UNCUYO
UNIVERSIDAD NACIONAL DE CUYO

FCE
FACULTAD DE CIENCIAS ECONÓMICAS

Carrera: Licenciatura en Economía

ÍNDICE DE ACTIVIDAD ECONÓMICA DE MENDOZA, POR ANÁLISIS DE COMPONENTES PRINCIPALES

Trabajo de Investigación

POR

Fernando Deluret
Número de Registro: 26619
E-mail: fernandodeluret@gmail.com

Profesor Tutor

Mónica Calderón

Mendoza - año 2019

RESUMEN

El presente trabajo tiene por objetivo la construcción de un Índice de Actividad Económica (IAEMEN) para la provincia de Mendoza, el cual será de periodicidad mensual. En la actualidad no se ha publicado para la provincia ningún indicador de actividad económica de periodicidad menor a la anual. Esta investigación intenta presentar un aporte útil para satisfacer la necesidad de información actualizada de la situación económica regional. La metodología usada para la construcción del índice es el Análisis de Componentes Principales a partir de una serie de variables que se consideran representativas de la actividad económica de Mendoza, las cuales se encuentran disponibles en forma oportuna y tienen una periodicidad mensual en su publicación. 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. Adicionalmente al objetivo inicial que era la construcción del índice se agregó, durante el desarrollo del trabajo, la compilación de la teoría del Análisis de Componentes Principales. Esto debido a que no existía una bibliografía en español que cubriera los diferentes aspectos del tema integrados. Como resultados del trabajo se logró construir un índice que captara los patrones de comportamiento particulares de la economía provincial, y se detectó un aparente rezago entre los movimientos de la economía de Mendoza respecto a la nacional.

Palabras clave: Índice de Actividad Económica de Mendoza, Análisis de Componentes Principales, Análisis Factorial, EMAE, Producto Bruto Geográfico.

ÍNDICE

Introducción 4

CAPÍTULO I – ÍNDICE DE ACTIVIDAD ECONÓMICA 5

1. Concepto 5
2. Estado actual e importancia 5
3. Metodología 6

CAPÍTULO II – CONCEPTOS PREVIOS 8

4. Vectores y espacios vectoriales 8
5. Ortogonalidad 9
6. Proyecciones ortogonales 9
7. Complemento ortogonal 13

CAPÍTULO III – ANÁLISIS DE COMPONENTES PRINCIPALES 14

CAPÍTULO IV – ANÁLISIS FACTORIAL 23

CAPÍTULO V – ROTACIÓN DE COMPONENTES 28

CAPÍTULO VI – CONSTRUCCIÓN DEL ÍNDICE 30

Conclusiones 36

Bibliografía consultada 37

INTRODUCCIÓN

En el presente trabajo se avanza en la construcción de un Índice de Actividad Económica (IAEMEN) para la provincia de Mendoza, el cual será de periodicidad mensual. En la actualidad no se ha publicado para la provincia ningún indicador de actividad económica de periodicidad menor a la anual; el indicador existente es el Producto Bruto Geográfico provincial. Esta investigación 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 método de Análisis de Componentes Principales a partir de una serie de variables que se consideran representativas de la actividad económica de Mendoza, las cuales se encuentran disponibles en forma oportuna y tienen una periodicidad mensual en su publicación. La metodología usada para comprobar la validez del índice creado será contrastarlo con respecto al EMAE, y ver si el comportamiento entre ambos se puede transpolar al comportamiento entre PBG y PBI.

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, y que por ende pueda ser representado como un conjunto menor de variables, llamadas factores o componentes.

La investigación se estructura de la siguiente manera, en el capítulo I se hace 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 el capítulo II se analizan una serie de conceptos previos necesarios tanto para el desarrollo matemático del ACP, como para su interpretación práctica. En el capítulo III se realiza el desarrollo algebraico y geométrico, a partir del cual surge la fórmula resultante del ACP. En el capítulo IV se analiza el tema desde la perspectiva del análisis factorial, propiamente dicho; lo cual permite añadir el componente fundamental de interpretabilidad de los componentes. En el capítulo V se analiza 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. En el capítulo VI se trabaja con el caso numérico, el cual tiene por objetivo construir un índice de actividad mensual para la provincia de Mendoza. Por último, se analizan las consecuencias prácticas del ACP y se sacan algunas conclusiones.

CAPÍTULO I

ÍNDICE DE ACTIVIDAD ECONÓMICA

En el presente capítulo se desarrollará el concepto de índice económico, se analizará que antecedentes hay sobre el tema en el país y se presentará la metodología con la que se construirá el índice. Esto nos dejará de manifiesto la importancia tanto teórica como práctica que se espera tenga la investigación.

1. Concepto

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.

2. Estado actual e importancia

Actualmente para la provincia de Mendoza el indicador de actividad económica usado es el Producto Bruto Geográfico (PBG), el cual tiene una periodicidad anual. Adicionalmente, existe una serie de variables que a priori se podrían imaginar como representativas de la actividad económica, como por ejemplo el consumo de electricidad, el consumo de gas, el patentamiento de vehículos, etcétera. Las cuales, si tienen una periodicidad mensual, por lo cual el objetivo es construir a partir de las mismas un indicador de actividad económica que sea también mensual. A nivel nacional el índice equivalente 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 nacional.

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.

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 (por ejemplo, consumos de energía, dan lugar a una nueva ecuación que se estima con una nueva variable explicativa, por ejemplo, consumo industrial de gas. Los residuos de esta nueva ecuación dan lugar a una nueva estimación con variables relacionadas y así sucesivamente hasta agotar la significatividad de los residuos. Otra forma de estimar el indicador compuesto es desarrollar una función de producción con sus componentes significativos referidos a mano de obra, materias primas, consumos intermedios y tal vez algún otro integrante significativo, con el objeto de estimar un valor bruto de producción que represente en forma acabada el indicador que represente la variable solicitada.

3. Metodología

La metodología que se utiliza para la construcción del índice será 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 de publicación al menos anual, buscándose como ideal que sean mensuales. Las variables que finalmente se usen deberán ser representativas de la actividad económica y tener cierto nivel de correlación con el índice construido, sobre estos aspectos se profundizará en el capítulo VI, donde se realizará la construcción del índice propiamente.

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 contra el EMAE y ver si la relación de comportamiento entre ambos se mantiene al traspolarlo a la comparación entre PBG provincial y PBI.

La situación del Análisis de componentes principales como herramienta teórica merece especial atención ya que influirá notoriamente en el desarrollo del presente trabajo. Esto se debe a que actualmente no existe una fuente en español que realice una explicación integral del tema. Sino que está tratado por partes, de acuerdo a la aplicación a la que esté dirigido. Por esto un objetivo muy importante en el presente trabajo será también presentar una integración teórica del tema, que tenga en cuenta los puntos de vista geométrico, matemático y conceptual. Así como también presentar todo esto siguiendo una sola línea lógica y manteniendo una notación unificada a lo largo del desarrollo que permita entender en mayor profundidad el tema, para a posteriori sacar conclusiones más interesantes a la hora de su aplicación.

En el capítulo se trató el concepto de índice económico los antecedentes en el tema a en el país y se presentó la metodología de ACP con la que se construirá el índice.

CAPÍTULO II

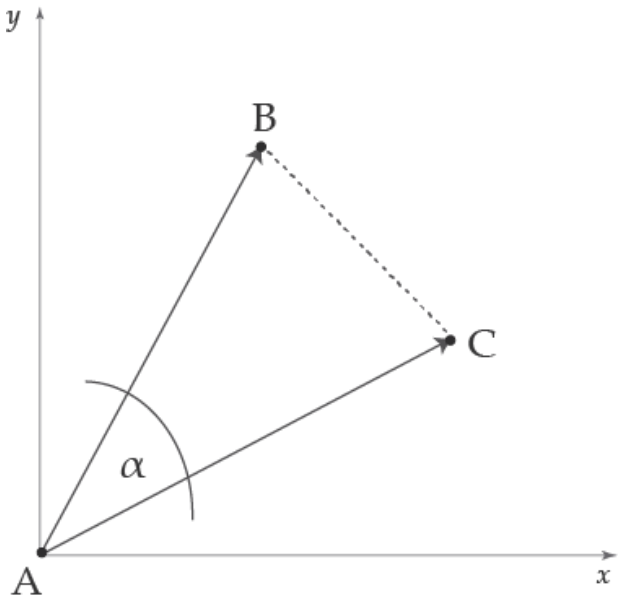
CONCEPTOS PREVIOS

En el presente capítulo se desarrollarán algunos conceptos teóricos claves para entender el desarrollo del ACP. Como la definición de vector y espacio vectorial, sus propiedades y operaciones, las bases de un espacio vectorial. También se desarrollará el concepto de proyección ortogonal que es el concepto central sobre el que se basa el ACP.

1. Vectores y espacios vectoriales

En el presente apartado se trabajará en $\mathbb{R}^2$ y dados 2 vectores $\overline{\mathrm{AB}}$ y $\overline{\mathrm{AC}}$ como se muestra en el Gráfico 1, podemos definir el módulo de un vector, la distancia entre vectores y el ángulo entre vectores de la siguiente manera:

Gráfico 1: Distancia entre vectores



Siendo $w_1 = \overline{AB}$ y $w_2 = \overline{AC}$, con $w_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$

Podemos definir el producto punto entre 2 vectores como:

$$w_1{}^T . w_2 = x_1 . x_2 + y_1 . y_2$$

El *módulo* de w_1 será $||w_1|| = \sqrt{w^T . w}$ y esto es igual al segmento $\overline{AB}$.

La *distancia* entre w_1 y w_2 será el modulo del vector diferencia entre ambos $||w_1 - w_2||$ lo cual va a ser igual al segmento $\overline{BC}$.

El *ángulo* α entre estos 2 vectores va a ser tal que: $\cos(\alpha) = \frac{w_1{}^T . w_2}{||w_1||.||w_2||}$

Un espacio vectorial es un conjunto de vectores que cumplen las siguientes propiedades:

- Son cerrados respecto de una operación (por ejemplo la suma): si $w_1, w_2 \in V \Rightarrow w_1 + w_2 \in V$
- Asociatividad: $w_1 + (w_2 + w_3) = (w_1 + w_2) + w_3$

- Elemento neutral (e): $\exists e \in V$ tal que $w_1 + e = e + w_1 = w_1$
- Elemento inverso: $\exists {w_1}^{-1} \in V$ tal que ${w_1} + {w_1}^{-1} = e$
- Multiplicación por un escalar: $w_1 \in V \Rightarrow \lambda . w_1 \in V$

Un subespacio vectorial es entonces un subconjunto de ese espacio tal que si se aplica cualquiera de las operaciones arriba mencionadas entre vectores del subespacio el vector resultante seguirá perteneciendo al mismo. Una consecuencia de esto es que el subespacio debe tener si o si el elemento neutro, por ejemplo, las rectas que pueden ser subespacio de $\mathbb{R}^2$ deben pasar si o si por el origen de coordenadas.

Dado un espacio vectorial V, un *conjunto generador* para este espacio es un conjunto de vectores que combinados linealmente permiten construir cualquier vector de dicho espacio. Si adicionalmente el conjunto generador cumple las siguientes propiedades se dice que este conjunto es una *base* de ese espacio vectorial y a cada uno de los vectores de la base se le llama *vector base:*

- Mínimo: $A = [w_1, w_2, \dots, w_k] \quad A \subseteq V$, será mínimo si no existe otro conjunto generador con menos de k vectores.
- Los vectores de A deben ser linealmente independientes

Una propiedad de una base es que 2 combinaciones lineales diferentes de sus vectores base darán como resultado 2 vectores diferentes entre sí siempre.

Una *base estándar* de un espacio $\mathbb{R}^n$, es una base tal que la matriz formada por sus vectores base es la matriz identidad de dimensión $n \times n$. Por ejemplo, la base estándar para $\mathbb{R}^3$ es:

$$A = \left\{ \underbrace{\begin{bmatrix}1\\0\\0\end{bmatrix}}_{\substack{vector\\base}}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}$$

Puede haber otras bases que no sean la estándar para un espacio vectorial, pero todas van a tener la misma cantidad de vectores.

2. Ortogonalidad

Se dice que dos vectores son ortogonales entre sí, si el ángulo que forman es de 90°. Por lo que, dada la fórmula del ángulo entre vectores:

$$\cos(90^o) = \frac{{w_1}^T . w_2}{||w_1||.||w_2||} = 0$$

Para que esto se cumpla el producto punto entre ambos vectores debe ser igual a cero. Una forma de interpretar la ortogonalidad es que dos vectores ortogonales son “lo más diferentes posible” entre sí.

Si los vectores de una base son ortogonales entre si ${w_i}^T . w_k = 0 \quad i \neq k$ y todos tienen módulo igual a 1 $||w_i|| = 1$ entonces la base es *ortonormal*.

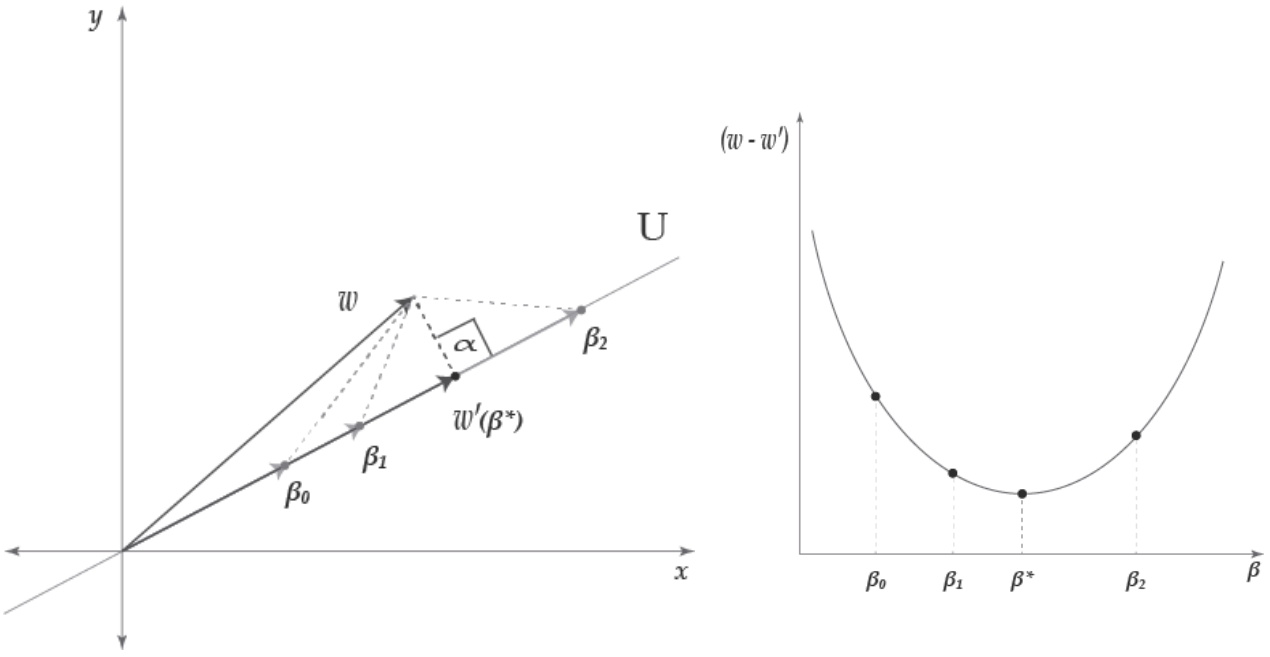
3. Proyecciones ortogonales

Dado el subespacio vectorial $U \subset \mathbb{R}^2$ y el vector $w \in \mathbb{R}^2$, el cual puede ser representado como una combinación lineal de los vectores base de $\mathbb{R}^2$. Supongamos que queremos representar w en el subespacio U, vamos a querer que el nuevo vector $w' \in U$ sea lo más parecido al vector original. Lo cual es equivalente a minimizar el segmento que representa el vector diferencia $(w - w')$, el cual en el Gráfico 2 está representado por las distintas líneas punteadas para cada caso.

Siendo $b_1 = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} \in U$ una base del subespacio U, el nuevo vector w' va a poder ser expresado como una combinación lineal de dicha base $w' = \beta . b_1$, donde β es un escalar y representa las coordenadas de w' en U.

La pendiente de U nos la va a dar la relación entre las componentes x_1 e y_1 del vector base b_1.

Gráfico 2: Proyección ortogonal de un vector



Como se puede observar en el grafico anterior de acuerdo al β que se elija va a ser el largo del vector w' resultante. Y por ende, el largo del vector diferencia $(w - w')$ va a depender también de β. Existirá un β^* que minimice esa distancia, el cual va a ser tal que haga que el vector diferencia y el vector base b_1 sean ortogonales ($\alpha = 90^o$). Esto como se vio en el apartado anterior es equivalente a decir que el producto punto entre ambos vectores es igual a cero, matemáticamente se puede expresar:

$$b_1{}^T . (w - w') = 0$$

$$\Rightarrow (b_1{}^T . w) - (b_1{}^T . w') = 0$$ Por propiedad de bilinearidad del producto punto

$$\Rightarrow ({b_1}^T.w) - ({b_1}^T.\beta.b_1) = 0$$ Usando la definición de w'

$$\Rightarrow ({b_1}^T.w) - \beta(\underbrace{{b_1}^T.b_1}_{||b_1||^2}) = 0$$

$$\Rightarrow \beta = \frac{{b_1}^T.w}{||b_1||^2}$$

$$\Rightarrow w' = \beta.b_1 = \underbrace{\frac{b_1.{b_1}^T}{||b_1||^2}}_{\substack{Matriz \\ de \\ proyección}} w$$ Dado que $({b_1}^T.w)$ es un escalar, podemos moverlo al final de la multiplicación

La matriz de proyección es aquella matriz (en este caso 2×2) que multiplicada por el vector original w nos da la proyección de este en el subespacio U (w'). Si b_1 es ortonormal ($||b_1||^2 = 1$) la matriz de proyección va a ser simétrica:

$$b_1.{b_1}^T = \left(b_1.{b_1}^T\right)^T$$

Se puede ver entonces, que antes para representar al vector w necesitábamos dos coordenadas (x_1 e y_1) y ahora necesitamos una sola (β).

Algo importante a destacar en este tema es que el nuevo vector w' va a seguir perteneciendo a $\mathbb{R}^2$, a pesar de que ahora solo dependa de un parámetro (β). Esto es debido a que el vector base de U (b_1) tiene dos componentes y por ende w' también. Por este motivo no sería lo mismo por más que se haga una proyección sobre una recta partir desde $\mathbb{R}^2$, que por ejemplo desde $\mathbb{R}^3$, en cuyo caso b_1 tendría tres componentes. Una forma de ver esto es que b_1 está dando información de acuerdo a los grados de libertad de rotación del espacio vectorial inicial, por ejemplo en $\mathbb{R}^2$ la recta U solo puede rotar a lo largo del plano XY en tanto que en $\mathbb{R}^3$ la recta podría rotar respecto al plano respecto a XY, YZ y XZ.

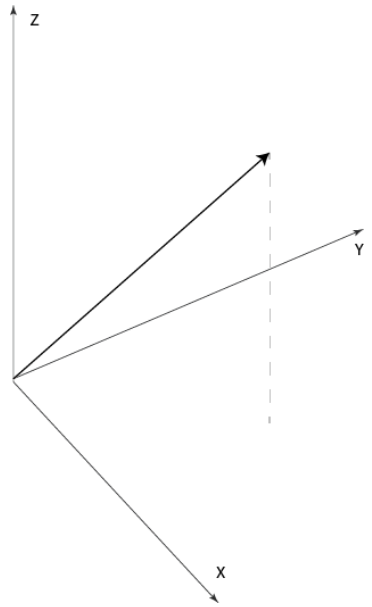
Se supone ahora que $w \in \mathbb{R}^D$ es decir $w = [w_1, w_2, \dots, w_D]$ y $U \subset \mathbb{R}^D$, pero ahora en vez de ser una recta (dimensión igual a 1) es un subespacio M-dimensional, con $M < D$. Es decir, tiene M vectores base $b = [b_1, b_2, \dots, b_M]$ donde a su vez cada b_i tiene dimensión $D \times 1$ ya que U es un subespacio de $\mathbb{R}^D$.

Para aclarar esto último, la cantidad de vectores base b_i que tenga U indica si el o los vectores que se proyecten lo harán sobre una recta, un plano etcétera. En tanto que "D" lo que indica es que esa recta o plano debe dibujarse en el espacio D-dimensional. Por ejemplo $b = \begin{bmatrix}0\\0\\1\end{bmatrix}$ genera una recta en $\mathbb{R}^3$ en tanto que el conjunto $b = \left\{\begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ genera un plano en $\mathbb{R}^3$, luego sobre ese plano o recta es sobre el que se proyectara el vector original.

En este caso multidimensional $w' = \sum_{i=1}^{M} \beta_i.b_i$ notado matricialmente $w' = B.\beta$, donde B es la matriz que representa el conjunto de los vectores base de U (con dimensión $D \times M$) y β es el vector $M \times 1$ con las M coordenadas.

La condición impuesta en el caso inicial para que la proyección fuera ortogonal es que el vector diferencia fuera ortogonal con el vector base del nuevo subespacio. Por ende lo que se va a exigir ahora es que $(w - w')$ sea ortogonal con cada uno de los M vectores base. Para verlo con un ejemplo, supóngase el caso en que se quiere proyectar un vector de $\mathbb{R}^3$ en el plano XY (es decir componente z=0) el cual se va a suponer horizontal como se ilustra en el gráfico 3, en tanto que la componente Z daría la altura. La forma de que el vector diferencia sea lo más pequeño posible seria bajando verticalmente, es decir en forma perpendicular al plano XY, con lo cual el vector diferencia va a ser ortogonal a todos los vectores de XY incluidos ambos vectores base.

Gráfico 3: Ortogonalidad de los vectores base



$$b_i{}^T.(w - w') = 0 \quad con\ \ i = 1, 2, \dots, M$$

Es decir, ahora hay un sistema de M ecuaciones simultáneas. Notado matricialmente:

$$\underbrace{B^T}_{MxD}\left(\underbrace{w}_{Dx1} - \underbrace{\underbrace{B}_{DxM}\underbrace{\beta}_{Mx1}}_{Dx1}\right) = 0$$

$$\Rightarrow B^T.w - B^T.B.\beta = 0$$

$$\Rightarrow \beta = (B^T.B)^{-1}B^T.w$$

$$\Rightarrow w' = B.\beta = \underbrace{B(B^T.B)^{-1}B^T}_{\substack{Matriz\ de\ proyeccion \\ (DxD)}}.w$$

Donde ahora w' y β son las fórmulas para calcular la proyección pero para el caso general de D dimensiones y llevar al vector a un subespacio con solo M coordenadas ($M < D$). Una vez más w' seguirá siendo un vector $D \times 1$.

4. Complemento ortogonal

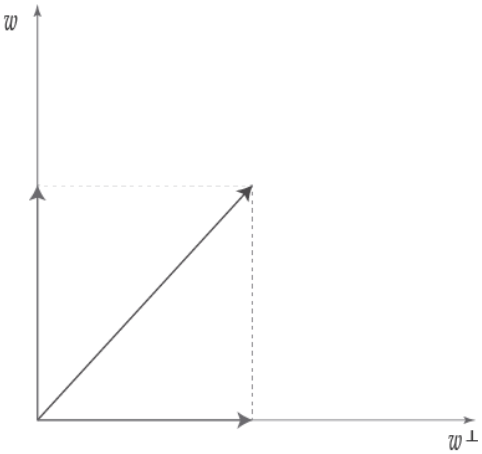
Dado un espacio vectorial V^n y un subespacio $W \subseteq V, \text{con } W^k$ (siendo n y k las dimensiones de los espacios). Entonces el complemento ortogonal de W es un subespacio $W^{\perp}$ de dimensión (n-k), tal que $W^{\perp}$ contiene todos los vectores de V que sean ortogonales a todos los vectores de W.

Cada vector $x \in V$ se puede *descomponer ortogonalmente* (de forma única) de la siguiente forma:

$$x = \sum_{i=1}^{k} \delta_i b_i + \sum_{i=1}^{n-k} \varphi_i {b_i}^{\perp}$$

Donde b_i y ${b_i}^{\perp}$ son los vectores base de los subespacios W y $W^{\perp}$ respectivamente. Lo único que se está diciendo es que cualquier vector x se puede expresar como una suma de un vector de W más un vector de $W^{\perp}$. Por ejemplo, supongamos que W y $W^{\perp}$ son 2 rectas pertenecientes a $\mathbb{R}^2$ perpendiculares entre sí, cualquier vector de $\mathbb{R}^2$ puede ser expresado como una suma de vectores como en el Gráfico 4.

Gráfico 4: Complemento ortogonal de un vector



Hasta aquí se analizado los conceptos geométricos y matemáticos necesarios para entender el desarrollo del análisis de componentes principales. En el próximo capítulo se iniciará el desarrollo de dicho tema en el que se aplicarán estos conceptos.

CAPÍTULO III

ANÁLISIS DE COMPONENTES PRINCIPALES

En el presente capítulo se realizará el desarrollo del ACP analizando sus aspectos gráfico, matemático y geométrico. A partir de esto se expondrán algunas conclusiones teóricas que serán usadas en la construcción del índice y que tendrán una influencia práctica en el mismo.

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 $\mathbb{R}^D$:

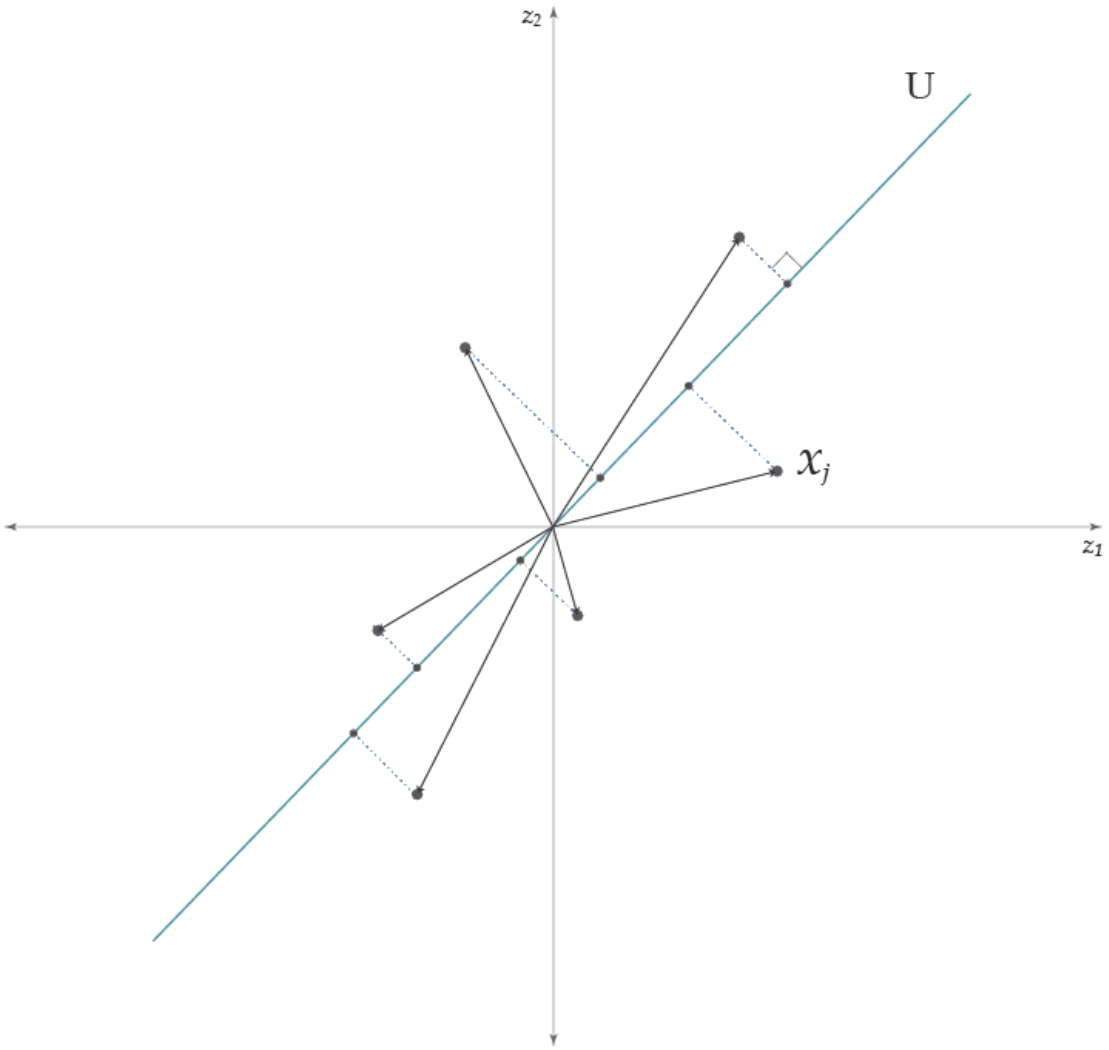
$$X = \{x_1, x_2, \dots, x_{N,}\} \quad x_j \in \mathbb{R}^D$$

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 5 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 5. 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 U[1].

[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 5: Caso de dos variables iniciales y un factor



De acuerdo a lo visto en el capítulo anterior, cada uno de los vectores x_j puede ser representado como una combinación lineal de los vectores base de $\mathbb{R}^D$:

$$x_j = \sum_{i=1}^{D} B_{ij} \, . \, b_i \qquad (1)$$

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 Dx1, 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 $\sum_{i=1}^{D} {b_i}^2 = 1$.

$$B_{ij} = {x_j}^T \, . \, b_i \qquad (2)$$

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. $B_{ij} \, . \, b_i$ 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 han hecho dos supuestos importantes, uno es que los datos deben tener media cero y el otro es que las bases van a ser ortonormales. A continuación, se profundizará un poco en este último supuesto.

El hecho que las bases sean ortonormales significa dos cosas, una ya vista es que la norma de cada vector base es 1 ($\sum_{i=1}^{D} {b_i}^2 = 1$), la otra es que los vectores base van a ser ortogonales entre sí, es decir, su producto punto va a ser igual a cero (${b_i}^T . b_j = 0 \text{ para } todo \; i \neq j$). Gráficamente esto último significa que los vectores base, que pueden ser vistos como los nuevos ejes de referencia para los datos transformados (ya que es respecto de quienes están dadas las coordenadas) serán perpendiculares entre sí.

Con esto en mente se retoma la fórmula obtenida para proyecciones ortogonales M-dimensionales:

$$\tilde{x}_j = B.\beta = B(B^T.B)^{-1}B^T.x_j$$

Donde $\tilde{x}_j$ es la proyección ortogonal del vector x_j en el subespacio M-dimensional y B es la matriz formada por el conjunto de los M vectores base de dicho subespacio $B = [b_1, b_2, \dots, b_M]$ donde cada b_i tiene dimensión $D \times 1$. Analizando la matriz resultante de $B^T.B$ tendremos que:

$$B^T.B = \begin{bmatrix} {b_1}^T.b_1 & \dots & {b_1}^T.b_M \\ \vdots & \ddots & \vdots \\ {b_M}^T.b_1 & \dots & {b_M}^T.b_M \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}$$

Dado que por ser bases ortonormales ${b_i}^T.b_j = 0 \text{ para } todo \; i \neq j \; y \; {b_i}^T.b_i = 1$, entonces el resultado es la matriz identidad $M \times M$. Por lo que se puede reescribir $\tilde{x}_j$ de la siguiente manera:

$$\tilde{x}_j = B.\beta = B.B^T.x_j$$

Se puede expresar el problema de reducir la dimensionalidad de la siguiente manera, de acuerdo a lo visto en complemento ortogonal de un subespacio vectorial:

$$\tilde{x}_j = \underbrace{\sum_{i=1}^{M} \beta_{ij}.b_i}_{\substack{Subespacio \\ principal}} + \underbrace{\sum_{i=M+1}^{D} \beta_{ij}.b_i}_{=0} \qquad (3)$$

Donde el primer término es el subespacio sobre el que vamos a proyectar los datos y el segundo término (que representa un subespacio $\mathbb{R}^{D-M}$ dimensional que es el complemento ortogonal del

subespacio principal) para nuestra proyección es cero ya que haremos la proyección en el subespacio principal.

Se puede definir una función J que represente el error de hacer la proyección ortogonal de los datos, que como ya se dijo, puede verse como la sumatoria de los vectores diferencia (representados en el gráfico 5 por las líneas punteadas).

$$J = \frac{1}{N}\sum_{j=1}^{N} ||x_j - \tilde{x}_j||^2$$

$$donde\ ||x_j - \tilde{x}_j||^2 = \left(x_j - \tilde{x}_j\right)^T . (x_j - \tilde{x}_j)$$

Lo que se busca aquí es encontrar un subespacio tal que minimice esta función de error. Ahora bien, $\tilde{x}_j$ va a depender de los valores que tomen β_{ij} y b_i entonces:

(1°) $$\frac{\partial J}{\partial \beta_{ij}} = \frac{\partial J}{\partial \tilde{x}_j} . \frac{\partial \tilde{x}_j}{\partial \beta_{ij}} = \underbrace{\frac{-2}{N}\left(x_j - \tilde{x}_j\right)^T}_{\frac{\partial J}{\partial \tilde{x}_j}} . \underbrace{b_i}_{\frac{\partial \tilde{x}_j}{\partial \beta_{ij}}} = 0$$

(2°) $$\Rightarrow \frac{\partial J}{\partial \beta_{ij}} = \frac{-2}{N}\left(x_j - \sum_{h=1}^{M} \beta_{hj} . b_h\right)^T . b_i = 0$$

(3°) $$\Rightarrow (x_j{}^T . b_i) - \left(\sum_{h=1}^{M} \beta_{hj} . b_h\right)^T . b_i = 0$$

(4°) $$\Rightarrow (x_j{}^T . b_i) - \beta_{ij} . \underbrace{b_i{}^T . b_i}_{=1} = 0$$

(5°) $$\Rightarrow \beta_{ij} = x_j{}^T . b_i$$

Donde el segundo término sale de derivar (3) respecto a β_{ij} suponiendo M=1, es decir para la proyección unidimensional sobre el subespacio generado por el vector b_i. En el segundo paso lo que se hace es usar (3) donde $\tilde{x}_j$ es la proyección de x_j en el subespacio principal, por eso la sumatoria solo llega hasta M. En la cuarta línea se simplificó la sumatoria ya que el producto punto resultante va a ser cero para todos los términos donde $i \neq h$ por la propiedad vista de bases ortonormales, y también por esto es que se puede decir que $b_i{}^T . b_i$=1.

La expresión a la que se llega coincide con la ecuación (2) que era la obtenida para proyecciones ortogonales. Entonces, lo que se está diciendo aquí es simplemente que las coordenadas β_{ij} que minimizan el error J son las que proyectan ortogonalmente el vector x_j en el eje generado por b_i.

Visto en términos del gráfico 5, se busca la pendiente de la recta U. Para simplificar la resolución matemática primero re se re expresará la función J para dejarla en función de b_i. Se parte de la definición de $\tilde{x}_j$ en (3) y se reemplaza β_{ij} de acuerdo a (2) podemos expresar:

$$\tilde{x}_j = \sum_{i=1}^{M} \beta_{ij} . b_i$$

$$\tilde{x}_j = \sum_{i=1}^{M} ({x_j}^T . b_i) . b_i$$

A partir de esta última expresión obtenida, y expresando x_j de la ecuación (1) como la suma del subespacio principal más el complemento ortogonal:

$$x_j = \sum_{i=1}^{M} ({x_j}^T . b_i) . b_i \; + \sum_{i=M+1}^{D} ({x_j}^T . b_i) . b_i$$

Se puede entonces, re escribir el vector diferencia como:

$$x_j - \tilde{x}_j = \sum_{i=M+1}^{D} ({x_j}^T . b_i) . b_i$$

Se reemplaza esto en la función J usando el hecho de que el producto punto ${x_j}^T . b_i = {b_i}^T . x_j$ se puede expresar vectorialmente de cualquiera de esas dos formas:

(1°) $$J = \frac{1}{N} \sum_{j=1}^{N} \left\| \sum_{i=M+1}^{D} ({b_i}^T . x_j) . b_i \right\|^2$$

(2°) $$\Rightarrow J = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=M+1}^{D} \| {b_i}^T . x_j \|^2 . \underbrace{\|b_i\|^2}_{=1}$$

(3°) $$\Rightarrow J = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=M+1}^{D} ({b_i}^T . x_j) . ({x_j}^T . b_i)$$

(4°) $$\Rightarrow J = \sum_{i=M+1}^{D} {b_i}^T . \underbrace{\left(\frac{1}{N} \sum_{j=1}^{N} x_j . {x_j}^T \right)}_{\hat{S}} . b_i \qquad (4)$$

A continuación, se explica brevemente el proceso matemático precedente. En la segunda línea se saca la sumatoria fuera de la norma. Ya que va a ser lo mismo sacar la suma de los vectores y sacarle la norma que sacar la norma de todos los vectores y sumar dichas normas. También se hizo uso de que la norma de cada vector base es igual a 1 por bases ortonormales. En la tercera línea se usó la definición de norma para expresarla como producto punto. En la cuarta línea se saca “afuera” la sumatoria que depende de *i* y quedan dentro de la sumatoria con subíndice j solo los términos que dependen de j.

Si se observa la matriz S, se puede ver que va a ser la suma de N matrices $D \times D$. Por lo que cada elemento de la matriz resultante va a ser la suma de ese elemento de cada una de las N matrices de la sumatoria. Adicionalmente se debe multiplicar cada uno de esos elementos de la matriz por $\frac{1}{N}$. Se puede ver entonces, que ya que los datos tienen media igual a cero, esta matriz es la matriz de varianzas y covarianzas ($D \times D$) de las D variables iniciales. Si los datos estuvieran estandarizados (adicionalmente estuvieran divididos por su desviación estándar) la matriz obtenida aquí, sería la matriz de correlaciones.

Para el caso en el que hay dos variables iniciales y se quiere reducir los datos a una sola componente principal ($\mathbb{R}^2 \to \mathbb{R}$). Si se plantea un *lagrangiano* para obtener la condición de minimización a partir de la fórmula obtenida en (4), y se tiene en cuenta la restricción de bases ortonormales quedará entonces:

(1°) $$J = {b_2}^T . S . b_2 \qquad s.a. \quad {b_2}^T . b_2 = 1$$

(2°) $$L = {b_2}^T . S . b_2 + \lambda(1 - {b_2}^T . b_2)$$

(3°) $$\frac{\partial L}{\partial \lambda} = 1 - {b_2}^T . b_2 = 0$$

(4°) $$\frac{\partial L}{\partial b_2} = 2 . {b_2}^T . S - 2 . \lambda . {b_2}^T = 0$$

(5°) $$\Rightarrow {b_2}^T . S = \lambda . {b_2}^T$$

(6°) $$\Rightarrow S . b_2 = \lambda . b_2 \qquad (5)$$

(7°) $$J = {b_2}^T . b_2 \, . \lambda = \lambda$$

Donde en la cuarta línea se usó la regla de derivada multivariante para derivar respecto a b_2. En la sexta línea por ser S una matriz simétrica puede conmutar de esa forma la multiplicación. En tanto que en la última línea solo reemplaza el resultado anterior en la función J.

Si se observa la expresión (5), se puede ver que la forma funcional es la misma que la de un problema de *eigen valores* y *eigen vectores*[2]. Para recordar, dada una matriz A ($n \times n$), v será eigenvector de A si se cumple que:

$$A . v = \lambda . v$$

Donde λ es un escalar y es el eigenvalor de la matriz A correspondiente a ese eigenvector. Para minimizar entonces J lo que se debe hacer es buscar los *eigen* vectores de la matriz S y elegir como subespacio principal el *eigen* vector que tenga el λ. De esta manera su complemento ortogonal que será el omitido tendrá el $J = \lambda$ de menor valor.

[2] Al hablar de *eigen valores* y *eigen vectores,* se habla de valores propios y vectores propios respectivamente. Esta forma de llamarlos será la usada en el resto del trabajo.

Notar que los *eigen* vectores serán los vectores base, por lo cual por bases ortonormales serán ortogonales entre sí, y el *eigen* vector con el *eigen* valor más alto es 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).

Si se amplía esta lógica para el caso D-dimensional se tendrá que:

$$S.b_j = \lambda_j.b_j \;\; con\, j = M+1, M+2, \dots, D$$

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

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 $x_j{}^T.b_i = \beta_{ij}$. Notando esto matricialmente para el caso de N vectores (recordar que se tendrá N vectores porque hay N observaciones) se obtiene lo siguiente:

$$\underbrace{X}_{N\times D} \times \underbrace{B}_{D\times M} = \underbrace{Z}_{N\times M}$$

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

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

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

$$Z_1 = X.b_1$$

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.

Si se analizan los resultados obtenidos para proyecciones ortogonales M-dimensionales, se puede observar que $\tilde{x}_j = B.\beta$. La única aclaración metodológica que corresponde hacer es que aquí cada uno de los N vectores está expresado horizontalmente como una de las filas de la matriz $\tilde{X}$, es decir que cada fila podría ser expresada de acuerdo a la ecuación de proyecciones ortogonales de la siguiente manera $\underbrace{\tilde{x}_j{}^T}_{1xD} = \underbrace{\beta^T}_{1xM} . \underbrace{B^T}_{MxD}$. Se expresa esto matricialmente para el caso de N vectores:

$$\underbrace{Z}_{N\times M} \times \underbrace{B^T}_{M\times D} = \underbrace{\tilde{X}}_{N\times D}$$

Está expresión significa que se pueden "reconstruir" los datos originales usando la matriz B como decodificador, la pérdida de información será la que produzca al omitir el subespacio D-M dimensional complementario al subespacio principal.

A partir de la expresión $Z_1 = X.b_1$ y se calcula la esperanza y la varianza de la nueva variable:

$$E(Z_1) = E(X.b_1) = E(X).b_1 = 0$$

$$VAR(Z_1) = \frac{1}{N}.{Z_1}^T.Z_1 = \frac{1}{N}.{b_1}^T.X^T.X.b_1 = {b_1}^T.\underbrace{\left(\frac{1}{N}.X^T.X\right)}_{S}.b_1$$

Donde en la primera ecuación se usó el hecho de que los datos iniciales deben estar centrados respecto a su media.

Si se observa la expresión obtenida para la varianza de Z_1 se puede ver que es igual a la obtenida cuando se buscaba cada una de las D-M dimensiones omitidas, comparando ambos resultados se tiene que:

$${b_1}^T.S = \lambda.{b_1}^T \quad \Rightarrow \quad \underbrace{{b_1}^T.S.b_1}_{VAR(Z_1)} = \lambda.\underbrace{{b_1}^T.b_1}_{=1} \quad \Rightarrow \quad VAR(Z_1) = \lambda$$

Interpretando esta última expresión se puede decir que se busca retener la mayor cantidad de varianza posible, o que se van a elegir los M componentes principales con mayor varianza. Esto es porque Z_1 es la proyección ortogonal de los datos sobre el eje generado por el vector b_1, y este eje es elegido de tal forma de maximizar la varianza de los datos sobre él. Este enfoque permite compatibilizar la idea de que al retener la mayor cantidad de varianza posible se está manteniendo la mayor cantidad de información, y es porque al mantener la mayor cantidad de varianza se está minimizando el error de proyección de los datos en sobre ese subespacio o eje.

Una medida de la variabilidad original de los datos podría ser la suma de las D varianzas de las variables originales X. Por teoría de diagonalización[3] se puede decir que:

$$Traza(S) = Traza(P'.V.P) = \sum_{j=1}^{D} \lambda_j$$

Por otra parte, la traza de S (la suma de todos los elementos de su diagonal principal) va a ser la suma de las D varianzas iniciales, que es como se dijo la variabilidad total de los datos iniciales. Adicionalmente si las variables iniciales están estandarizadas esta varianza total será igual a D (cada una de las D variables iniciales tiene var=1). Se tiene entonces aquí una relación importante entre la varianza original y la varianza retenida después del ACP. Por ejemplo, si se retuvieran todas las componentes principales (M=D) no habría perdida de varianza, lo que puede ser visto como que no habría perdida alguna de información.

[3] Cuya demostración escapa al alcance de este trabajo.

Se puede decir que la proporción o porcentaje de variabilidad que explica cada componente va a ser:

$$\%\ de\ Var\ explicado\ por\ Z_j = \frac{\lambda_j}{Traza(S)}$$

Algo que hasta acá se ha dado por sentado es M, es decir la cantidad de componentes que se van a retener. Si bien hay distintos criterios de decisión uno de los más usados es retener todas las componentes cuya varianza sea mayor a 1 ($\lambda_i > 1$). Es decir, que sea mayor a la varianza de cualquiera de las variables iniciales (todas igual a 1). La interpretación conceptual de esto es que la nueva componente va a ser "mejor" que cualquiera de las variables iniciales por si sola ya que va a contener más información.

Aquí finaliza el desarrollo del modelo ACP estrictamente hablando. En el siguiente capítulo se analizará el tema desde la perspectiva del análisis factorial, que va a permitir incorporar como elemento de análisis, la interpretabilidad de los componentes o nuevas variables obtenidas.

CAPÍTULO IV

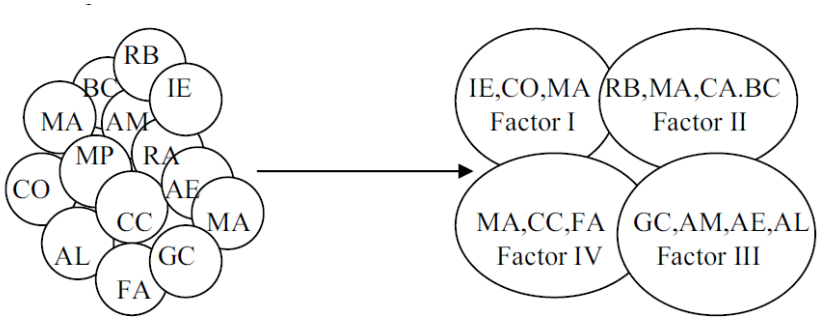
ANÁLISIS FACTORIAL

En el presente capítulo se estudiará en enfoque del análisis factorial, que se complementa con el ACP. Agrando al mismo la importancia de la interpretación de los factores obtenidos e intentando por lo tanto facilitar la misma.

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.

Para aclarar esto obsérvese el siguiente ejemplo, se quiere analizar la importancia que los consumidores dan a 14 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). Se observa que las 14 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) y no observables directamente. Por lo tanto, 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. En el gráfico 6 se puede ver este análisis.

Gráfico 6: Ejemplo de variables subyacentes



A continuación, se plantea matemáticamente el modelo factorial:

$$x_1 = l_{11}.F_1 + \cdots + l_{1M}.F_M + e_1$$
$$\vdots$$
$$x_D = l_{D1}.F_1 + \cdots + l_{DM}.F_M + e_D$$

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:

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

Expresando esto matricialmente se obtiene:

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

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

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

Donde ${h_1}^2$ es el porcentaje de varianza de la variable x_1 explicado por los factores comunes, y se llama ***comunalidad*** y ${\omega_1}^2$ es la parte de la varianza de x_1 que la explica su factor especifico 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 Nx1 que contiene todas las observaciones. Si se tiene esto en cuenta se puede re expresar la solución obtenida como:

$$\underbrace{X}_{N\times D} \times \underbrace{B}_{D\times M} = \underbrace{Z}_{N\times M}$$

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

$$\overbrace{Z_1}^{Nx1} = \overbrace{X_1}^{Nx1} . \overbrace{b_{11}}^{1x1} + \cdots + X_D . b_{1D}$$
$$\vdots$$
$$Z_M = X_1 . b_{M1} + \cdots + X_D . b_{MD}$$

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:

$$\tilde{X}_1 = Z_1 . b_{11} + \cdots + Z_M . b_{M1}$$
$$\vdots$$
$$\tilde{X}_D = Z_1 . b_{1D} + \cdots + Z_M . b_{MD}$$

Como se vio en el capítulo anterior la diferencia entre $\tilde{X}_i$ y X_i es la pérdida de información que existe por la reducción de dimensionalidad, si se supone entonces que se realiza el ACP sin reducir dimensionalidad (M=D), es decir conservando los D componentes principales. Este sistema de ecuaciones quedaría de la siguiente forma:

$$X_1 = Z_1 . b_{11} + \cdots + Z_D . b_{D1}$$
$$\vdots$$
$$X_D = Z_1 . b_{1D} + \cdots + Z_D . b_{DD}$$

Pero el análisis factorial exige que los factores estén estandarizados, entonces estandarizando las variables (teniendo en cuenta los resultados obtenidos anteriormente $E(Z_i) = 0\ y\ var(Z_i) = \lambda_i$) y llamando Y_i a las variables estandarizadas quedará:

$$Y_i = \frac{Z_i}{\sqrt{\lambda_i}}$$

El cual puede ser re expresado como:

$$X_1 = \underbrace{Z_1.b_{11} + \cdots + Z_M.b_{MD}}_{\hat{X}_1} + \underbrace{Z_{M+1}.b_{(M+1)D} + \cdots + Z_D.b_{D1}}_{e_1}$$

$$\vdots$$

$$X_D = \underbrace{Z_1.b_{1D} + \cdots + Z_M.b_{MD}}_{\hat{X}_D} + \underbrace{Z_{M+1}.b_{(M+1)D} \ldots + Z_D.b_{DD}}_{e_D}$$

Si se re expresa cada $Z_i = Y_i.\sqrt{\lambda_i}$ para cada una de las D ecuaciones:

$$X_j = \underbrace{Y_1.\sqrt{\lambda_1}\,.b_{1j} + \cdots + Y_M.\sqrt{\lambda_M}.b_{Mj}}_{\hat{X}_j} + \; e_D$$

Si se compara este resultado con el sistema de ecuaciones planteado al inicio del análisis factorial:

$$x_1 = l_{11}.F_1 + \cdots + l_{1M}.F_M + e_1$$
$$\vdots$$
$$x_D = l_{D1}.F_1 + \cdots + l_{DM}.F_M \; + e_D$$

Donde cada $l_{ij} = \sqrt{\lambda_i}\,.b_{ij} \qquad con \;\; \mathrm{i} = 1,2,..,\mathrm{M} \qquad \mathrm{j} = 1,2,\ldots,\mathrm{D}$

Se puede ver entonces que estandarizando cada uno de los componentes principales obtenidos del ACP se puede interpretar los resultados del mismo como un análisis factorial, obteniendo así la ventaja de que los factores resultantes puedan tener una interpretación a partir de la rotación de las componentes, lo cual se analizará en la siguiente sección.

Para finalizar esta sección se demostrará cómo cada uno de los l_{ij} puede ser interpretado como la correlación entre la variable inicial "j" y la componente "i". Si se nota vectorialmente X_j y Z_i como los vectores con las N observaciones para la variable "j" y la componente "i" respectivamente se puede calcular su covarianza como :

$$cov(X_j, Z_i) = \frac{1}{N}\, X_j'.\, Z_i$$

Si se agrega un vector δ de dimensión $1xD$ que tenga 0 en todas las posiciones y 1 en la posición "j" se puede notar a X_j' como:

$$X_j' = \delta\,.X' = \; [0 \quad \ldots \quad 1 \quad \ldots \quad 0]\; x \begin{bmatrix} x_{11} & \ldots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{D1} & \ldots & x_{DN} \end{bmatrix}$$

Donde X' es la matriz de datos iniciales traspuesta. Adicionalmente dado que $Z_i = X.b_i$:

$$cov(X_j, Z_i) = \frac{1}{N}\,\delta\,.X'.\, X.b_i = \underbrace{\frac{1}{N}X'.\, X}_{S}\;.\;\underbrace{\delta\;.\;b_i}_{b_{ij}} = \lambda_i\,.b_{ij}$$

$$corr(X_j, Z_i) = r_{ij} = \frac{\lambda_i\,.b_{ij}}{\underbrace{\sqrt{var(X_j)}}_{=1}\,.\sqrt{\lambda_i}}$$

$$r_{ij} = \sqrt{\lambda_i} \, . \, b_{ij}$$

Donde se hace uso de la expresión obtenida para la matriz de correlaciones S y de la definición de coeficiente de correlación para 2 variables. De esta forma se llega a la demostración de que cada uno de los coeficientes l_{ij} representa la correlación entre la variable inicial "j" y la componente "i".

La matriz formada por todos los coeficientes l_{ij} es entonces la **matriz de cargas factoriales** que contiene cada una de las correlaciones entre los factores retenidos y las variables iniciales.

A modo de resumen, los conceptos vistos en este capítulo incluyeron el desarrollo intuitivo y estadístico del análisis factorial, su relación con el ACP, y la unificación de ambos en cuanto a la notación matemática. Por último, se analizó la correlación entre factores y variables, y la matriz de cargas factoriales que incluye todas las combinaciones de estas correlaciones.

CAPÍTULO V

ROTACIÓN DE COMPONENTES

En el presente capítulo se desarrollará el concepto de rotación de componentes y la forma de aplicarlo para permitir que los factores obtenidos del análisis factorial puedan ser interpretados con mayor facilidad.

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 componente[4] va a ser recogida por la segunda. Se puede

[4] 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.

generalizar esto diciendo que la información perdida por las primeras K componentes va a ser recogida por las ultimas M-K componentes.

Matemáticamente la rotación Varimax lo que hace es calcular una variable que llama **simplicidad** ($C_i{}^2$), que es la varianza de los cuadrados de las cargas factoriales (r_{ij}) para un determinado factor (i). Adicionalmente lo que se hace comúnmente es aplicar lo que se llama la Normalización de Kaiser, donde cada r_{ij} se divide por la comunalidad de la variable inicial "j" ($h_j{}^2$). Esto se hace para evitar que las x con mayor comunalidad tengan más influencia (más peso) en la solución final. Una vez calculada esta simplicidad normalizada para cada uno de los factores lo que se hace es **maximizar la sumatoria de todas estas simplicidades** (C^2). Matemáticamente la expresión que se maximiza es:

$$Max\ C^2 = \sum_{i=1}^{M} C_i{}^2 \qquad con \quad C_i{}^2 = \frac{1}{D}\sum_{j=1}^{D}\left(\frac{r_{ij}{}^2}{h_j{}^2}\right)^2 - \left(\frac{1}{D}\sum_{j=1}^{D}\frac{r_{ij}{}^2}{h_j{}^2}\right)^2$$

Dos notas interesantes de casos extremos que se pueden hacer sobre este tema es que**, si no se reduce la cantidad de variables** (M=D) se pueden rotar los ejes como se quiera sin perder información, solo cambian las coordenadas (coeficientes). La otra es que, justamente por la forma iterativa en que se realiza, si solo existiera un componente no se podría realizar ninguna rotación, como mínimo se deben tener dos.

Con la rotación de componentes ya se han desarrollado todos los elementos teóricos sobre el tema necesarios para el desarrollo del índice de actividad económica mensual para la provincia de Mendoza. El mismo se realizará en el siguiente capítulo.

CAPÍTULO VI

CONSTRUCCIÓN DEL ÍNDICE

En este capítulo se desarrollará el índice de actividad económica para la provincia de Mendoza usando el Análisis por Componentes Principales (ACP), a partir de una serie de variables iniciales (las cuales se detallan en la tabla 1). Las cuales se consideran representativas para la actividad económica de Mendoza se intentará obtener una cantidad reducida de componentes que las representen y a partir de los cuales se pueda construir un índice de actividad económica para la provincia. Las series usadas se obtuvieron de la Dirección de Estadísticas e Investigaciones Económicas (DEIE) de la provincia de Mendoza. En tanto que las series para la validación del índice (EMAE, PBG, PBI) se obtuvieron del Instituto Nacional de Estadística y Censos (INDEC) de la Nación

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, que hoy solo existe anualmente.

Tabla 1: Presentación de las series de datos usadas

Variable	Descripción
AUTOMOTRIZ	Ventas Mensuales de Automotores Cero Km por Segmento. Mendoza. Enero 2010 - Noviembre 2018
EELECTRESIDENCIAL	Consumo de Energía Eléctrica Residencial en MWh. Años 2004 - 2019
EELECTCOMERCIAL	Consumo de Energía Eléctrica General/Comercial en MWh. Años 2004 - 2019
EELECTINDUSTRIAL	Consumo de Energía Eléctrica Grandes Demandas/Industrial en MWh. Años 2004 - 2019
ENARGAS TOTAL SISTEMA	Total - En miles de m^3 de 9300 kcal y en porcentaje. Años 2004-2018
HOTELMDZ	Demanda hotelera por mes y condición de residencia de los viajeros hospedados. Ciudad de Mendoza. Años 2008-2018
INDUSTRIA	Índice de ventas industriales a valores constantes y variación porcentual.
INMUEBLES	Total de inmuebles involucrados en operaciones en el Registro Público de la Propiedad, a través de escrituras públicas (corresponden a la 1ª,3ª y 4ª Circunscripción Judicial) y variación porcentual. Mendoza. Enero 2006-Noviembre 2018

PATVEHI	Patentamiento de Vehículos (Autos, Motos y Maquinarias Agrícolas)
SHOPPING	Índice mensual de ventas de mercaderías y servicios en centros de compras a valores corrientes. Año base 2010. Mendoza. Enero 2010-Noviembre 2018
SUPERMERCADOS	Ventas a precios constantes por grupo de artículos, en pesos de 2004. Mendoza. Años 2010-2018
VINO	Despachos de vinos autorizados para ser liberados al consumo. Mendoza. Años 2004 - 2017

Fuente: Dirección de Estadísticas e Investigaciones Económicas (DEIE) de la provincia de Mendoza

Inicialmente se realizó un ACP sobre este conjunto de variables con el software SPSS. Se obtuvo los resultados que se muestran a continuación (tabla 2). Una aclaración es que en todos los casos se utilizó como metodología para rotar los ejes la rotación Varimax.

Tabla 2: Resultado del ACP inicial

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.816	23.469	23.469	2.816	23.469	23.469	2.298	19.150	19.150
2	2.447	20.393	43.862	2.447	20.393	43.862	2.155	17.961	37.111
3	1.753	14.608	58.470	1.753	14.608	58.470	2.070	17.251	54.362
4	1.500	12.502	70.971	1.500	12.502	70.971	1.649	13.742	68.103
5	1.056	8.801	79.773	1.056	8.801	79.773	1.400	11.669	79.773
6	.706	5.884	85.656						
7	.491	4.095	89.752						
8	.403	3.360	93.112						
9	.344	2.868	95.979						
10	.207	1.722	97.701						
11	.177	1.472	99.173						
12	.099	.827	100.000						

Rotated Component Matrix

	Component				
	1	2	3	4	5
AUTOS	.128	.152	.302	-.830	.174
EELECTCOMER	.040	.136	.898	-.048	-.037
EELECTIND	-.102	-.725	.170	.365	-.062
EELECTRES	-.389	-.151	.586	.016	.594
GASTOTAL	.138	.134	-.226	-.163	.856
HOTELMDZ	.017	-.110	.653	-.281	-.211
INDUSTRIA	.853	.082	-.086	.002	-.098
INMUEBLES	.849	.061	.199	.169	.109
PATVEHI	-.068	.890	-.053	-.043	.025
SHOPPING	.149	.855	.228	.107	.016
SUPERMERCADOS	.449	-.028	-.026	.823	-.017
VINO	.649	-.031	-.470	-.026	.460

Communalities

	Initial	Extraction
AUTOS	1.000	.849
EELECTCOMER	1.000	.830
EELECTIND	1.000	.702
EELECTRES	1.000	.871
GASTOTAL	1.000	.848
HOTELMDZ	1.000	.562
INDUSTRIA	1.000	.751
INMUEBLES	1.000	.805

PATVEHI	1.000	.802
SHOPPING	1.000	.816
SUPERMERCADOS	1.000	.881

VINO	1.000	.856
Extraction Method: Principal Component Analysis.		

Dado que estos resultados nos dejan cinco componentes principales, se dificultaría obtener un índice único. El proceso que se fue haciendo implicó sacar algunas variables en base al peso que tuvieran en los componentes principales, descartando las de menor peso relativo. También se usó como criterio el número de observaciones disponibles para esa variable. Ya que no de todas se disponía de la misma cantidad de observaciones, se priorizó mantener aquellas que tuvieran más. El resultado final de este proceso de iteración es el que se muestra en la tabla 3.

Tabla 3: Resultado de ACP final

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.403	40.054	40.054	2.403	40.054	40.054	2.325	38.753	38.753
2	1.673	27.879	67.933	1.673	27.879	67.933	1.751	29.180	67.933
3	.881	14.689	82.622						
4	.564	9.398	92.020						
5	.368	6.132	98.152						
6	.111	1.848	100.000						

Rotated Component Matrix

	Component	
	1	2
EELECTCOMER	.929	.052
EELECTRES	.908	.172
GASTOTAL	.126	.733
INMUEBLES	.038	.701
VINO	-.433	.761
HOTELMDZ	.657	-.333

Communalities

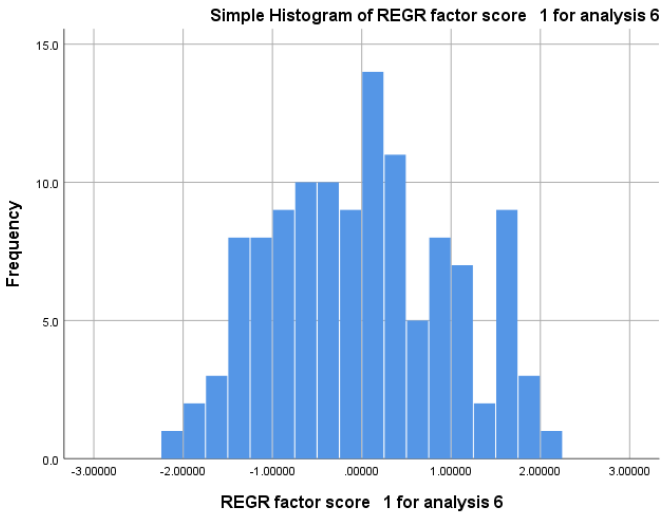
	Initial	Extraction
EELECTCOMER	1.000	.866
EELECTRES	1.000	.855
GASTOTAL	1.000	.553
INMUEBLES	1.000	.492
VINO	1.000	.767
HOTELMDZ	1.000	.542

De acuerdo a estos resultados finales se usan seis variables representativas del nivel de actividad mensual y se obtienen seis factores principales. De las seis variables EELECTCOMER, EELECTRES y HOTELMDZ están muy representadas en el primer factor, en tanto que GASTOTAL, INMUEBLES y VINO lo están en el segundo factor.

Dados estos dos factores obtenidos, cuyos histogramas están representados en el gráfico 7, se los combina algebraicamente sumándolos, para así obtener el índice de actividad que se estaba buscando.

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 la Tabla 4 y Gráfico 8 respectivamente

Gráfico 7: Histogramas de los factores



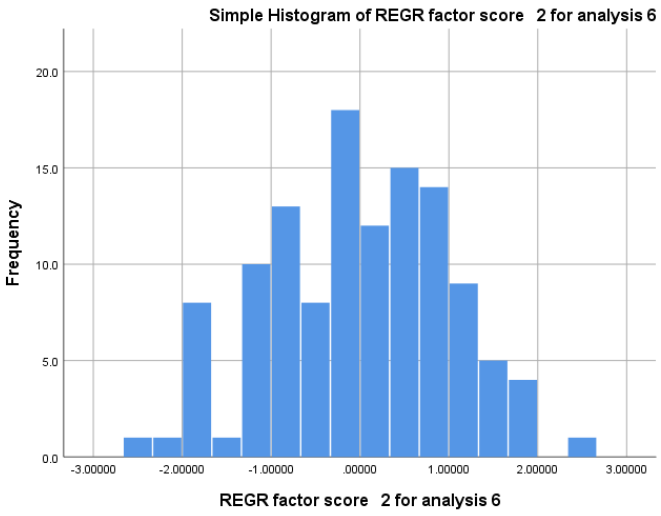
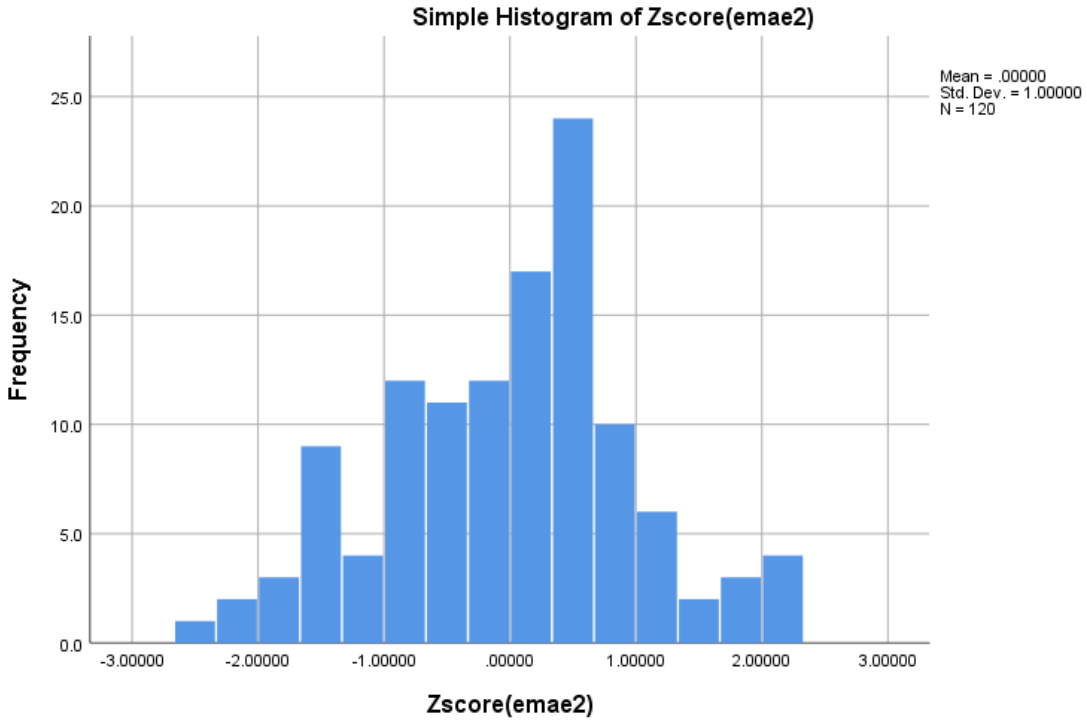


Tabla 4: Estadísticas descriptivas EMAE

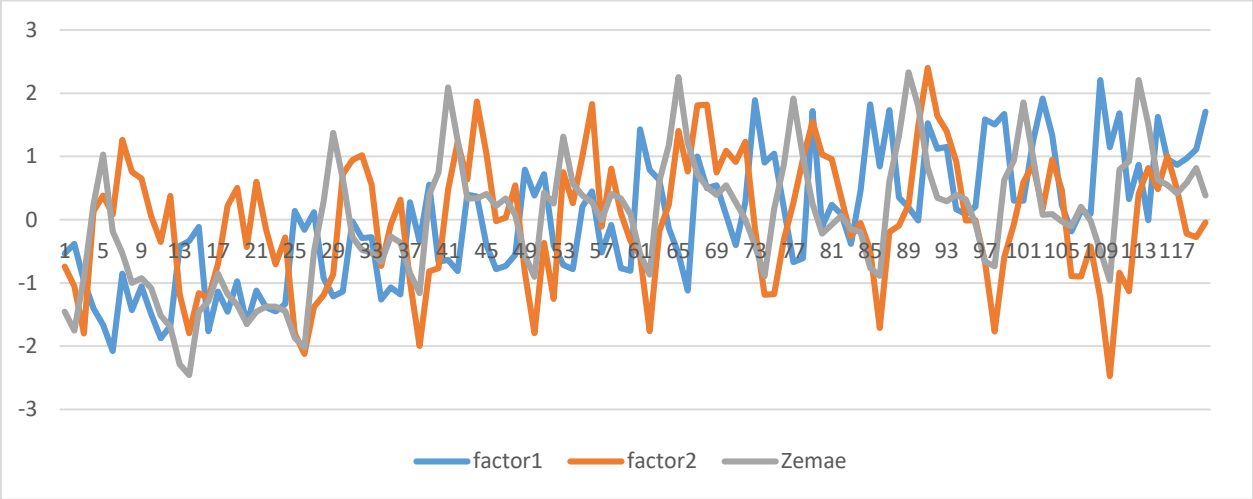
	N	Minimum	Maximum	Mean	Std. Deviation
emae2	120	114.90	168.90	142.5817	11.28243
Valid N (listwise)	120				

Gráfico 8: Histograma EMAE



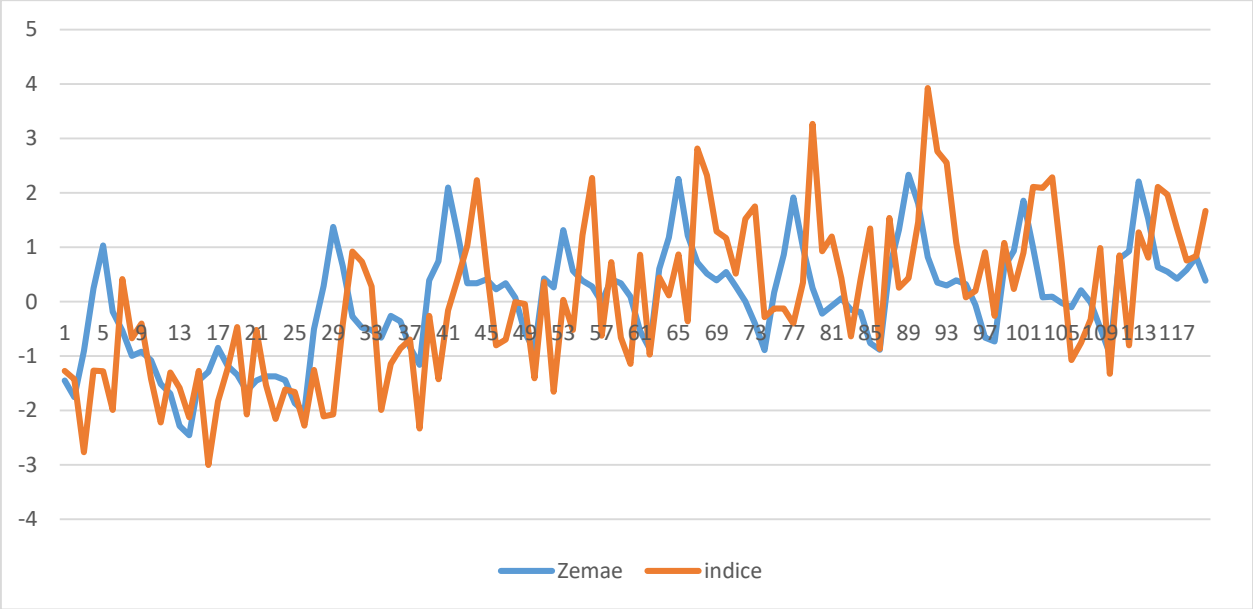
Luego de estandarizado el EMAE se lo contrasta gráficamente con respecto a ambos factores (Gráfico 9) y con respecto al índice construido (Gráfico 10).

Gráfico 9: Contrastación entre los factores y el EMAE



Fuente: elaboración propia a partir de datos del INDEC

Gráfico 10: Contrastación ente el IAEMEN y el EMAE



Fuente: elaboración propia a partir de datos del INDEC

Como se puede observar en el Gráfico 11, sí hay una equivalencia en las variaciones de los niveles de actividad a nivel nacional con respecto a lo que muestra el índice construido para Mendoza. Adicionalmente se puede observar un aparente rezago del índice respecto del EMAE, para corroborar esto se intentará ver si este rezago existe también entre el PBG de la provincia y el PBI nacional que serían los

equivalentes de estas series, pero con periodicidad anual. Esto último se muestra en los gráficos 11 y 12 respectivamente.

Gráfico 11: Contrastación entre PBG y PBI

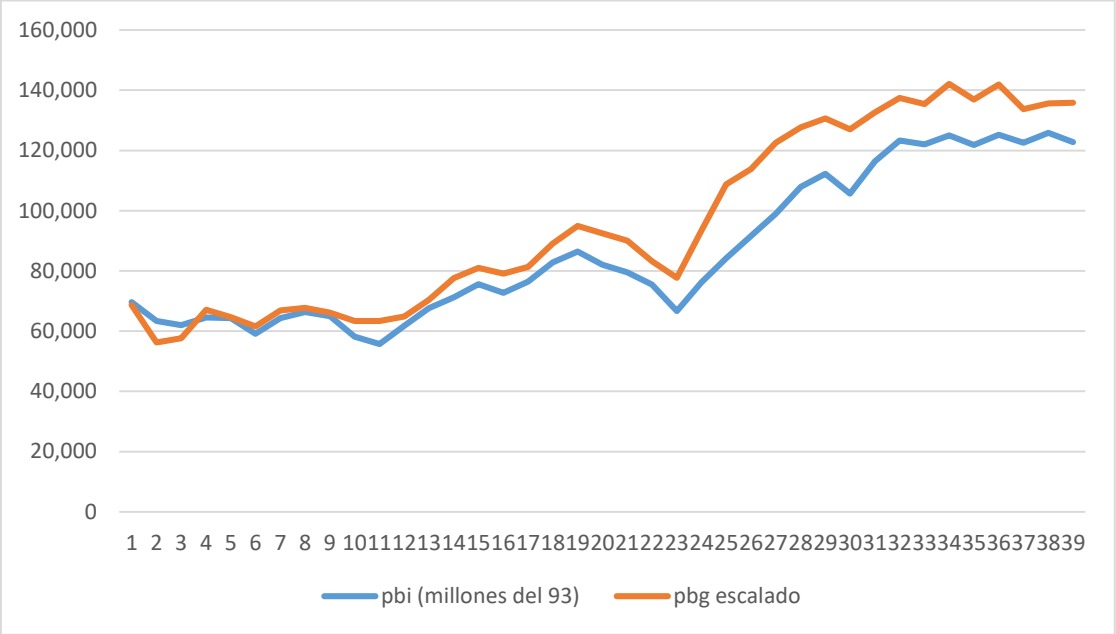
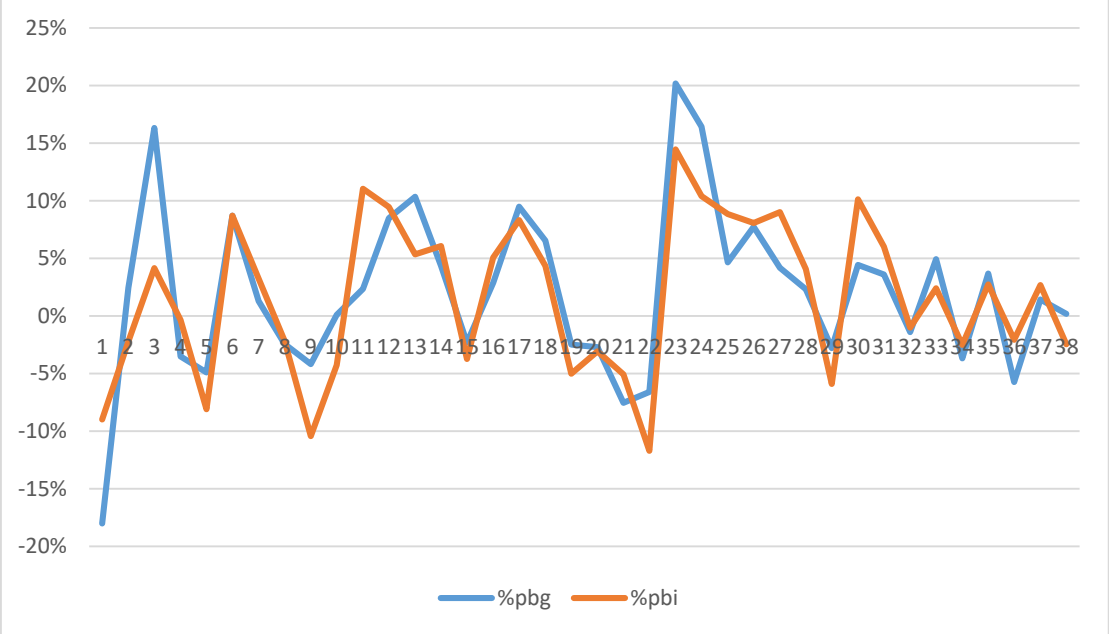


Gráfico 12: Variaciones porcentuales entre PBG y PBI



Fuente: elaboración propia a partir de datos del INDEC

Al analizar los resultados obtenidos, se puede concluir que a través del método de ACP se logró crear un índice que representa la actividad económica de la provincia. Al comparar este índice respecto al EMAE hay una equivalencia con algunos períodos de atraso, lo cual indicaría que las variaciones que se

producen a nivel nacional se transmiten luego a la provincia. Al momento de intentar contrastar esta hipótesis haciendo uso del PBG y PBI, la misma pareciera no cumplirse. Se debe tener en cuenta que la periodicidad es distinta en ambos casos ya que justamente el índice se crea porque no hay ninguna estadística con esa periodicidad para la provincia. Para el caso del PBG y PBI ambas series tienen variaciones que son bastante simultaneas. Quedaría como posible caso de estudio profundizar en el por qué de este rezago al hacer el análisis mensual.

Finaliza aquí este último capítulo donde se desarrolló el índice que era objetivo del presente trabajo, y en el cual se lograron captar los patrones de comportamiento propios de la economía local y su relación con la actividad económica del país.

CONCLUSIONES

Para concluir el presente trabajo se expondrán algunas conclusiones sobre el tema y algunos 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 ($\mathbb{R}^2 \to \mathbb{R}$). 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 ($B_{ij} = {x_j}^T . b_i$).

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 atraso, lo cual indicaría a priori que las variaciones que se producen a Nivel Nacional se transmiten luego a la Provincia de Mendoza. Al momento de intentar contrastar esta hipótesis haciendo uso del Producto Bruto Geográfico de Mendoza y Producto Bruto Interno de Argentina, la misma pareciera no cumplirse (teniendo en cuenta que la periodicidad es distinta en ambos casos ya que justamente el índice se crea porque no hay ninguna estadística con esa periodicidad para la provincia) sino que son bastante simultaneas las variaciones; esto podría deberse a la estacionalidad propia de los indicadores mensuales. 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*. Editorial Académica Española. Saarbrücken.

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

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

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

PÉREZ, Cesar (2004). Técnicas de análisis multivariante de datos: Aplicaciones con SPSS. Pearson Prentice Hall. Madrid.

RAICHMAN, Silvia & Totter, Eduardo (2016). Geometría analítica para ciencias e ingenierías. Universidad Nacional de Cuyo. Mendoza.

UNCUYO UNIVERSIDAD NACIONAL DE CUYO — FACULTAD DE CIENCIAS ECONÓMICAS

DECLARACIÓN JURADA RESOLUCIÓN 212/99 CD

El autor de este trabajo declara que fue elaborado sin utilizar ningún otro material que no haya dado a conocer en las referencias que nunca fue presentado para su evaluación en carreras universitarias y que no transgrede o afecta los derechos de terceros.

Mendoza, 09-12-19

Fernando Delorat

Firma y aclaración

26619

Número de registro

35184347

DNI