representa el trabajo de Muñoz & Trombetta (2014), los cuales, construyen un indicador mensual sintético de actividad económica, para las veinticuatro provincias y para la República Argentina utilizando la metodología de la Conference Board realizando modificaciones en el procedimiento de cálculo del índice, sin cambiar la intención del algoritmo. Los autores expresan que estos han sido construidos a partir de información estadística de acceso público, con frecuencia y actualización trimestral. Aseverando que su uso como proxy del nivel de actividad provincial, permitiría abordar los objetivos de monitoreo, evaluación y diseño de políticas públicas para las economías regionales.

Como intentos aislados, la necesidad de contar con información actualizada y permanente, lleva a las provincias a la búsqueda de un sustituto o bien un complemento del cálculo del PBG; así, Chaco y Formosa discontinuaron el cálculo del PBG presentando la elaboración de un indicador compuesto. En Corrientes, se realiza el cálculo del PBG de forma trimestral, con varios períodos de rezago en su publicación. Jorrat (2003) calcula un índice mensual compuesto de actividad económica para Tucumán, usando la metodología desarrollada por programa de investigación Ciclos Económicos de Argentina de la Universidad Nacional de Tucumán. Michel Rivero (2007), construye un indicador de actividad económica mensual para la provincia de Córdoba, mediante el uso de la metodología de construcción de indicadores de actividad propuesta por Jorrat, determinando sus puntos de giros mediante un índice de difusión.

Cuadro 1

Provincia	Nombre indicador	Periodicidad	Descripción	Metodología
Buenos Aires	ITAE-PBA *Nivel General *Desestacionalizado	Trimestral	El Indicador Trimestral de la Actividad Económica de la provincia de Buenos Aires (ITAE-PBA), desarrollado por la Dirección Provincial de Estadística, tiene como objeto disponer de un indicador que permita observar el comportamiento de la actividad económica provincial en el corto plazo y anticipar la evolución anual del Producto Bruto Geográfico.	El Indicador ITAE-PBA es una agregación ponderada de un conjunto de actividades que resultan explicativas de la variación trimestral del Producto Bruto Geográfico (PBG) a precios constantes de productor. Las ponderaciones corresponden a las cuentas provinciales base 1993 y se utilizan, en todos los casos en que resulta posible, las mismas fuentes de información que la estimación del PBG anual.
CABA	ITAE	Trimestral		Este indicador se elabora en base al seguimiento de los principales sectores productivos de la Ciudad, con una metodología semejante a la utilizada en la estimación anual de dichos sectores para el cálculo del Producto Geográfico Bruto.
Santa Fe	ISAE	Mensual	Indicador Sintético de la Actividad Económica: pretende estimar los movimientos de una variable no observable, denominada “el estado de la economía” a partir de los movimientos de un conjunto de variables coincidentes observables.	Se basa en la metodología propuesta por Stock y Watson, que postula un modelo probabilístico de espacio de estados que se utiliza para predecir un proceso latente y estimar, así, un indicador agregado coincidente de la actividad económica. Se fundamenta en la hipótesis de que los comovimientos observados en las series indicadoras son capturados por una única variable no observable, común a todas ellas, denominada el “estado de la economía”. La estimación de los parámetros del modelo se realiza mediante el filtro de Kalman.
Tucumán	IAT	Mensual	Índice Mensual de Actividad Económica de Tucumán	Este es un índice compuesto creado bajo la metodología del National Bureau of Economic Research de Estados Unidos (Conference Board, 2001), adaptada para parámetros y disponibilidad de datos locales, siguiendo Jorrat (2002, 2003 y 2005). Se construye en base a series coincidentes y tiene una periodicidad mensual, teniendo una correlación tanto en niveles como en variaciones logarítmicas con respecto al PBI, siendo estos 0,98 y 0,80 respectivamente.
Tucumán	IDT	Mensual	Índice de Difusión de Tucumán	Asimismo, se elabora el Índice de Difusión de Tucumán (IDT), el cual aporta información respecto a la intensidad de los movimientos cíclicos propuestos por el IAT, necesario para confirmar el inicio de una expansión o recesión.
Córdoba	InAEC	Mensual	Indicador de Actividad Económica de la Provincia de Córdoba	El InAEC busca compatibilizar dos criterios: Explicar la actividad utilizando la mayor información posible para cubrir una parte importante del PGB con las actividades consideradas. Para su cálculo el InAEC: Utiliza una metodología enmarcada en el enfoque contable.
Misiones	IEM	Mensual	Indicador Compuesto de Actividad Económica de Misiones (ICAEM)	En su construcción, se utilizan metodologías propuestas por Conference Board (2001) y Jorrat (2003), elaborado en base a once series componentes provenientes de fuentes públicas y representativas de los distintos sectores de la actividad económica provincial.

Fuente: Elaboración propia sobre la base de diversas páginas Web.

Igual camino siguieron D’Jorge et al (2007), con un índice mensual coincidente para Santa Fe. Lapelle (2013), construye un indicador sintético para la región de Rosario, se trata de un indicador coincidente y, para su construcción, sigue a The Conference Board. Como antecedente reciente de un trabajo sobre el tema se puede mencionar el trabajo de Heredia y Álvarez (2017) “Indicador

Compuesto de Actividad Económica para la provincia de Misiones”. Otros antecedentes sobre trabajos anteriores que se refieren a la obtención de indicadores compuestos son la estimación del indicador a través de modelos uniecuacionales, donde los residuos no explicados de la primera estimación, dan lugar a una nueva ecuación que se estima con una nueva variable explicativa, los residuos dan lugar a una nueva estimación con variables relacionadas y así sucesivamente hasta agotar la significatividad de los residuos. Una forma alternativa de estimar el indicador compuesto es desarrollar una función de producción, los antecedentes en Argentina se vinculan a los respectivos PBG provinciales.

El cuadro 1 sintetiza el estado del arte en cuanto a los indicadores a nivel de Argentina. Como puede observarse las series de indicadores no son homogéneos ni en su construcción, ni períodos considerados. La mayoría se discontinuó o la demora en su publicación es mayor a los dos años, perdiendo con ello su utilidad inicial. Como puede verse las metodologías de estimación existentes, se basan en el Análisis por Series de Tiempo con una importante cantidad de variables, disponibles como series históricas largas, de publicación generalmente restringida a Organismos Oficiales; tal es el caso del EMAE, con enormes equipos de trabajo, resultando costoso, requiriendo mucha información cuantitativa, de lo contrario no se alcanza el Ruido Blanco. La otra fuente de estimación de indicadores son los estudios de los Ciclos Económicos: iniciados por Jorrat (2003) para Argentina y después Tucumán, le siguen varias provincias, los más recientes son en 2014 un Índice Sintético de Actividad Provincial de Muñoz&Trombetta y en 2017 un Indicador Compuesto de Actividad Económica de Misiones (ICAEM) estimado por Heredia&Alvarez.

La importancia del índice que se busca construir se desprende de la necesidad de datos actualizados para la toma de decisiones, y del hecho de que, como se mencionó, actualmente no existe ningún indicador que tenga una frecuencia de publicación mayor a la anual para la provincia (Producto Bruto Geográfico). Debido a esto es que se busca que el índice construido tenga tanto una alta periodicidad, idealmente mensual. Así como que tenga el menor rezago posible en su publicación.

METODOLOGÍA

La metodología propuesta en esta investigación consiste en aplicar Análisis Factorial por Componentes Principales: es un método de reducción de dimensionalidad lineal de los datos. El objetivo es poder reducir el número de variables necesarias para representar un conjunto de datos, seleccionando las que mejor se adapten al fenómeno que se quiere medir. Los requerimientos en cuanto a datos son menores, lo que lleva a conformar equipos de personas más chicos y continuar su publicación de forma sencilla por parte de técnicos. El diagrama 1 sintetiza el proceso de estimación.

Diagrama 1



El método que se utiliza para la construcción del índice es el Análisis de Componentes Principales. A partir de una serie de variables que se seleccionarán en función de su disponibilidad. Las mismas deberán tener una frecuencia buscándose como ideal que sea mensual, con una publicación al

menos semestral. Las variables que finalmente se usen deberán ser representativas de la actividad económica y tener cierto nivel de correlación con el índice construido. Esta metodología si bien no es reciente tiene un uso poco difundido dentro del ámbito de las ciencias económicas, por lo que desde ese punto de vista se puede considerar como innovador su uso para un caso como el presente. Se espera también poner de manifiesto la utilidad que esta técnica puede aportar. Así como interpretar el comportamiento de los datos de formas que otros tipos de análisis no permitirían.

El criterio de validación del índice que se construya será justamente contrastarlo con el EMAE y ver si la relación de comportamiento entre ambos se mantiene al extrapolarlo a la comparación entre PBG provincial y PBI de Argentina.

El procedimiento metodológico implica someter las variables originales a un análisis de comunalidades, extraídos por Componentes Principales; para luego presentar la varianza total explicada desagregada en factores. La selección de componentes factoriales se hace estableciendo un eigen value igual o superior a uno. Para interpretar los resultados se utiliza Análisis Factorial que descompone la varianza total en porciones de varianza de cada uno de los factores, maximizando la misma. Para obtener las puntuaciones de las variables factoriales a partir de las originales, se debe rotar la Matriz de Componentes. A continuación, se desarrolla con mayor amplitud las etapas de estimación.

ANÁLISIS DE COMPONENTES PRINCIPALES

Dado un conjunto de datos de D variables y N observaciones de cada variable, los datos pueden ser representados como un conjunto de N vectores pertenecientes a \mathbb{R}^D :

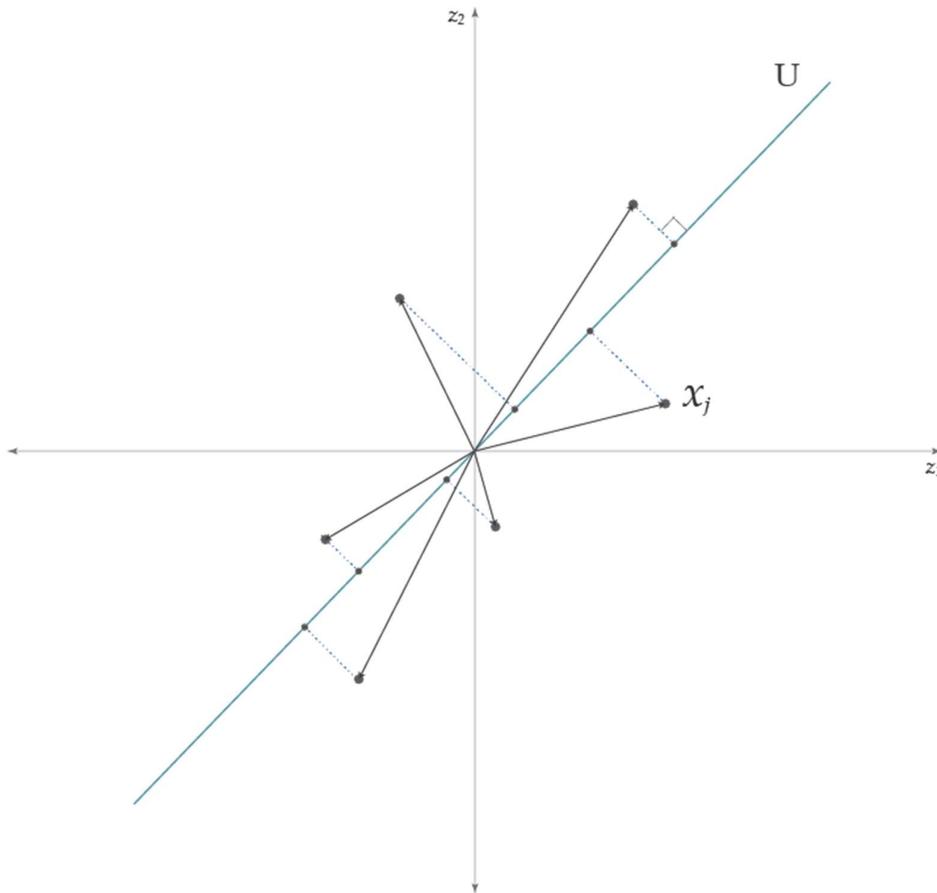
$$X = \{x_1, x_2, \dots, x_N\} \quad x_j \in \mathbb{R}^D$$

Donde cada x_i está formado por una observación de cada una de las D variables iniciales. El análisis de componentes principales busca encontrar una representación con una dimensión menor. Pero que sea lo más parecida posible a X. Para lograr esto lo que se hace es realizar una proyección ortogonal de los datos, ya que esto minimiza el vector diferencia con respecto a los datos originales.

Para ver esto con un ejemplo, se supone que como muestra el gráfico 1 existen dos variables iniciales (D=2) z_1 y z_2 . Se puede representar gráficamente el par de datos como un vector en \mathbb{R}^2 , entonces habrá tantos vectores como cantidad de observaciones haya en las series de datos. El objetivo será entonces encontrar una representación de menor dimensionalidad de este conjunto de datos, en este caso la única dimensionalidad menor a dos es uno, o sea que el objetivo sería ajustar esa serie de vectores a una recta. Existirán entonces dos incógnitas a resolver. La primera de las incógnitas será cómo representar cada uno de los vectores iniciales de tal forma que haya el menor error posible. La segunda será cómo elegir la pendiente de esa recta de tal forma que se minimice el conjunto de los errores. Es decir, minimizar la suma de los segmentos punteados del gráfico 1. Ahora bien, recordando lo visto en proyecciones ortogonales eso da respuesta a la primera de las incógnitas, la forma de representar a cada vector dentro del subespacio U es proyectarlo ortogonalmente, ya que de esta manera se asegura que se minimiza el error de proyección. Queda entonces por resolver el problema de encontrar la pendiente de la recta U¹.

¹ Una acotación metodológica a hacer en este punto, es que, si bien se habla de la recta U, esta es en realidad un subespacio vectorial. Por lo que no puede ser cualquier recta, debe pasar por el origen de coordenadas. La consecuencia práctica de esto es que a la hora de trabajar con ACP los datos deben tener media igual a cero.

Gráfico 1



Cada uno de los vectores x_j puede ser representado como una combinación lineal de los vectores base de \mathbb{R}^D :

$$x_j = \sum_{i=1}^D B_{ij} \cdot b_i \quad (1)$$

Donde b_i es cada uno de los vectores base y B_{ij} es un escalar que representa la coordenada correspondiente al vector b_i . Cada b_i va a tener dimensión $D \times 1$, ya que es un vector base de \mathbb{R}^D . Supondremos para trabajar con ACP que las bases van a ser ortonormales, es decir que $\sum_{i=1}^D b_i^2 = 1$.

$$B_{ij} = x_j^T \cdot b_i \quad (2)$$

La ecuación (2) surge de la deducción de proyección ortogonal para las coordenadas, donde $\|b_i\| = 1$ por ser las bases usadas para ACP ortonormales. $B_{ij} \cdot b_i$ puede ser visto como la proyección ortogonal de x_j en el subespacio unidimensional generado por b_i (es la proyección sobre esa recta o eje). Una aclaración a hacer aquí es que la dimensión inicial de los datos es D (se necesitan D coordenadas para representarlos), y lo que se busca es reducir esa dimensión a M minimizando el error en el proceso. La ecuación (2) muestra cuál va a ser la magnitud de una de esas M coordenadas, más específicamente la coordenada sobre el subespacio unidimensional generado por el vector b_i . Para el ACP se hacen dos supuestos importantes, uno es que los datos deben tener media cero y el otro es que las bases van a ser ortonormales. Para la reducción de datos se requieren los *eigen*

vectores, los cuales serán los vectores base, por lo cual por bases ortonormales serán ortogonales entre sí, y el *eigen* vector con el *eigen* valor más alto, se trata de un vector que apunta en la dirección en la que los datos tienen más variación (y el *eigen* valor va a ser el valor de esa variación).

Para el caso D-dimensional se tiene que:

$$S \cdot b_j = \lambda_j \cdot b_j \text{ con } j = M + 1, M + 2, \dots, D$$

$$J = \sum_{j=M+1}^D \lambda_j$$

Es decir, se minimiza J eligiendo los D-M *eigen* valores de menor valor y, por lo tanto, proyectando los datos en el subespacio formado por los vectores correspondientes a los M *eigen* valores mayores. Algo a remarcar en este punto, es el hecho de que habrán D *eigen* valores (ya que la dimensión de la matriz S es DxD), de los cuales se elegirán los M mayores y se descartarán los D-M con el menor valor.

Si se analizan los resultados obtenidos al proyectar los vectores ortogonalmente sobre el subespacio principal $x_j^T \cdot b_i = \beta_{ij}$. Notando esto matricialmente para el caso de N vectores (recordar que se tendrá N vectores porque hay N observaciones) se obtiene lo siguiente:

$$\begin{matrix} X & \times & B & = & Z \\ \underbrace{N \times D} & & \underbrace{D \times M} & & \underbrace{N \times M} \end{matrix}$$

$$\begin{bmatrix} x_{11} & \dots & x_{D1} \\ \vdots & \ddots & \vdots \\ x_{1N} & \dots & x_{DN} \end{bmatrix} \times \begin{bmatrix} b_{11} & \dots & b_{M1} \\ \vdots & \ddots & \vdots \\ b_{1D} & \dots & b_{MD} \end{bmatrix} = \begin{bmatrix} z_{11} & \dots & z_{M1} \\ \vdots & \ddots & \vdots \\ z_{1N} & \dots & z_{MN} \end{bmatrix}$$

$$z_{11} = x_{11} \cdot b_{11} + \dots + x_{D1} \cdot b_{1D}$$

$$z_{1N} = x_{1N} \cdot b_{11} + \dots + x_{DN} \cdot b_{1D}$$

$$Z_1 = X \cdot b_1$$

Cada una de las N observaciones de la nueva variable z_1 va a ser una combinación lineal de las D variables iniciales, combinada usando como coeficientes los valores del *eigen* vector b_1 . Cada columna de la matriz Z va a ser una de las nuevas M variables, que no es ni más ni menos que una de las M coordenadas vistas desde el punto de vista geométrico de las proyecciones ortogonales. Aquí finaliza el desarrollo del modelo ACP estrictamente hablando. En el siguiente título se analiza el análisis factorial, que va a permitir incorporar como elemento de análisis, la interpretabilidad de los componentes o nuevas variables obtenidas.

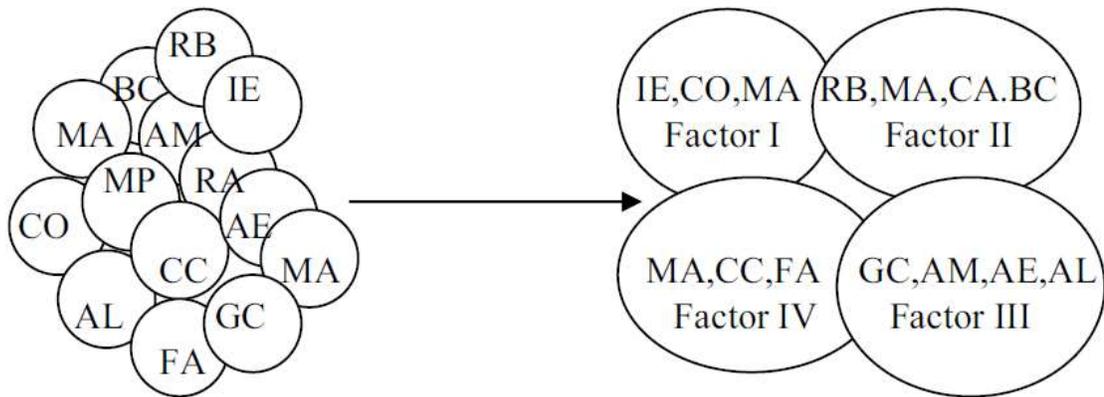
ANÁLISIS FACTORIAL

El análisis factorial da un enfoque diferente del problema de reducción de dimensionalidad. El planteo del que se parte es el siguiente, dado un conjunto de D variables iniciales existen un conjunto M de variables subyacentes o factores que no son observables directamente sino a través de combinaciones de las variables iniciales. Lo que se busca es representar los datos originales con este conjunto de factores, teniendo la menor pérdida de información posible. Al igual que en el análisis de componentes principales lo que se busca es que los factores sean obtenidos de tal forma de que no estén correlacionados entre sí. Esto agrega dos elementos de juicio al análisis realizado hasta acá, uno es el *principio de parsimonia*, por el cual se busca que la cantidad de factores sea lo menor

posible. El otro elemento es el *principio de interpretabilidad*, el cual busca que los factores obtenidos puedan ser interpretables.

En el diagrama 2 se puede ver este análisis, obsérvese el siguiente ejemplo, se quiere analizar la importancia que los consumidores dan a catorce variables que se consideran relevantes para la compra de un automóvil. Estas variables son: reparaciones baratas (RB), amplia gama de colores (GC), interior espacioso (IE), bajo consumo de gasolina (BC), manejabilidad (MA), aspecto moderno (AM), valor de recompra alto (RA), confortable (CO), motor potente (MP), aspecto elegante (AE), cómodo de conducir (CC), atractivo de línea (AL), maletero amplio (MA) y fácil de aparcar (FA). Las catorce variables pueden caracterizarse por cuatro dimensiones subyacentes relacionadas respectivamente con el confort (factor I), con el coste-eficiencia (factor II), con la elegancia (factor III) y con el manejo fácil (factor IV) no observables directamente. En vez de considerar las 14 variables, se simplifican las cosas, de forma que sólo cuatro factores deban considerarse para caracterizar la estructura subyacente de los datos.

Diagrama 2



A continuación, se plantea matemáticamente el modelo factorial:

$$\begin{aligned} x_1 &= l_{11} \cdot F_1 + \dots + l_{1M} \cdot F_M + e_1 \\ &\vdots \\ x_D &= l_{D1} \cdot F_1 + \dots + l_{DM} \cdot F_M + e_D \end{aligned}$$

Se supone a los factores comunes F_1, F_2, \dots, F_M como variables estandarizadas (media cero y varianza unitaria) y que además no están correlacionadas entre sí. Se supone también que la matriz de covarianzas de los factores específicos es una matriz diagonal (factores únicos incorrelacionados entre sí) y tienen media igual a cero ($E[e] = 0$, $E[Fe'] = 0$)

Dado que las variables X son variables tipificadas, su matriz de covarianzas es igual a la matriz de correlación poblacional S, matriz que puede descomponerse de la siguiente forma:

$$\begin{aligned} S &= E(XX') = E((LF + e)(LF + e)') = LE(FF')L' + E(ee') + LE(fe') + E(ef')L' \\ &= LIL' + \Omega + L0 + 0L' = \mathbf{LL}' + \mathbf{\Omega} \end{aligned}$$

Expresado matricialmente se obtiene:

$$\underbrace{\begin{bmatrix} 1 & \dots & \rho_{1D} \\ \vdots & \ddots & \vdots \\ \rho_{1D} & \dots & 1 \end{bmatrix}}_{D \times D} = \underbrace{\begin{bmatrix} l_{11} & \dots & l_{1M} \\ \vdots & \ddots & \vdots \\ l_{D1} & \dots & l_{DM} \end{bmatrix}}_{D \times M} \times \underbrace{\begin{bmatrix} l_{11} & \dots & l_{D1} \\ \vdots & \ddots & \vdots \\ l_{1M} & \dots & l_{DM} \end{bmatrix}}_{M \times D} + \underbrace{\begin{bmatrix} \omega_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \omega_D^2 \end{bmatrix}}_{D \times D}$$

Si se analiza el resultado para el primer elemento de S (que es la varianza de x_1) se tiene que:

$$VAR(x_1) = 1 = \underbrace{l_{11}^2 + \dots + l_{1M}^2}_{h_1^2} + \omega_1^2$$

Donde h_1^2 es el porcentaje de varianza de la variable x_1 explicado por los factores comunes, y se llama *comunalidad* y ω_1^2 es la parte de la varianza de x_1 que la explica su factor específico y se la llama *especificidad*.

Ahora bien, cómo se relaciona este análisis de factores con el análisis de componentes principales. Se parte de la solución expresada matricialmente a la que se llega en el ACP. Es importante prestar atención al cambio en la notación, ya que hasta aquí se usó notación matricial donde cada elemento era una de las N observaciones, la transformación que se hace aquí es pasar a expresarlo vectorialmente, donde cada variable Z_i y X_i es un vector $N \times 1$ que contiene todas las observaciones. Si se tiene esto en cuenta se puede re expresar la solución obtenida como:

$$\underbrace{\begin{bmatrix} x_{11} & \dots & x_{D1} \\ \vdots & \ddots & \vdots \\ x_{1N} & \dots & x_{DN} \end{bmatrix}}_{N \times D} \times \underbrace{\begin{bmatrix} b_{11} & \dots & b_{M1} \\ \vdots & \ddots & \vdots \\ b_{1D} & \dots & b_{MD} \end{bmatrix}}_{D \times M} = \underbrace{\begin{bmatrix} z_{11} & \dots & z_{M1} \\ \vdots & \ddots & \vdots \\ z_{1N} & \dots & z_{MN} \end{bmatrix}}_{N \times M}$$

$$\begin{matrix} N \times 1 & N \times 1 & 1 \times 1 \\ \tilde{Z}_1 = & \tilde{X}_1 \cdot & \tilde{b}_{11} + \dots + X_D \cdot b_{1D} \\ \vdots & & \end{matrix}$$

$$Z_M = X_1 \cdot b_{M1} + \dots + X_D \cdot b_{MD}$$

Ahora bien, también se había llegado a la conclusión de que $\underbrace{Z}_{N \times M} \times \underbrace{B^T}_{M \times D} = \underbrace{\tilde{X}}_{N \times D}$, lo cual notado vectorialmente quedaría:

$$\begin{matrix} \tilde{X}_1 = Z_1 \cdot b_{11} + \dots + Z_M \cdot b_{M1} \\ \vdots \end{matrix}$$

$$\tilde{X}_D = Z_1 \cdot b_{1D} + \dots + Z_M \cdot b_{MD}$$

La matriz formada por todos los coeficientes b_{ij} es entonces la matriz de cargas factoriales que contiene cada una de las correlaciones entre los factores retenidos y las variables iniciales. Para continuar la etapa de interpretabilidad, se hace necesario rotar los componentes, según se explica en la sección que viene.

ROTACIÓN DE COMPONENTES

En el proceso de rotación de componentes lo que se hace gráficamente es girar la dirección de estos nuevos ejes de referencia que van a definir el subespacio sobre el que se proyectarán los datos. Existen 2 tipos de rotaciones, las rotaciones ortogonales que mantienen los ejes perpendiculares, con lo cual las componentes resultantes seguirán teniendo la característica de no estar correlacionadas entre sí. El segundo tipo de rotaciones son las rotaciones oblicuas las cuales sacrifican un poco de esa independencia con el objetivo de obtener una mayor interpretabilidad de las componentes. A su vez dentro de cada tipo hay varios métodos que se pueden usar. En este apartado se profundizará en la rotación Varimax, que es una rotación ortogonal y es la de uso más extendido. Lo que se busca con la rotación de componentes es que cada una de las variables tenga una correlación máxima con uno de los factores (es decir, uno) y cero con el resto de los factores o componentes. De tal forma que esto facilite la tarea de asociar a cada factor como esa “variable subyacente” representada por un grupo de las variables iniciales.

Se supone que existe un conjunto de datos en \mathbb{R}^3 (es decir tres variables iniciales) que se podrían agrupar dentro de un determinado elipsoide, y que se van a representar estos datos en un plano \mathbb{R}^2 (es decir dos componentes). La metodología hasta aquí usada por el ACP es ir eligiendo iterativamente como ejes de cada componente el eje sobre el que el conjunto de datos tenga mayor varianza. En este caso como se va a tener los datos representados en un plano, los nuevos dos ejes serán los dos ejes sobre los que los datos tengan máxima varianza. Los cuales van a coincidir con los dos ejes más grandes de este elipsoide imaginario. Lo que se hace al rotar los componentes es rotar estos dos ejes (manteniéndolos perpendiculares entre sí) sin variar el plano que generan. La consecuencia de esto es que todos los puntos originales se van a seguir representando sobre el mismo plano, pero respecto a unos ejes diferentes. Es decir, solo van a cambiar sus coordenadas. Esto permite que al rotar los componentes no se pierda nada del total de varianza explicada por los componentes en su conjunto. Metodológicamente lo que sucede es que la información perdida por la primera componente² va a ser recogida por la segunda. Se puede generalizar esto diciendo que la información perdida por las primeras K componentes va a ser recogida por las últimas M-K componentes.

Con la rotación de componentes ya se han desarrollado todos los elementos teóricos sobre los temas necesarios para el desarrollo del índice de actividad económica mensual para la provincia de Mendoza.

ÍNDICE DE ACTIVIDAD ECONÓMICA

Actualmente para la provincia de Mendoza el indicador de actividad económica usado es el Producto Bruto Geográfico (PBG), el cual tiene una apertura a nivel de sectores o ramas de actividad con una periodicidad anual; disponible en la página Web de la DEIE, desde 1991 a 2018, sin empalme entre 2003 y 2004. Adicionalmente, existe una serie de variables que a priori se podrían pensar como representativas de la actividad económica, que permitirían estimar con cierta robustez un indicador mensual ajusta a la realidad mendocina. A nivel nacional el índice que existe es el Estimador Mensual de Actividad Económica (EMAE). El cual en muchos casos es usado como *proxy* para la actividad económica provincial por la mencionada falta de disponibilidad de un instrumento mejor. La desventaja de esto es que se pierden ciertas particularidades del comportamiento económico que tiene la provincia y no se dan a nivel país.

² Ya no está en la dirección en la que los datos tienen máxima varianza, por lo tanto, va a disminuir ese nivel de varianza captada.

El concepto adoptado en este trabajo es que un índice compuesto resume en un único número mensual el comportamiento de un grupo de series seleccionadas que se mueven sincrónicamente entre ellas y presentan un comportamiento homogéneo respecto al ciclo económico. En este caso el interés radica en series que se muevan con el nivel de actividad económica, es decir, series coincidentes con el ciclo económico en Mendoza. Una característica importante es que los índices no representan los valores totales, sino el comportamiento. Lo que se busca es que las variaciones producidas en este caso en el nivel de actividad económica, sean captadas por este indicador.

El procedimiento implica una serie de pasos, fundamentales para lograr un indicador consistente de la actividad económica.

- Selección de variables.
- Tratamiento estadístico de las series. Corrección de observaciones atípicas.
- Estimación de los factores o componentes.
- Construcción de un índice único.

A continuación, se hace una descripción de los pasos seguidos en la investigación.

SELECCIÓN DE VARIABLES

La búsqueda de variables se orienta a que las mismas satisfagan algunos requerimientos, en este trabajo adoptan algunos de los criterios de la Conference Board (2001).

- Adecuación Estadística: los datos deben ser recolectados y procesados de una manera estadísticamente confiable. Suavidad: su movimiento de un periodo a otro no puede ser muy errático. Deben ser variables que presenten una fuerte correlación con la actividad económica, medida a través del PBG.
- Disponibilidad o demora de la información: la serie debe publicarse con una periodicidad razonablemente rápida. La información debe ser de acceso público y de periodicidad mensual. Asimismo, es necesario que la frecuencia de actualización no exceda, usualmente, el semestre vencido.
- Los datos para cada una de ellas deben provenir de una misma fuente.

El primer requisito garantiza que los cambios del indicador reflejen efectivamente variaciones en el nivel (inobservable) de actividad económica de la provincia. El segundo responde a la necesidad de mantener un proceso permanente de actualización mensual con una demora razonable en su publicación. El tercero apunta a garantizar la homogeneidad metodológica en la construcción de las series, imprescindible para la posterior obtención y continuidad del mismo en el tiempo.

TRATAMIENTO ESTADÍSTICO DE LAS SERIES

La literatura en general recomienda aplicar un proceso de ajuste estadístico de las variables seleccionadas antes de proceder a la construcción de cualquier índice. La primera razón para ello, es que, por tratarse de series de tiempo, están naturalmente sujetas a variabilidad estacional que es deseable eliminar, ya que el indicador debe funcionar como un proxy de la actividad económica excluyendo cambios de naturaleza puramente estacional. La segunda es que, la volatilidad de las variables empleadas es alta en relación a lo que típicamente se observaría en series de estimadores de actividad. Por lo tanto, se aplican tratamientos por datos con comportamientos irregulares o extremos, con la consiguiente corrección de observaciones atípicas.

Varias de las series consideradas están expresadas en valores nominales (pesos corrientes); para su inclusión en un índice sintético de nivel de actividad tenga sentido económico, todas las series que lo componen deben ser expresadas en términos reales, ya sea moneda constante o unidades físicas. Lo que hace necesario deflactar las mismas utilizando uno o más índices de precios. Este procedimiento supone dificultades, dada la nefasta intervención política del INDEC ocurrida en 2007, que hizo inviable la utilización de los índices de precios oficiales. El sitio InflacionVerdadera.com creado en 2007 para proveer índices de precios alternativos a los oficiales en Argentina, con la posterior creación del Billion Prices Project en el Massachusetts Institute of Technology. Los resultados demostraron la manipulación de las estadísticas oficiales en Argentina, los cuales fueron publicados en un *paper* académico en el Journal of Monetary Economics, en agosto de 2012, según reza la página Web del sitio. En virtud de esta limitación se opta en esta presentación por deflactar todas las series nominales por una combinación de índice de precios. Para el tratamiento de la inflación hasta el año 2006 inclusive el Índice de Precios al Consumidor (IPC) de GBA publicado por el INDEC, luego se empalma con el IPC de la provincia de San Luis desde enero 2007 a diciembre del 2019.

A efectos de describir las variables recopiladas, se las agrupa en Bienes, Servicios y Remuneraciones, detalladas en las tablas y gráficos 1, 2 y 3. Las mismas se consideran representativas para la actividad económica de Mendoza, además de cumplir con los criterios de inclusión de la Conference Board. Luego, a partir de las variables seleccionadas, se aplica el análisis factorial para obtener una cantidad reducida de componentes que las representen. Con los factores estimados se construye un índice sintético de actividad económica de la provincia. La utilidad práctica que se busca con este caso es que, dado que las variables usadas están disponibles con periodicidad mensual el índice resultante también lo será, con lo cual podría servir como un estimador de la actividad económica de la provincia, solo disponible anualmente, con un retraso de dos años. El denominador común de estas variables es que capturan lo que habitualmente se conoce como “el lado de la demanda” de la actividad económica y no de la perspectiva de la oferta. Esto es así, dado que se cuenta con pocos índices de producción; sumado al hecho de ser discontinuos, con cambios metodológicos y una periodicidad anual o apenas trimestral en la mayoría de los casos. Hecha esta aclaración las series compiladas son de consumo en todos los casos de Bienes, tal como, cemento, vino, energía eléctrica, gas natural, combustibles; igual situación en los servicios salvo la recaudación del impuesto a los sellos. Las remuneraciones representarían el precio de equilibrio entre la demanda y la oferta del ángulo productivo.

En este *paper*, se describirán someramente las variables elegidas; se destaca que, la investigación corresponde a un equipo mayor³, que presenta un trabajo en estas mismas Jornadas.

Tabla 1: BIENES

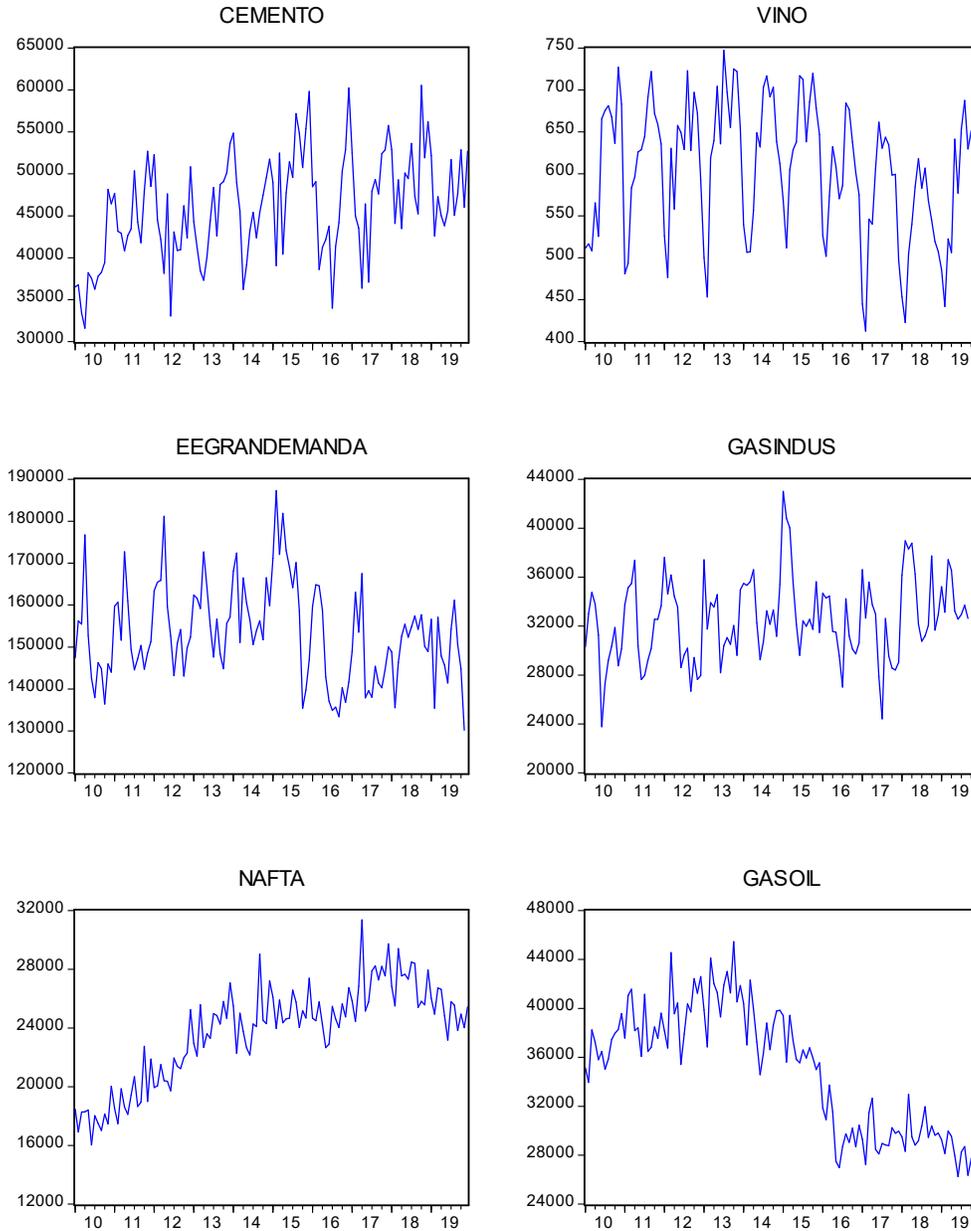
Variable	Descripción
CEMENTO	Consumo de Cemento en Toneladas
VINO	Despachos de vinos autorizados para ser liberados al consumo
EEGRANDEMANDA	Consumo de Energía Eléctrica Grandes Demandas/Industrial en Mwh
GASINDUS	Consumo de Gas Natural de la Industria - En miles de m3 de 9300 kcal.
NAFTA	Ventas estaciones de servicio de Nafta común, súper y premium en mts3
GASOIL	Ventas estaciones de servicio de Gas Oil grado 1, 2 y 3 en mts3

Así, la variable CEMENTO, presenta variaciones importantes a lo largo del período, la estacionalidad en los meses de verano e invierno; la tendencia es prácticamente constante. Los picos más altos se registran en diciembre de 2015 y 2016, luego octubre de 2018, la presidencia de Mauricio Macri se caracterizó por una menor volatilidad, respecto de las gestiones de Cristina Fernández.

³ Se agradece a los alumnos de la carrera de Economía, por su contribución en el análisis descriptivo de las series en cuestión.

En el VINO se observa que los puntos más altos se establecen en la época invernal, esto se debe a que la cosecha de la vid se produce aproximadamente a principios de marzo, para luego comenzar con la elaboración del vino, que concluye en invierno. Los primeros cinco años, con sus picos más altos, se mantuvo relativamente estable, para luego presentar una disminución, a partir del 2016, para finalmente hacia el invierno del 2019 observar un repunte en la producción máxima por estación respecto del año anterior.

Gráfico 1: BIENES



El Consumo de Energía Eléctrica de las Grandes Demandas (EEGRANDEMANDA) muestra picos máximos durante febrero y abril; en julio y octubre, se observan los valles. La tendencia tiene un comportamiento levemente creciente desde el 2010 hasta el 2015, compatibles con tarifas sumamente bajas, luego una caída significativa del consumo hacia finales del año 2015. Éste es un

síntoma negativo de la economía, ya que el consumo de energía eléctrica por parte de las grandes demandas está relacionado con el nivel de producción.

Para el Consumo de Gas Natural de la Industria los datos presentan un patrón económico estable, ya que no se detecta una tendencia positiva o negativa. Se observa la existencia de una componente estacional en verano, debido al funcionamiento de las industrias conserveras.

En la NAFTA hay un crecimiento del consumo, con salvedades en distintos momentos del tiempo. La caída de los precios del petróleo en el 2010, marcó precios bajos de los combustibles, 2014-2015, acelerándose en 2016. En el último período, la caída del consumo de nafta ha sido constante, impulsado por el aumento de su precio, a efectos de alcanzar la competitividad de los precios internacionales, evitando el déficit energético.

Las ventas en las estaciones de servicio de GAS OIL muestran un crecimiento hasta el 2014, sin incluirlo; a partir de allí, se registró una importante caída de litros de gasoil vendidos. El principal motivo de estas caídas es el aumento constante del precio del combustible y el impuesto sobre los Combustibles Líquidos que en algunos períodos se complementa con una menor capacidad de compra por parte de los usuarios. Cada elección, las ventas aumentan. Esto es por los esfuerzos que hacen los distintos gobiernos para que crezca el consumo de bienes, entre ellos los combustibles. Otro factor relevante es el campo. En épocas de inclemencias climáticas, ya sea por heladas o granizo, que afectan la actividad agrícola, el consumo de gasoil sube o cae, respectivamente; por ello, la serie presenta un comportamiento cíclico cada enero y julio.

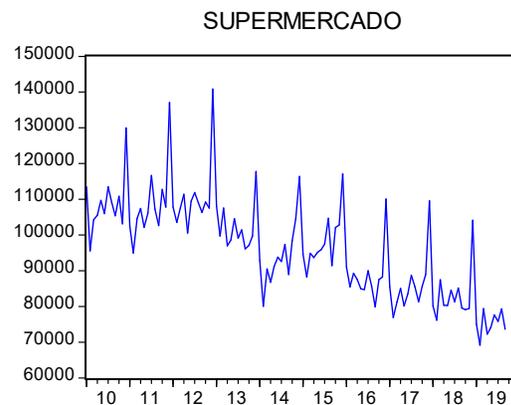
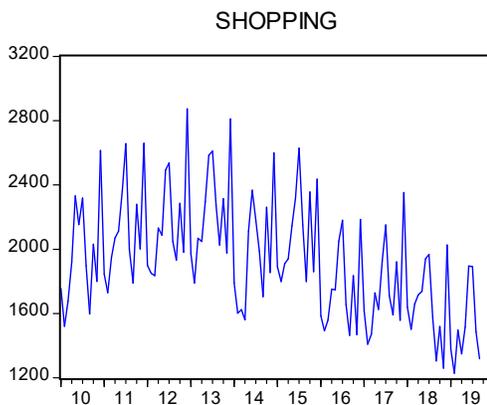
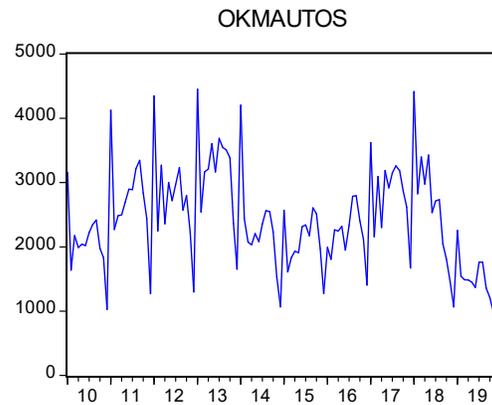
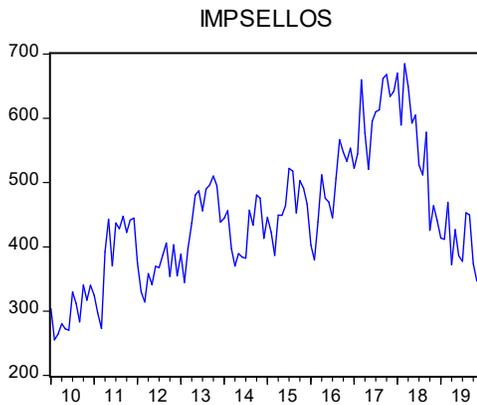
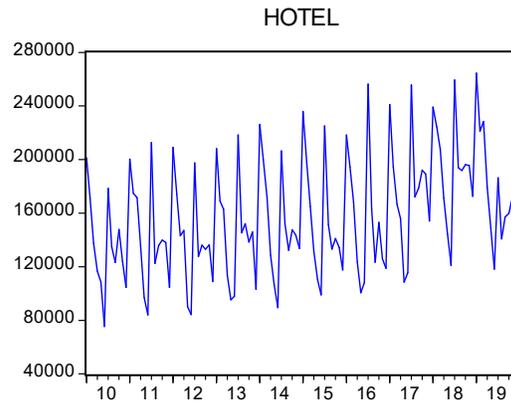
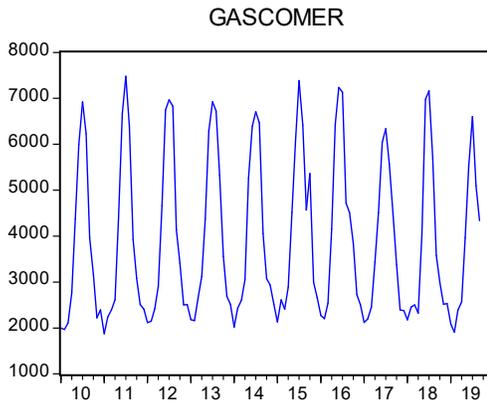
La tabla 2 enuncia los servicios incluidos en el análisis. Se comienza con el Consumo de Gas Natural Comercial. Los datos presentan un patrón estable, ya que no se detecta una tendencia positiva o negativa. Al analizar la serie, solamente a partir de la gráfica y a simple vista, se observa una estacionalidad en el consumo, con máximos en las estaciones de invierno y mínimos en las de verano.

La ocupación de plazas en HOTEL presenta una clara e ininterrumpida tendencia alcista; si se traza una línea imaginaria se ve que, tanto los picos como los valles, alcanzan valores más altos con el paso del tiempo. La serie presenta un patrón repetitivo, los valores mayores en ocupación hotelera, dentro de los doce meses del año se dan en el mes de enero y los más bajos en junio aproximadamente.

La recaudación del Impuesto a los SELLOS se aplica a los contratos de alquiler, transacciones financieras; ventas de inmuebles, de automóviles; y todo tipo de actividad económica. Para todas ellas, se cobra por un porcentaje del 1,5% del monto del contrato, salvo excepciones. Se observa que, la recaudación alcanza valores mayores en 2011, 2013, 2015 y 2017, en cambio en los años pares se muestra una caída. La tendencia es creciente, aunque no muy pronunciada, hasta el año 2017 inclusive, en ese mismo año se llega a un pico máximo de recaudación; luego hay un período de caída o decrecimiento hasta 2019. No presenta evidencia de algún comportamiento repetitivo o estacional.

Tabla y gráfico 2: SERVICIOS

Variable	Descripción
GASCOMER	Consumo de Gas Natural Comercial - En miles de m3 de 9300 kcal.
HOTEL	Demanda hotelera: pernoctaciones equivalente a plazas ocupadas
SELLOS	Recaudación Impuesto a los Sellos en millones de pesos constantes dic2019
OKMAUTOS	Ventas de Automotores Cero Km Particulares, Comercial Liviano y Pesado
SHOPPING	Índice de ventas en centros de compras a valores constantes del año 2004
SUPERMERCADO	Índice de ventas en supermercados a valores constantes del año 2004



Las Ventas de Automotores Cero Km presentan distintos ciclos en las cantidades vendidas en los últimos diez años. Entre el 2010 y el 2012 se aprecia una tendencia creciente, para luego permanecer estable hasta el 2014. En el 2016, las ventas comienzan a crecer hasta 2018, siguiendo una tendencia decreciente hasta el 2019. La estacionalidad se da todos los eneros, esto se verifica porque la gente compra en diciembre, pero patenta los vehículos al año siguiente, evitando pagar todo período fiscal de impuesto automotor.

Los Índices de ventas en centros de compras (SHOPPING) y en supermercados (SUPER) presentan una tendencia algo creciente en los primeros años observados, hasta 2013; mientras que luego decrece en forma constante. La Economía argentina se desaceleró en 2014, Mendoza no escapó de ese

efecto; la desaceleración se debió a factores tanto internos como externos. En 2013, hubo un aumento de la inflación, producida por un aumento en la emisión monetaria y por una sobrecontratación de empleados públicos. También al existir el Cepo Cambiario, las inversiones del exterior disminuyeron, lo que se ve reflejado en la disminución en las ventas. La agudización de las tensiones cambiarias en el primer trimestre, que dieron lugar a una brusca devaluación, y al consiguiente impacto sobre el desempeño durante el resto de 2014. Los segundos corresponden entre otros, a las crisis mundiales del euro, los bancos, las elecciones presidenciales en EE.UU., Francia, Japón y cambio de guardia en China; la desaceleración de la Economía china; sin olvidar, el fallo desfavorable de la Corte Suprema de los Estados Unidos en la disputa del país con los “fondos buitres”. En ambas series la estacionalidad se da todos los diciembres, con las fiestas de fin de año y los aguinaldos.

El análisis de las remuneraciones requirió transformar las series, dividiendo cada una de ellas con la remuneración promedio total de Mendoza. Esto se hizo a efectos de salvar el fuerte problema inflacionario y el pago de los aguinaldos, que cada seis meses traen una perturbación no estocástica. De esta manera, los salarios sectoriales están expresados en moneda constante. Por lo cual, los mismos no corresponden a valores monetarios, sino a índices porcentuales respecto de la remuneración promedio de la economía como un todo. Dicho esto, se describen cada una de las percepciones salariales.

Las remuneraciones del sector agropecuario de Mendoza tienen una tendencia decreciente hasta abril de 2013, sin recuperarse en todo el período analizado. El valor porcentual respecto del total es muy inferior a cien, con apenas un pico que no llega a sesenta, es decir, las percepciones del hombre que trabaja el campo son sumamente bajas. El efecto estacional se puede notar cada doce meses, entre febrero y marzo de cada año, con picos bajos, se especula que esto se debe a un aumento de la oferta laboral por parte de los trabajadores golondrina, en los meses fuertes de cosecha.

La variable remuneración promedio de los trabajadores registrados en el sector privado de la Construcción es menor al valor 100, es decir, el salario de los trabajadores de Construcción se encuentra por debajo del salario promedio de la economía. En el primer período hasta el año 2012, presenta una tendencia creciente, esto puede deberse a que, el índice de construcción aumentaba y hubo arreglos salariales. Luego, en el segundo período 2013-2015, se nota una tendencia decreciente que se sitúa por debajo del salario promedio. A partir del 2016, se puede ver una tendencia constante, esto puede deberse a que por más de que el índice de construcción sufrió de variaciones en esos años, la remuneración de la construcción sobre la remuneración promedio se mantuvo constante. Si bien, ciertos valores sobresalen, no se puede identificar un comportamiento estacional.

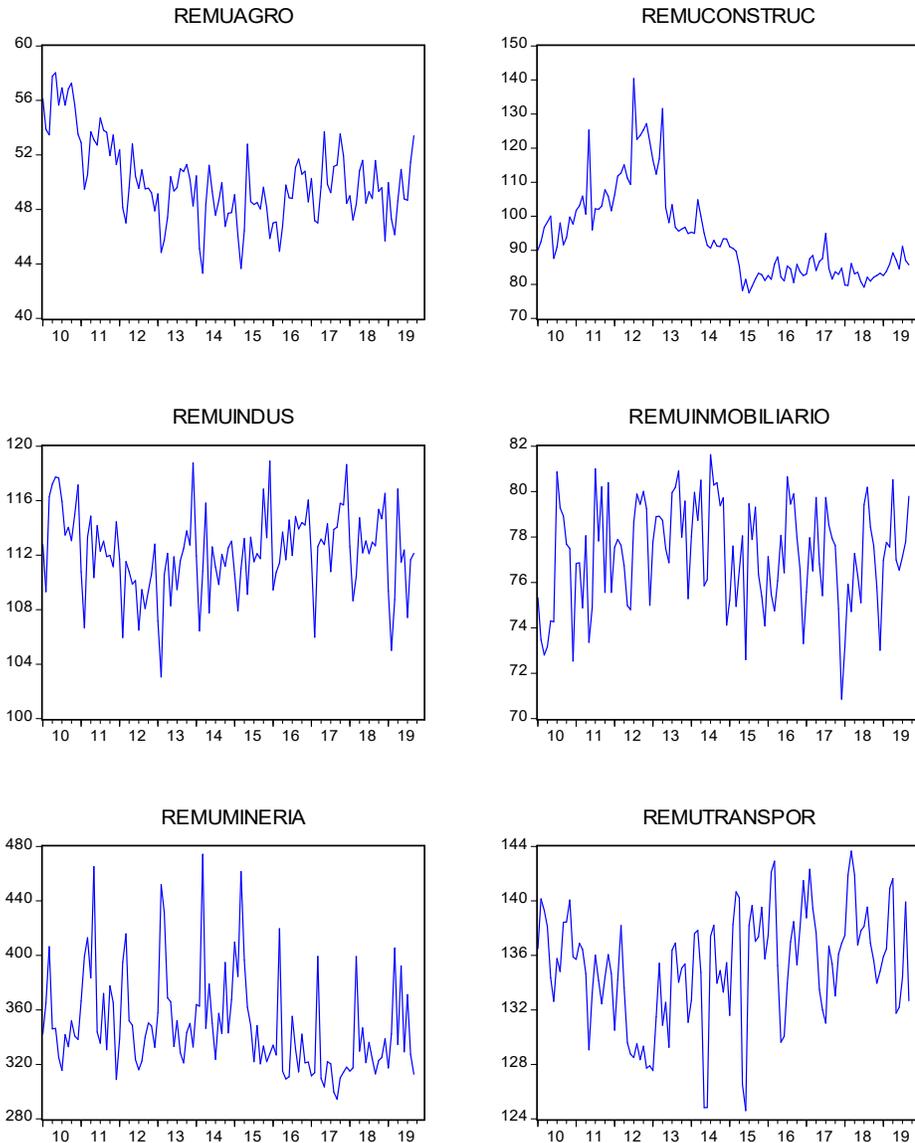
Tabla 3: REMUNERACIONES

Variable	Descripción: remuneración promedio de los trabajadores registrados del sector privado sectorial respecto a la remuneración promedio total de Mendoza, expresado en porcentaje
REMUAGRO	Remuneración Agricultura a dos dígitos Código CIU
REMUCONSTRUC	Remuneración Construcción a dos dígitos Código CIU
REMUINDUS	Remuneración Industria Manufacturera a dos dígitos Código CIU
REMUINMOBILIARIO	Remuneración Inmuebles a dos dígitos Código CIU
REMUMINERIA	Remuneración Petróleo, Minas y Canteras a dos dígitos Código CIU
REMUTRANSPOR	Remuneración Transporte, Almacenamiento y Comunicaciones a dos dígitos Código CIU

Las remuneraciones promedio del sector industria de Mendoza, definiendo al sector industrial como el conjunto de actividades que implican la transformación de materias primas a través de procesos productivos. Los salarios del sector industria son en promedio mayores que los salarios del resto de

los sectores. La serie muestra una tendencia decreciente, hasta el 2013; luego, le sigue una tendencia constante, sin cambios hasta la actualidad. La estacionalidad presente se debe a que en el mes de diciembre, se pagan aguinaldos, salario extra que corresponde al 50% de la mayor remuneración mensual percibida; al ser las remuneraciones mayores que el promedio de la sociedad, en dicho mes se potencia la diferencia salarial.

Gráfico 3: REMUNERACIONES



Las remuneraciones del sector inmobiliario permiten distinguir tres periodos: el primero que va desde enero de 2010 hasta mayo de 2014 se ve una tendencia creciente de las remuneraciones; el segundo, mayo del 2014 hasta enero de 2018 se ve un decaimiento de los datos. El tercer periodo comprendido entre enero de 2018 y septiembre del 2019 una tendencia nuevamente creciente. Se puede ver que los datos siempre son menores a 100, es decir que en el sector inmobiliario las remuneraciones en promedio son menores a las remuneraciones promedios del país. En cuanto, a la estacionalidad, no se verifica ningún patrón claro.

La remuneración promedio a los trabajadores del sector minería y extracción de petróleo y gas natural, se encuentra notablemente por encima de 100, es decir que las remuneraciones de dicho sector son muy superiores a la remuneración promedio, para ser específicos, este sector remunera al menos tres veces y media más al del promedio de Mendoza. La serie presenta una tendencia decreciente, es decir que la remuneración tiende a caer levemente con el paso del tiempo. No se puede comprobar un componente estacional sistemático.

La remuneración promedio de los trabajadores registrados en el sector privado de transporte, de almacenamiento y de comunicaciones sobre la remuneración promedio de Mendoza, son siempre mayor a 100. Esto quiere decir que, por lo menos por el período comprendido, los trabajadores de estos sectores son remunerados en mayor medida que el promedio de la provincia. El comportamiento de la variable muestra tendencias decrecientes en el subperíodo comprendido entre enero de 2010 y enero de 2013; luego muestra una tendencia creciente hasta aproximadamente enero de 2016; a partir de ese punto, la variable oscila alrededor de un valor constante hasta septiembre del 2019. Cada doce meses la variable muestra comportamientos similares. Esto es explicado porque en enero las remuneraciones por el pago de aguinaldos son mayores. Téngase en cuenta que cómo se remunera en mayor medida a los sectores en cuestión respecto al promedio, también el aguinaldo será mayor para estos sectores, generando picos cada doce meses.

ESTIMACIÓN DE LOS FACTORES O COMPONENTES

En esta sección se desarrollará el Análisis por Componentes Principales (ACP), a partir de las variables, agrupadas como Bienes, Servicios y Remuneraciones, ya descritas en el punto anterior. La selección de factores, se basa en considerar aquellos cuyos *eigen value* sea igual o superior a uno. La reducción de factorial de los datos, se hizo rotando los ejes por procedimiento Varimax, buscando cada una de las variables originales tenga una correlación máxima con uno de los factores (es decir, uno) y cero con el resto de los componentes. De tal forma que, esto facilite la tarea de asociar a cada factor con esa “variable subyacente” representada por un grupo de las variables iniciales. Otra de las opciones del software, es seleccionar puntuar los factoriales estimados. Para esta investigación se opta dar las puntuaciones por el método de regresión lineal.

Las etapas llevadas adelante son:

- Inicialmente se realizar un ACP sobre el conjunto completo de variables.
- Observar los resultados de la matriz de componentes rotados, para evaluar la conformación de los distintos factores. Si algunas de las variables iniciales, tienen baja ponderación en todos los factores, se puede eliminar. Otro caso es cuando una variable original pesa igual en dos o más factores, sin marcar una preponderancia significativa.
- Eliminar aquellas variables que no cumplan la condición de significatividad, de acuerdo a lo antes citado, como criterio de selección.
- Correr nuevamente la estimación, revisando los criterios ya vistos.
- El proceso termina cuando cada variable, en la matriz de componentes rotados, tiene una correlación superior a 0,60 en un factor y menos de ese valor en los otros componentes que hubieran surgido.

Cuando se incluyeron la totalidad de las variables, estas mostraban factoriales, con bajas ponderaciones, o representadas en varios factoriales. Por lo que se eliminan cinco variables, resultando en cuatro factores significativos. Los resultados que se muestran en la tabla 4.

Tabla 4

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4,171	32,086	32,086	4,171	32,086	32,086	2,735	21,040	21,040
2	2,399	18,454	50,539	2,399	18,454	50,539	2,580	19,843	40,883
3	1,618	12,445	62,984	1,618	12,445	62,984	2,290	17,614	58,497
4	1,051	8,083	71,067	1,051	8,083	71,067	1,634	12,569	71,067
5	,861	6,625	77,692						
6	,690	5,309	83,001						
7	,569	4,380	87,381						
8	,452	3,477	90,857						
9	,442	3,396	94,254						
10	,331	2,545	96,798						
11	,217	1,667	98,466						
12	,126	,971	99,437						
13	,073	,563	100,000						

Método de extracción: Análisis de Componentes principales.

Matriz de componentes

	Componente			
	1	2	3	4
CEMENTO	,401	,373	,473	,479
VINO	-,580	,524	,098	-,094
GASINDUS	,583	-,559	,213	-,045
NAFTA	,693	,485	,389	-,113
EEGRANDEMANDA	,089	-,582	,437	-,399
GASCOMER	-,398	,648	-,138	-,405
HOTEL	,625	-,091	-,064	,135
IMPSELLOS	,587	,626	,129	-,028
SHOPPING	-,667	,238	,545	,168
SUPERMERCADO	-,726	-,181	,387	,453
REMUAGRO	-,510	,025	-,610	,307
REMUCONSTRUC	-,623	-,420	-,288	-,080
REMUTRANSPOR	,597	-,114	-,257	,350

Método de extracción: Análisis de componentes principales.

a. 4 componentes extraídos

Matriz de componentes rotados

	Componente			
	1	2	3	4
CEMENTO	,784	-,229	,274	-,100
VINO	,037	,710	,276	-,217
GASINDUS	,067	-,838	-,157	,512
NAFTA	,870	-,012	-,273	,218
EEGRANDEMANDA	-,184	-,165	,074	,794
GASCOMER	,047	,836	-,153	-,193
HOTEL	,290	-,489	-,311	,028
IMPSELLOS	,779	,083	-,332	-,082
SHOPPING	,089	,471	,774	,010
SUPERMERCADO	-,254	,083	,914	-,090
REMUAGRO	-,511	,102	,078	-,670
REMUCONSTRUC	-,508	,157	,525	,293
REMUTRANSPOR	,205	-,617	-,298	-,221

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 7 iteraciones.

Matriz de transformación de las componentes

Componente	1	2	3	4
1	,566	-,571	-,568	,177
2	,640	,615	-,120	-,445
3	,485	,063	,614	,620
4	,187	-,541	,535	-,622

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

Los resultados del output de la tabla 4, dejan cuatro componentes principales. Al observar la Matriz de Componentes Rotados algunas de las variables tienen una correlación con el factor dominante inferior a .600. en toda la fila; tal el caso de HOTEL, REMUCONSTRUC; lo que permite reducir la cantidad de variables input, a efectos de lograr una reducción factorial. El proceso que se hace implica probar sacar algunas variables en base al peso que tuvieron en los componentes principales rotados, descartando las de menor peso relativo, o aquellas que ponderan de similar manera en dos de las variables factoriales; esto se hace hasta encontrar lo que se considera un óptimo. El resultado final de este proceso de iteración es el que se muestra en la tabla 5.

Tabla 5

Varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,289	32,895	32,895	3,289	32,895	32,895	2,517	25,166	25,166
2	2,053	20,530	53,425	2,053	20,530	53,425	2,207	22,073	47,239
3	1,356	13,556	66,981	1,356	13,556	66,981	1,974	19,742	66,981
4	,946	9,465	76,446						
5	,703	7,031	83,477						
6	,621	6,213	89,690						
7	,419	4,189	93,879						
8	,334	3,335	97,214						
9	,201	2,011	99,225						
10	,077	,775	100,000						

Método de extracción: Análisis de Componentes principales.

Matriz de componentes^a

	Componente		
	1	2	3
CEMENTO	,406	,381	,627
GASINDUS	,576	-,604	,265
NAFTA	,690	,488	,308
EEGRANDEMANDA	,054	-,617	,370
GASCOMER	-,387	,621	-,301
HOTEL	,638	-,032	-,038
IMPSELLOS	,600	,630	,122
SHOPPING	-,693	,283	,541
SUPERMERCADO	-,738	-,132	,501
REMUTRANSPOR	,610	-,212	-,098

Método de extracción: Análisis de componentes principales.

a. 3 componentes extraídos

Matriz de componentes rotados^a

	Componente		
	1	2	3
CEMENTO	,124	,816	,150
GASINDUS	-,350	,116	,794
NAFTA	-,282	,854	-,016
EEGRANDEMANDA	,116	-,124	,701
GASCOMER	,184	-,020	-,769
HOTEL	-,520	,315	,199
IMPSELLOS	-,312	,781	-,254
SHOPPING	,906	,094	-,151
SUPERMERCADO	,867	-,202	,144
REMUTRANSPOR	-,557	,158	,302

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 5 iteraciones.

Matriz de transformación de las componentes

Componente	1	2	3
1	-,772	,559	,303
2	,119	,595	-,795
3	,625	,577	,526

Método de extracción: Análisis de componentes principal:

Método de rotación: Normalización Varimax con Kaiser.

CONSTRUCCIÓN DEL ÍNDICE

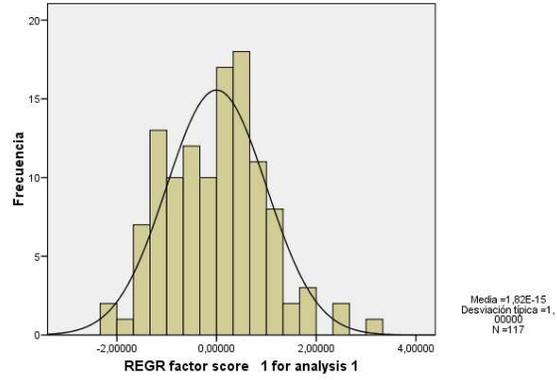
El output de la tabla 5, deja diez de las variables originales agrupadas en tres componentes:

- FAC1_1: pesan preponderantemente HOTEL; las ventas de SHOPPING; SUPER y las remuneraciones al sector transporte, almacenamiento y comunicaciones (REMUTRANSP).
- FAC2_1: ponderan alto el CEMENTO, las ventas de NAFTA y la recaudación del impuesto a los SELLOS.
- FAC3_1: tienen mayor peso el consumo de Gas Industrial (GASIND), de Energía Eléctrica de las Grandes Demandas (EEGRANDEMANDA) y de Gas Comercial (GASCOMER).

Estas tres variables factoriales se usan como representativas del nivel de actividad mensual. Dados estos dos factores obtenidos, cuyos histogramas están representados en los histogramas 1 al 3.

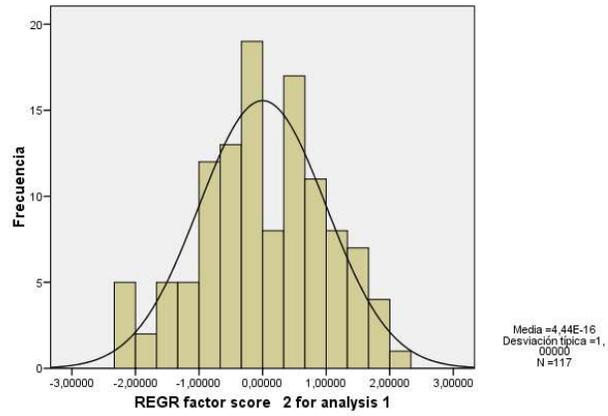
Histograma 1

REGR factor score 1 for analysis 1



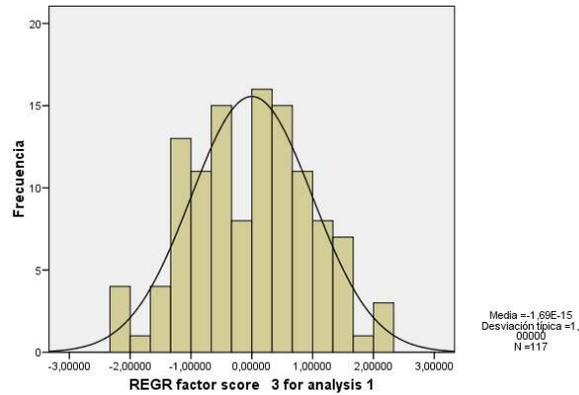
Histograma 2

REGR factor score 2 for analysis 1



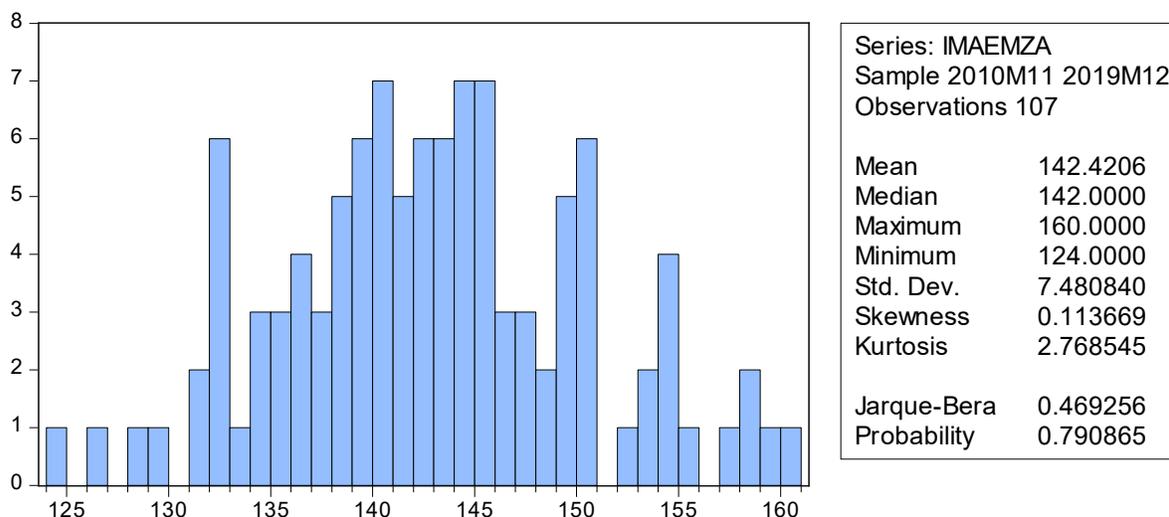
Histograma 3

REGR factor score 3 for analysis 1



El paso siguiente es combinarlos algebraicamente sumándolos, para así obtener el índice de actividad que se estaba buscando, se lo denomina como IMAEMZA. Luego se realiza una transformación al mismo, haciéndolo positivo; generando a partir de esto, un índice con base en 100, a efectos de compararlo con el EMAE. Como ya se señaló en la sección descriptiva de las variables, la mayoría de ellas tiene un comportamiento disímil para el año 2010, y los cambios de tendencia se producen períodos de elecciones presidenciales. Esto justificó reducir el tamaño de la serie de tiempo de noviembre de 2011 a setiembre de 2019.

Histograma 4



El histograma 4 y las estadísticas descriptivas que lo acompañan muestran la normalidad del índice construido, aceptando el test de Jarque-Bera de distribución Normal, Como se puede observar la serie tiene una media igual a 142,4206 y un desvío estándar de 7,48084. En lo que respecta a la distribución de la variable, el gráfico podría asemejarse a la forma de campana, sigue una distribución normal tiene una asimetría igual a cero y un valor de curtosis igual a 3. Para el caso de IMAEMZA, se puede apreciar que el valor del coeficiente de asimetría (Skewness) es igual a 0,113669; mientras que el valor de la curtosis (Kurtosis) es de 2,768545. Este último valor difiere del valor deseado (tres). Para ver si esta discrepancia es significativa se realizará el test de Jarque-Bera, cuyo estadístico contempla en su cálculo al valor de la asimetría y al de curtosis. Dicho estadístico, para determinar si es o no significativo, se compara con el valor crítico que arroja el cuantil de la distribución Chi-cuadrado con dos grados de libertad. De esta manera, el valor obtenido del estadístico es el siguiente

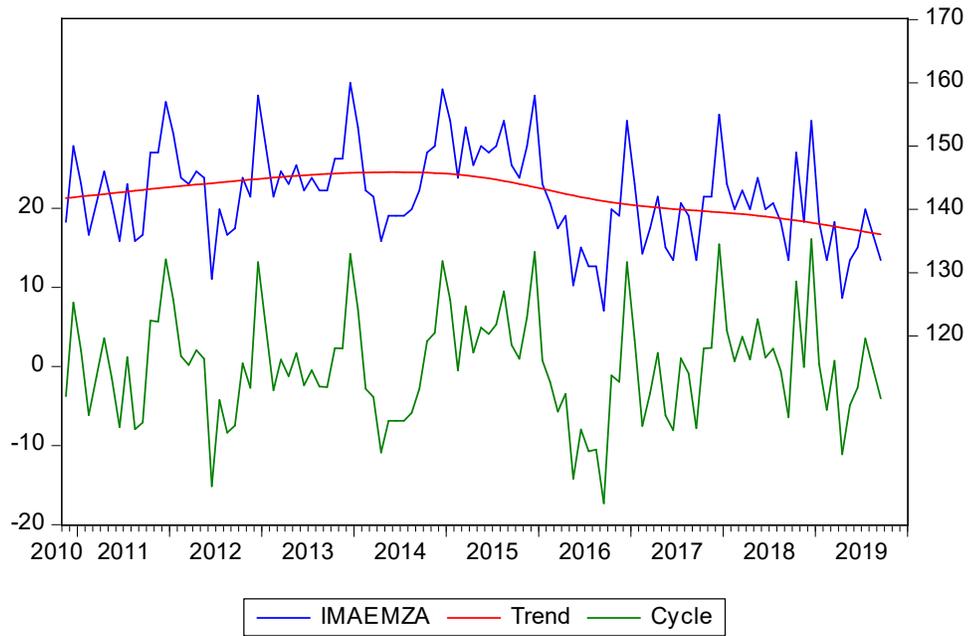
$$JB = 0,469256 < \chi^2(2) = 5,9915$$

Es decir, el estadístico utilizado para el test de Jarque-Bera es menor que el cuantil Chi-cuadrado con 2 grados de libertad con un nivel de significación del 5%. Luego se acepta la hipótesis nula de normalidad.

A efectos de evaluar el comportamiento tendencial del índice, se filtra la variable con el método de Hodrick-Prescott. El gráfico 4 muestra una tendencia creciente hasta el año 2014, para luego decrecer suavemente. Las fluctuaciones son muy marcadas en el ciclo anual, producto de la estacionalidad intrínseca en las distintas variables iniciales.

Gráfico 4

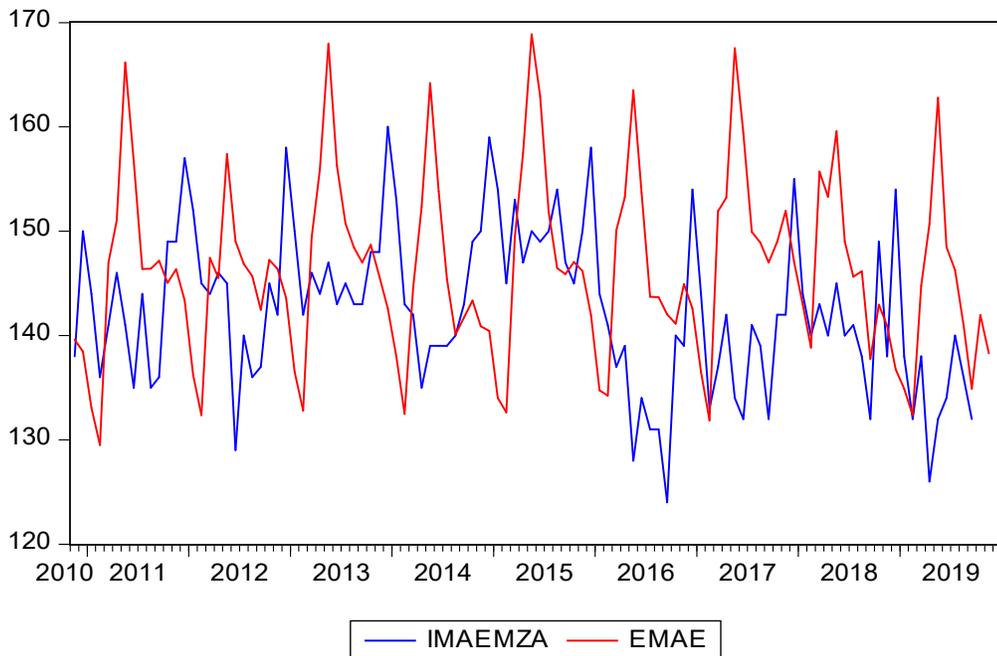
Hodrick-Prescott Filter (lambda=14400)



INTERPRETACIÓN DEL INDICADOR CONSTRUIDO

Ahora bien, para probar la validez del índice se lo va a comparar respecto al EMAE en el mismo periodo de tiempo. Para que esta comparación tenga sentido lo que se hizo fue estandarizar el EMAE como se muestra en el gráfico 5.

Gráfico 5



Las fluctuaciones en las series IMAEMZA y EMAE, que aparecen en el gráfico 5, parecieran presentarse en sentido contrario; esto es, cuando una sube, la otra variable cae. Para contrastar esto estadísticamente, se estimarán las correlaciones de primer orden de EMAE con el índice IMAEMZA rezagada cinco períodos y adelantadas también cinco meses. Como criterio para probar el aparente comportamiento se seguirán algunos criterios utilizado por Heredia&Alvarez (2017). En este paso se busca conocer cuál es el *timing* y conformidad de las series con respecto al EMAE, por lo que se procede analizar las correlaciones (ρ) entre las variables componentes (rezagadas y adelantadas).

La serie será:

- Contracíclica si $\rho(t + i) < 0$
- Acíclica si $0 \leq \rho(t + i) \leq 0.2$
- Procíclica si $\rho(t + i) \geq 0.2$

Luego teniendo en cuenta la intensidad de su correlación se considera procíclica débil si $0.2 \leq \rho(t + i) \leq 0.5$, y se considera procíclica fuerte si $\rho(t + i) \geq 0.5$. Por otro lado, se considera que un indicador es adelantado, rezagado o coincidente. Se considera adelantada cuando la correlación es máxima en $t < 0$, coincidente cuando el máximo se encuentra en el momento $t = 0$, y rezagada cuando la correlación es máxima en $t > 0$.

Entonces, para la validación del indicador sintético compuesto, se realiza un análisis de correlación cruzada entre estos y la serie de referencia, EMAE. Este procedimiento se observa en la tabla 6, donde aparece para el índice un comportamiento contracíclico, dado que el ρ observado es de -0,12 menor al valor crítico. Se trata, además, de un índice con un *timing* adelantado, puesto que la correlación mayor a 0,5 se da en el momento -5, con una intensidad fuerte dado que el valor del coeficiente estimado de correlación es de 0,60. Se interpreta que la Economía de una provincia, muestra signos de recuperación un tiempo antes de hacerlo a nivel nacional; lo mismo sucede, al contrario, cuando se ingresa a una recesión económica, las regiones menores, sufren antes que el país. En el caso de Mendoza, se adelanta alrededor de un cuatrimestre. En el caso de las tasas de crecimiento de IMAEMZA, se muestra como acíclico en el momento 0 y una fuerte correlación de índice con el EMAE, en el momento -5, de las correlaciones, afirmando la asociación de “adelanto” existente, en cuanto al *timing*.

Tabla 6

Correlación EMAE respecto a	-5	-4	-3	-2	-1	0	1	2	3	4	5
Índice IMAEMZA	0,60	0,37	0,12	-0,19	-0,25	-0,12	-0,09	0,00	-0,06	0,06	0,22
Tasas crecimiento IMAEMZA	0,65	-0,05	0,09	-0,32	-0,17	0,10	-0,09	0,20	-0,23	-0,04	0,08

Fuente: Elaboración propia.

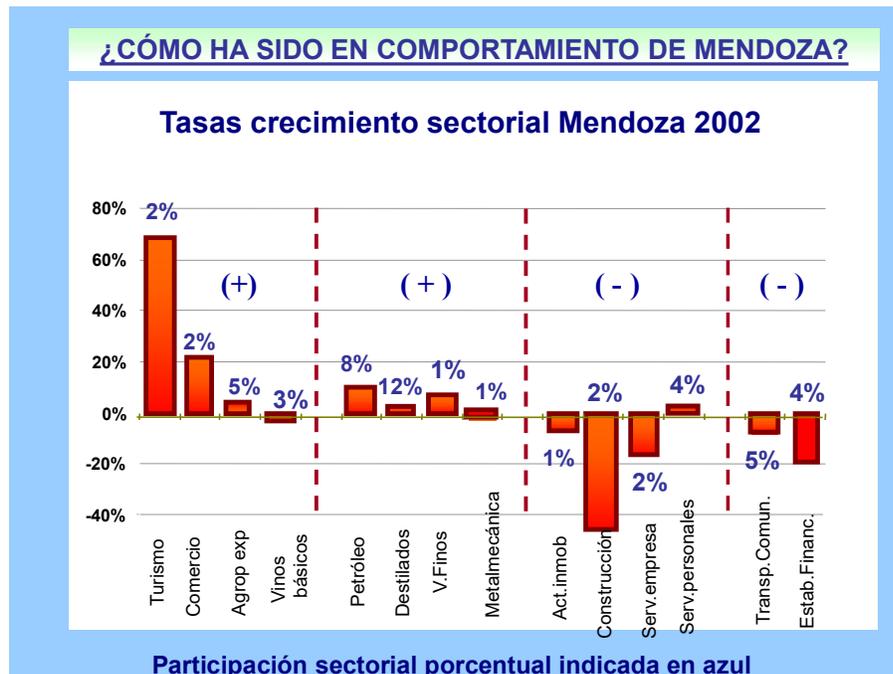
Esto es congruente, con la idea de que Mendoza, depende de las políticas macroeconómicas del país; sin poder tomar medidas de base fundamentales que, prevengan las caídas o pueda aprovechar las cimas, durante el auge en el ciclo. Esto, se podría asociar con la transabilidad de los bienes y servicios. El diagrama 3 ilustra el concepto de transables y no transables, en asociación con la intensidad en el uso de la mano de obra y del capital, como factores de la producción. Los sectores “ganadores” de una economía están asociados con la posibilidad de ser bienes o servicios transables; además, en mejor posición si son intensivos en mano de obra, cuando los salarios reales son relativamente baratos respecto del insumo capital. En cambio, si se trata de bienes o servicios no transables, su recuperación es lenta, y más aún, si son intensivos en capital.

Diagrama 3

	Intensivos en TRABAJO	Intensivos en CAPITAL
Transables	<ul style="list-style-type: none"> •Turismo y vinculadas •Comercio ** •Agropecuario •Vinos básicos <p>(+)</p>	<ul style="list-style-type: none"> •Petróleo •Derivados del petróleo •Vinos finos y especiales •Metalmeccánica <p>(+)</p>
No transables	<ul style="list-style-type: none"> •Act. Inmobiliaria •Construcción •Serv a empresas •Serv personales <p>(-)</p>	<ul style="list-style-type: none"> •Transportes, Almacén y Comunicaciones •Establecimientos Financieros y Seguros <p>(-)</p>

A modo ilustrativo, el diagrama 4 representa las tasas de crecimiento sectorial del año 2002, “el día después de la crisis del 2001”, donde Mendoza reaccionaba positivamente en el mes de noviembre de 2002, mientras que, a nivel nacional, los efectos de la recuperación después de la recesión, se hizo sentir en febrero de 2003. Esto da una idea amplia, de lo expresado en cuanto a la baja versatilidad provincial, frente a las crisis sufridas en las dos últimas décadas. Los sectores de la Economía mendocina, han reaccionado a la suba o a la baja, por lo menos en cuatrimestre antes, de la reacción nacional.

Diagrama 4



Fuente: Informe de coyuntura presentado en FCE-UNCuyo, noviembre de 2002.

En general, las provincias, se componen en menor proporción de bienes transables; por lo cual, no tienen muchos sectores “favorecidos” con las devaluaciones recurrentes del tipo de cambio: como los exportadores; los que comercializan productos agropecuarios, tales como el ajo, la uva; mineros o industrializados, a través de los vinos finos y especiales, las frutas desecadas. El caso del petróleo, al ser intensivo en capital, ha sufrido con las continuas crisis del tipo de cambio y la falta de inversión sistemática.

Mendoza, también enfrenta obstáculos importantes al momento de conseguir inversiones extranjeras, dado que no pueden ofrecer flexibilización laboral o impositiva, a nivel de Ganancias o IVA.

A efectos de probar que el adelanto no supera el cuatrimestre, se calculan las tasas de crecimiento anuales del Producto Bruto Interno de Argentina y el Producto Bruto Geográfico de Mendoza desde el año 1946 al 2018⁴. La tabla 7 muestra las correlaciones de primer orden entre el PBI de Argentina y el PBG de Mendoza; la asociación entre ambos es muy alta, en todos los casos; poniendo en evidencia la dependencia que existe entre las economías regionales con la nacional. En cambio, al observar las tasas de crecimiento, se ve que la Economía de Mendoza tiene un comportamiento procíclico, con intensidad moderada y un *timing* coincidente.

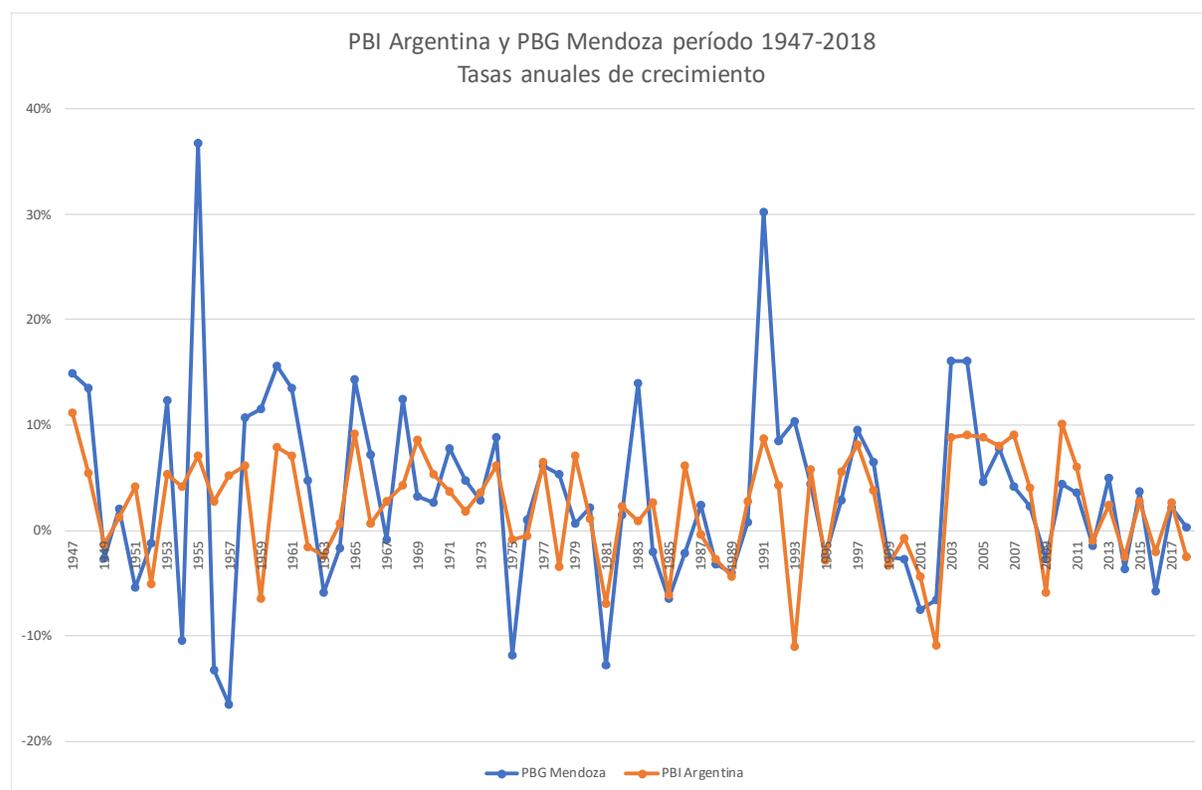
Tabla 7

Correlación PBI ARG respecto a	-5	-4	-3	-2	-1	0	1	2	3	4	5
PBGMZA	0,94	0,94	0,95	0,95	0,96	0,96	0,94	0,92	0,91	0,90	0,89
Tasas crecimiento PBGMZA	0,04	-0,23	0,15	-0,09	0,09	0,46	0,06	-0,17	-0,13	-0,06	0,05

Fuente: Elaboración propia.

El gráfico 6 remarca que, las variaciones entre el PBI de Argentina y el PBG de Mendoza se acompañan en general, aunque las de la provincia, son las intensas en sus variaciones. Esta misma situación se observa en el gráfico 5, al comparar EMAE con el indicador estimado en este trabajo, IMAEMZA.

Gráfico 6



⁴ Fuente: Armado y consistencia de las series por parte del Equipo para el Convenio de Cálculo y Metodología del PBG de Mendoza. Se utilizaron diversas fuentes de cálculo, surgidas especialmente de la DEIE y MECON.

CONSIDERACIONES FINALES

Para concluir el presente trabajo se expondrán algunas conclusiones sobre el tema y aspectos teóricos para resaltar. Un punto importante es la posibilidad de analizar la matriz de correlaciones S de un cierto conjunto de variables y a partir de esta, determinar a priori si puede ser potencialmente útil el ACP sin una pérdida de información significativa. Se supone que existen dos variables iniciales, para un análisis gráfico simple. El primero de los casos que se puede dar es el de alta correlación entre las variables. Entre más correlacionadas estén las variables (si se las graficara) la nube de puntos formada se parecería más a una recta, por ende, menor sería el error de aproximar los datos a dicha recta. Si, por ejemplo, la nube de puntos estuviera alrededor del eje X, es decir la recta tendría pendiente igual a cero. En este caso lo que está pasando es que la variable Y tiene muy poca varianza. El resultado del ACP para este caso va a ser que el subespacio principal va a ser el eje X, cuyo vector base es $b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Esto quiere decir que el coeficiente que le se está asignando a los datos de la variable Y es cero, se está eliminando esta variable. Este es el segundo caso de utilidad, cuando haya variables con varianza muy bajas van a tener coeficientes muy cercanos a cero y $J \cong var_y$ (en el caso extremo J es exactamente igual a la varianza de Y, la perdida que existirá de información es la varianza de la variable que se está eliminando). O sea que, observando la matriz S a priori se podría deducir que tanta pérdida de información habría por reducir su dimensionalidad. Valores altos de correlaciones o valores bajos de varianzas indicarían casos donde se podría aplicar ACP con muy poca pérdida de información.

Una visión alternativa para el uso del ACP sería como método de compresión de datos para ser almacenados. Lo cual puede tener mucha utilidad para casos que, por la evolución de la tecnología, son cada vez más comunes y donde la cantidad de datos es enorme. Se continua con el supuesto de dos series iniciales ($\mathbb{R}^2 \rightarrow \mathbb{R}$). Inicialmente se tendrán dos series de N datos y el ACP dejará una sola serie de N coordenadas. Si a esa serie de coordenadas la se la multiplica por el vector base da los N puntos que nuevamente serán bidimensionales. Pero ahora todos sobre la recta que los ajusta, porque van a ser las estimaciones de los valores iniciales. Visto para grandes cantidades de datos esto permitiría almacenarlos solo con las series de coordenadas que tienen una dimensionalidad menor, en este caso la mitad, y la matriz de coordenadas haría las veces de decodificador para volver a obtener la información inicial.

Otra conclusión teórica obtenida, y no poco importante, es que la componente principal obtenida (la "nueva variable") va a ser una combinación lineal de las variables iniciales. El valor que va a tener la coordenada, que puede ser vista como una nueva variable Z va a ser los datos de cada par (X,Y) combinados linealmente con los coeficientes que se definan en cada caso. Esto recordemos sale del resultado de la coordenada para proyecciones ortogonales ($B_{ij} = x_j^T \cdot b_i$).

Respecto a la investigación del ACP como tema teórico, si bien representó un desafío mayor al planteado inicialmente, debido a la falta de bibliografía en español y de un texto que desarrollara el tema integralmente y en profundidad. Tuvo como resultado que el trabajo incorporara como objetivo compilar el tema teórico, para luego poder interpretar con más claridad los resultados obtenidos al usar esta metodología.

Por último, con respecto al objetivo práctico inicial del trabajo, que era construir un índice de actividad para la provincia de Mendoza a partir de esta metodología que se puede considerar como innovadora, o al menos como de uso poco frecuente en el ámbito de la Economía. Se puede decir que el objetivo fue cumplido. El índice obtenido puede ser sujeto a un análisis en mayor profundidad para determinar el por qué de su comportamiento. Al analizar los resultados obtenidos en el caso de estudio, se observa que, a través del método de Análisis por Componentes Principales, se logró crear un índice que representa la actividad económica de la provincia. Al comparar gráficamente este índice respecto al Estimador Mensual de Actividad Económica (EMAE) para Argentina, pareciera haber una equivalencia con algunos periodos de adelantados, lo cual indicaría a priori que las variaciones que se producen a Nivel Nacional se hacen evidentes, antes en la región. El peso de los bienes y servicios no transables de la actividad económica

mendocina, tanto intensivos en capital como en mano de obra, hacen que reaccione anticipadamente frente a las crisis o recuperaciones.

Esta investigación no se agota en el presente trabajo, puesto que queda la puerta abierta a mayores contrastaciones empíricas, que permitan evaluar la robustez del índice estimado a través del Análisis Factorial por Componentes Principales.

BIBLIOGRAFÍA CONSULTADA

CALDERÓN, M., GUTIERREZ, R. & TOLEDO, M. (2016) *Nivel socioeconómico por georeferenciación*. Editado por Editorial Académica Española. Saarbrücken.

D´JORGE, M. L., COHAN, P. P., HENDERSON, S. J., & SAGUA, C. E. (2007). Proceso de construcción del Índice Compuesto Coincidente Mensual de Actividad Económica de la Provincia de Santa Fe (ICASFe). Asociación Argentina de Economía Política (AAEP). C.A.B.A.

GÉRON, Aurélien (2017). Hands-on machine learning with Science-kit Learn & TensorFlow. *O'Reilly*.

GROPPER, D. (s.f.). Propuesta Metodológica para el Cálculo del Producto Bruto Geográfico en Argentina. Buenos Aires.

HASTIE, T., Tibshirani R. & Friedman, J. (2008) . The elements of statistical learning.

HEREDIA&ALVAREZ (2017) Indicador Compuesto de Actividad Económica para la provincia de Misiones Asociación Argentina de Economía Política (AAEP). C.A.B.A.

IMPERIAL COLLEGE OF LONDON (2019) Curso On Line *Mathematics for machine learning: PCA* Recuperado de: <https://es.coursera.org/learn/pca-machine-learning>

INDEC. (2016). *Metodología del Estimador Mensual de la Actividad Económica, EMAE*. Ciudad Autónoma de Buenos Aires: Instituto Nacional de Estadística y Censos.

JORRAT, J. M. (2003). *Indicador Económico Regional: El Índice Mensual de Actividad Económica de Tucumán (IMAT)*. Asociación Argentina de Economía Política (AAEP). C.A.B.A.

LAPELLE, H. C. (2013). *El Indicador Sintético de Actividad para la Región Rosario (ISARR) 1993-2012*. Asociación Argentina de Economía Política (AAEP). C.A.B.A.

LÓPEZ, C. (2004). *Técnicas de Análisis Multivariante de Datos: Aplicaciones con SPSS*. Recuperado de: https://www.researchgate.net/publication/40944326_Tecnicas_de_Analisis_Multivariante_de_Datos_Aplicaciones_con_SPSS

MATHEMATICS for machine learning: PCA (2016). Curso. *Imperial College of London*.

MUÑOZ, F., & TROMBETTA, M. (2014). *Indicador Sintético de Actividad Provincial (ISAP): Un Aporte al Análisis de las Economías Regionales*. Varias. Asociación Argentina de Economía Política (AAEP). C.A.B.A.

PÉREZ, Cesar (2004). *Técnicas de análisis multivariante de datos: Aplicaciones con SPSS*.

RAICHMAN, Silvia & Totter, Eduardo (2016). *Geometría analítica para ciencias e ingenierías*.

THE CONFERENCE BOARD (2001). *Business Cycle Indicators Handbook*. New York, U.S.A.