VNIVERSIDAD NACIONAL
AVTONOMA DE
MEXICO

FACULTAD DE ECONOMIA.

DIVISION DE POSTGRADO.

LOS METODOS ESTADISTICOS MULTIVARIADOS EN ENCONOMIA: ALCANCES Y LIMITACIONES.

-una crítica-

TESIS PARA OBTENER EL GRADO DE MAESTRIA EN ECONOMIA.

HERNANDO ENRIQUE MUTIS GAITAN.

MEXICO, D.F. 1987.

INTRODUCCION.

Este trabajo se propone exponer con relativa profundidad la naturaleza de los principales métodos estadísticos multivariados para entender los alcances y limitaciones a los que da lugar cuando se los quiere utilizar en el dominio económico.

Desde hace mucho los métodos estadísticos univariados han tenido importancia indiscutible en el proceso de consecusión, organización, procesamiento y análisis de la información económica -así como su tratamiento inferencial-, todo dentro de un marco que se plantea adentrarse en la realidad con un arsenal teórico determinado. No pocas veces los problemas teóricos y técnicos del tratamiento de la información han traído como consecuencia la comprobación o refutación de postulados económicos. Para no mencionar sino un ejemplo, en el campo de la política económica esta actividad de búsqueda y ordenación de la información económica ha conllevado a definir en lo más general, propósitos y en lo más concreto, determinar metas.

Los análisis estadísticos multivariados no han gozado de popularidad semejante. En parte, por su relativa novedad; en parte por su aparente difícil tratamiento estadístico-matemático; en parte, por los problemas de computación y cálculo que ellos implican; y, en parte, porque la interpretación de sus resultados no siempre es muy provechosa. La mera posibilidad de resumir grandes volúmenes de información no es un problema referido sólo a su computación y tratamiento aritmético y/o contable. De ser sólo eso, revestiría interés secundario. Lo que principalmente importa es unir este aspecto a la potencialidad que encierra el tratamiento simultáneo de múltiples variables bajo un manejo conceptual estadístico-matemático de relativa complejidad que sea capaz de arrojar conclusiones lógicas y distinguibles para el análisis económico.

Los adelantos computacionales han hecho posible compendiar, organizar y analizar grandes cantidades de información con relativa facilidad. Con alguna anterioridad a estos adelantos y posteriormente favorecidos por ellos, se han desarrollado un conjunto de técnicas estadísticas que se conocen como métodos estadísticos multivariados (M.E.M).

En principio podríamos considerar que las técnicas estadísticas multivariadas abandonan el conocido "ceteris paribus" de la teoría económica convencional en el sentido de que permiten incorporar en el análisis muchas variables en forma simultánea en tanto se trata de ordenar, organizar y presentar las múltipes relaciones establecidas entre ellas: de esta forma, se supone, obtendríamos un cuadro panorámico más articulado con la multiplicidad de las relaciones que se establecen entre las variables y, a posteriori, nos permite estudiarlas separadamente.

Podemos acercarnos más al significado de las técnicas estadísticas multivariadas remitiéndonos a todos aquellos métodos que analizan simultáneamente múltiples mediciones; la notación multivariada se refiere a tcdas aquellas variables que están interrelacionadas de tal forma que sus diferentes efectos no pueden ser estudiados fácilmente de manera separada, o no es conveniente hacerlo. Por este camino, un propósito del análisis multivariado es medir, describir, explicar y predecir el grado de relación entre las variables.

Con los análisis multivariados podemos abordar el tratamiento simultáneo de diversos problemas -explicitados y cuantificados en distintas variables medidas todas ellas en diferentes unidades o elementos-, y obtener resultados que no riñan con la lógica y que puedan formar parte -y enriquecer empíricamente los planteamientos o hipótesis- en el campo de la economía. Aspecto importante de nuestro estudio es escudriñar hasta dónde podemos acudir empíricamente a los M.E.M. cuando se trate de apoyar algún planteamiento teórico.

Los problemas de la aplicación de los M.E.M. a la economía nos conducen a las contradicciones que se presentan entre la coherencia propia del análisis estadístico-matemático y la coherencia propia de las cuestiones económicas. El sentido del trabajo y por ende, la inquietud fundamental tiene que ver precisamente con la posibilidad de combinar los M.E.M. con la investigación económica. Este primer gran problema se parte en varios interrogantes: cuál es la naturaleza de los métodos multivariados ? Qué problemas especiales se presentan para su aplicación en el campo de la economía ? Cuáles aplicaciones han tenido mayor eco en los estudios económicos en México ?

Ante estas primeras preguntas brotan otras entre las que se cuenta la referente a las relaciones más generales entre economía y estadística, definida esta relación en una primera instancia, como el vínculo entre un planteamiento teórico y su comprobación empírica; en segunda instancia, por el enriquecimiento teórico que puede obtenerse al contemplar una cierta información económica a la luz de un método estadístico determinado. Desde luego, ambas esferas dan como resultado productos teóricos diferentes en consonancia a los distintos enfoques de la economía y la matemática que se sostienen. En esta perspectiva, un interés es discutir un poco las formas como ciertas escuelas de economía han enfrentado el problema. El abordar estas discusiones conlleva una referencia -explícita en algunas ocasiones e implícita en las más de las veces- a la forma como se concibe esta relación economía- estadística desde el punto de vista del método en su sentido más general. Es esta la razón por la que se vio la necesidad, sin abandonar el marco propuesto, de hacer una reconstrucción personal sobre los enfoques tradicionales y modernos de los problemas del método y de la comprobación empírica de una formulación teórica.

Dentro de este panorama se discuten algunos temas asociados al papel mítico que ordinariamente se le concede a la matemática como culmen de la pureza científica, junto con el cuestionamiento de la correlación y la predicción como afirmación y comprobación de los nexos causales.

expectativas y los resultados, entre las propuestas y los logros de una particular técnica multivariada. Por otro lado, al interior de la técnica en sí misma, se dilucidan cuestiones referentes al análisis e interpretación de los parámetros obtenidos. Lo primero es un problema que se trata en la esfera que vincula a la teoría económica con el instrumental estadístico. Lo segundo, es un análisis sobre el conocimieto y manejo de la técnica invocada. Es decir, la inquietud es la comparación entre lo esperado y lo obtenido, sopesando el rol de comprobación empírica que puede admitir una técnica determinada.

Dentro de la presentación de los métodos multivariados corresponde el turno al análisis discriminante. Se enfatiza mediante una exposición gráfica los propósitos del procedimiento y se hace especial insistencia en las condiciones que lo hacen posible, ubicando de manera particular ciertos problemas del que está impregnado por su intención modelizadora, justipreciando la validez de sus resultados.

En el análisis de cúmulos o de conglomerados, además de presentar los objetivos para los cuales fue diseñado, se esbozan los algoritmos más usados para su solución. Subrayamos en este capítulo la importancia de la habilidad analítica para escudriñar y encontrar las agrupaciones haciendo un especial énfasis en los diversos conceptos de distancias y jerarquizaciones anidadas que forman parte integrante del procedimiento mencionado.

Como la última técnica multivariada que se presenta, se definen los propósitos y la matemática matricial que hacen posible la correlación canónica. La intención inicial era enjuiciar las aplicaciones de estas tres técnicas -corrrelación canónica, cúmulos y discriminantes- con un enfoque similar al realizado en el capítulo sobre los componentes y factores. Desafortunadamente, esta aspiración quedó irrealizada ante la ausencia de trabajos de esta índole en investigaciones económicas.

Aparte de estos aspectos, que forman la espina dorsal de esta propuesta, se debe mencionar que una de las pretensiones de este trabajo es el que pueda cumplir con la tarea de ofrecer una guía introductoria a los M.E.M. Por este motivo se ha hecho un deliberado esfuerzo por privilegiar la explicación intuitiva con el fin de que el lector se pueda apropiar con rapidez de la idea central de cada procedimiento y, por lo menos en principio, no se vea entorpecido por la matemática que se está manejando. En otro contexto, hubiese bastado simplemente con un tratamiento formalizado que culminara con la sola mención de los resultados.

No obstante, se reconoce que para entender a cabalidad la lógica interna de los métodos -y por lo tanto, también sus limitaciones- es imprescindible comprender el desarrollo matemático. Si bien es cierto que por lo general estos temas no forman parte del plan de estudios de la carrera de economía -y a veces tampoco de los cursos de postgrado- es conveniente tener una cierta familiaridad con aspectos básicos de la estadística y del álgebra matricial, formación que sí se incluye en los planes regulares de estudios. Se anexa al final del material un apéndice que intenta responder a los temas no contemplados en estos planes y que tiene el propósito de facilitar la lectura y comprensión de algunos aspectos centrales de la matemática multivariada.

Este objetivo pedagógico adquiere plena validez ante la carencia de textos que respondan a estas inquietudes. La bibliografía sobre estudios multivariados exige una preparación especializada para abordarla y por lo regular no intenta discutir, más allá de un nivel estrictamente empírico, los problemas de la aplicación de una determinada técnica a un campo teórico específico, pues su objetivo básico es responder a la consistencia interna de la matemática; su prescripción por excelencia para la aplicación de cualquier procedimiento no va más allá del "úsese en caso de necesidad". Esa es, tal vez, su principal limitación.

TEMATICA.

El propósito de cualquier análisis multivariado es obtener una cierta comprensión de la estructura de la información y de sus principales rasgos. Dentro de estas situaciones caben varias alternativas según las preguntas que nos hagamos y el tipo de análisis que queremos desarrollar. En principio, con los métodos multivariados podemos afrontar tareas como las siguientes:

1.- Estudiar las relaciones entre grupos de variables. Se busca encontrar las correlaciones entre dos grupos de información -cada uno de ellos multivariado- destacando las más altas, detectando la estructura de correlaciones entre ambos considerándolos simultáneamente; el principio subyacente es desarrollar una combinación lineal de cada grupo de variables con el otro de una manera tal que maximice la correlación entre estos dos conjuntos. Este análisis se considera la expresión multivariada de la regresión múltipe y se conoce como correlación canónica.

A partir de la correlación canónica puede construirse un modelo predictivo que busque explicar o predecir simultáneamente varios fenómenos, basado en el conjunto de correlaciones canónicas: variaciones en las tarifas arancelarias, en el tipo de cambio, en los precios internacionales por tipo de producto, en la tasa de expansión del producto y mercado mundiales, en las tasas de interés mundiales.... pueden estar correlacionados con otro grupo de variables como saldos de balanza comercial y presupuestal internas, tasa de crecimiento de sectores o tipos de producto de la industria nacional, de precios, de salarios, dé productividad, de empleo.... La idea es que dada la estructura simultánea de correlaciones podemos encontrar las variables que más pesen en el conjunto y además establecer el comportamiento de un grupo ante cambios en el otro. Desde luego, en ésta y en las otras alternativas, el proceso de definición del problema y de las variables así como su posible interpretación le compete a la teoría económica.

2.- Discriminar y clasificar nuevos elementos. Se pueden investigar las relaciones entre unidades y mediciones de diversas variables con agrupamientos ya conocidos de la información para encontrar una regla de asignación que nos indique la pertenencia de un nuevo elemento a uno de los grupos ya considerados. La regla de asignación se asocia con cierta probabilidad de clasificación incorrecta de tal manera que podemos manejar, o mejor, reducir la probabilidad de clasificar erróneamente los elementos. Es útil este análisis en situaciones en las cuales el conjunto de observaciones puede ser dividido en grupos basados en clases ya de antemando conocidas. Este método se conoce como análisis discriminante.

Si queremos comparar el nivel de desarrollo o de crecimiento industrial de algún país africano o asiático con los países de Latinoamérica a los que previamentae hemos supuesto que se subdividen en tres niveles de desarrollo o de crecimiento industrial -alto, medio y bajo- podríamos considerar una función discriminante que retomando variables homogéneas a las consideradas en el análisis de Latinoamérica, nos informe sobre el grupo al que comparativamente más se parece el país inicialmente mencionado. Otro ejemplo lo podríamos obtener del mundo de los negocios: el método discriminante para concesiones de crédito. Al solicitante -individuo, entidad o país-, se le evalúan mediciones de varias variables y de acuerdo a los resultados arrojados por la función discriminante se decide otorgarle o no el préstamo.

3.- Reducir la dimensionalidad. Se trata de expresar el contenido principal de la información en dimensiones más reducidas -menores y/o nuevas variables-, con el objeto de hacer más entendible, manejable y funcional la información, mediante la búsqueda de variables linealmente relacionadas para evitar la redundancia que se presenta en la información original. Estos métodos se conocen como análisis de componentes principales.

Pueden ser utilizados en econometría y regresión como método de selección de variables; pueden simplificar la información para realizar correlación canónica o análisis discriminante. En estos últimos casos nos pueden ayudar a la escogencia de un subgrupo de variables dentro de un conjunto más amplio para los ejemplos mencionados anteriormente. La reducción de la dimensionalidad es la columna vertebral de las técnicas multivariadas.

4.- Analizar factores. Se trata de explicar un grupo de información multivariada en términos de un pequeño número de factores subyacentes a esa estructura de la información -y que no son de fácil detección-, analizando las interrelaciones e interdependencias de un grupo de variables y explicar éstas en términos de los factores comunes subyacentes: se "condensa" la información contenida en las variables originales en un subconjunto de nuevas variables que son combinación lineal de las más importantes de las originales. A este subconjunto lo llamamos factores y la técnica se conoce como análisis de factores.

5.- Búsqueda de agrupaciones. El propósito es encontrar subgrupos estadísticamente significativos de un conjunto de unidades multivariadas cuando no tenemos información a priori sobre la existencia de tales agrupaciones. A diferencia del análisis discriminante, que ubica un nuevo elemento en alguno de los grupos ya definidos, aquí tratamos de encontrar o corroborar estos grupos. Se trata de agrupar de acuerdo a una medida de similaridad o asociación que nos haga mínima la diferencia entre subconjuntos con el fin de determinar las posibles agrupaciones latentes en toda la información. Tales métodos corresponden a lo que se ha llamado como análisis de cúmulos.

Un ejemplo es el tratar de encontrar en los paises de América Latina la existencia de cúmulos según el nivel de desarrollo de acuerdo a un conjunto de información multivariada de cada uno de ellos, cuando ignoramos las posibles agrupaciones. El análisis de cúmulos intenta encontrar estos grupos minimizando las similitudes y maximizando las diferencias. De paso se pueden configurar subgrupos anidados en otros más grandes que a su vez forman parte de otro conjunto.

* * * * * *

Quiero agradecer, en primer lugar, al Mtro. Carlos Salas Páez por la asesoría y apropiada dirección en la realización de este trabajo. Así mismo, a los profesores del I.I.M.A.S., Dr. Rubén Hernández Cid y Mtro. Raymundo Peralta Herrera por las facilidades ofrecidas que hicieron posible la culminación de este material.

También quiero agradecer la colaboración y apoyo tan especial que recibí de mis compañeras Eva y Patricia. Finalmente, agradezco la cooperación de mi esposa, Leticia.

Capitulo 1.

COMENTARIOS GENERALES.

1.1 NOTAS SOBRE LAS RELACIONES ECONOMIA Y ESTADISTICA.

De entrada podemos plantear que es innegable la necesidad de la economía en acudir a la estadística. Pero este incuestionable punto de partida va a someterse a distintas modalidades y tratamientos de acuerdo, por una parte, a la concepción teórica que adoptemos para la economía, y por otra, al sentido, perspectivas, posibilidades y limitaciones que brinda la propia estadística.

Es menos discutible ubicar la utilidad de la estadística cuando la concebimos como proceso de organización, recopilación, clasificación, síntesis y tratamiento de la información -tareas a las que responde la estadística descriptiva tradicional- que cuando la asumimos como procedimientos inferenciales y probabilísticos. Esto también acontece con técnicas más sofisticadas -métodos econométricos, análisis de series de tiempo, técnicas multivariadas...-, las cuales, a pesar de tener una gran carga descriptiva han sido objeto de un gran debate acerca de su utilización e importancia.

Los problemas de la articulación economía-estadística se deben abordar según la concepción teórica que se adopte para la economía. Es un problema que atañe básicamente a los diferentes enfoques que existen sobre la economía. Las escuelas económicas más importantes -clásica, marxista, neoclásica y lo que llamamos la gran corriente postkeynesiana, conformadas éstas en muy

> utilidad marginal, fueron considerados científicos porque se les disfrazó de expresiones matemáticas y permitían derivar leyes que tenían cierta analogía formal con la lógica."[2]

Estos, llamemos, 'desajustes' entre matemáticas y economía no se circunscriben al reino de los neoclásicos. Piero Sraffa, por ejemplo, ofrece un enfoque de la determinación de valores y precios diferente a la construcción neoclásica y también cae en una utilización inapropiada de las matemáticas.[3]

1.1.1 ACERCA DE LA COMPROBACION EMPIRICA

Los planteamientos pioneros de los neoclásicos no fueron fácilmente objeto de contrastación empírica. Esta característica se desarrolla mucho después. Con la elaboración de datos y de indicadores, con los métodos de estimación y de prueba de hipótesis, la estadística va a vincularse con los planteamientos económicos, inspirados éstos en las nociones positivistas. A estas alturas la demostración y validez de una teoría tendría que ver no sólo con la propia lógica de su desenvolvimiento, sino también con las manifestaciones empíricas que el hecho teórico presente. Es por esto que la especificación de las relaciones entre categorías toma la forma de expresiones funcionales entre variables, convirtiendo el procedimiento de estimación y prueba de hipótesis en parte integrante de las teorías económicas neoclásicas postkeynesianas.

Esta relación entre teoría y comprobación ha encaminado a la teoría económica de una manera que su construcción es realizada de tal forma que pueda ser susceptible de una verificación estadística sistemática. De ahí el papel tan importante que tenga la cuantificación de los conceptos económicos. Una buena

----

2. Valle, p. 18.

3. Véase el interesante trabajo que Salas y Valle desarrollan criticando el uso que hace Sraffa de las matemáticas. Cfr. bibliografía.

ilustración de esta dimensión que ha tomado la teoría es la discusión entre monetaristas y keynesianos: es la elastidad de la demanda de dinero igual o menor que la unidad ? Es la elasticidad demanda de inversiones significativamente menor que uno ? Es la velocidad de circulación del dinero relativamente estable ?... La respuesta estadística a estos interrogantes alínea a los teóricos en alguna de las corrientes, de esta manera, el problema es cómo formular teóricamente la pregunta para que pueda ser sujeto de contrastación empírica.[4]

Para algunos este cambio de ruta significó que la economía apropiándose del método de las ciencias naturales se elevaba a un plano de refinación nunca antes alcanzado. El premio Nobel Ragnar Frisch en un artículo titulado "De la Teoría Utópica a las Aplicaciones Prácticas: el Caso de la Econometría", comenta lo siguiente:

> "...En la primera mitad del siglo XX ... los propios teóricos se dedicaron en forma sistemática a la construcción de la teoría en forma tal que ésta pudiera ponerse en contacto inmediato con el material de observación. Podría afirmarse que a partir de entonces la economía pasó a la etapa en que las ciencias naturales habían estado durante largo tiempo, o sea la **etapa en que la teoría obtiene sus conceptos de la técnica de observación y a su vez la teoría influye sobre la técnica de observación...**"

> "...Por primera vez en la historia, ahora parecia que el trabajo realizado en el frente teórico de la economía ahora formulado en gran medida en términos matemáticos y el trabajo realizado en el frente descriptivo exterior debían convergir y apoyarse de una manera recíproca, lo que produciría una teoría lo suficientemente refinada para retener al material concreto de la observación, y al mismo tiempo una masa

----

4. En el libro de Desai hay una buena exposición sobre el camino que ha tenido esta discusión. En particular, véase el capítulo tercero. Cfr. bibliografía.

> de observaciones planeadas y ejecutadas con la intención de marcarlas dentro de la estructura teórica."[5].

Esta aparente carga de cientificidad adquirida por la matematización de la economía lo hace afirmar que:

> "Mientras la teoría económica trabaje sobre una base puramente cualitativa , sin tratar de medir la importancia numérica de los diversos factores, podrá obtenerse y defenderse cualquier 'conclusión'."[6].

De reconocer el aporte, el premio Nobel hace su apología, desvirtúa el problema del método en la economía y se ubica dentro de una particular corriente del pensamiento económico. Desde luego, dentro de esta visión positivista difícilmente aceptaría categorías sin cuantificación: la plusvalía tendría un sentido real si es posible cuantificarla y si pasa las más duras pruebas de hipótesis de la estadística inferencial.

Por nuestra parte no se trata de negar ciertos aportes de la teoría estadístico-matemática, al contrario reconocemos su necesidad, lo que no significa empotrarla como el criterio máximo de verdad. Sobre este punto esperamos volver más adelante.

1.1.2 EL SIGNIFICADO DE LA PRUEBA DE HIPOTESIS.

Sigamos examinando estas relaciones entre teoría económica y estadística, haciendo un comentario sobre el significado de lo que es una prueba de hipótesis y su uso en los análisis económicos. Si de un cierto planteamiento teórico hemos llegado

----

5. Frisch, p. 32. Enfatizado en el original.

6. Frisch, p. 33.

a postular que hay una asociación fuerte entre la tasa de crecimiento de la oferta monetaria y la variación en el nivel de precios y nos interesa la comprobación estadística de esta afirmación, levantamos información muestral de estas dos variables [7] y con una medida de la asociación como el coeficiente de correlación -R- de Pearson [8] determinamos si este valor encontrado puede darle respaldo empírico a la afirmación inicial.

La prueba de hipótesis establece una regla de procedimiento que nos permitiría definir -con una cierta probabilidad- si la correlación particular observada obedece a un patrón sistemático o si su valor se debe a efectos aleatorios. Hemos partido de información muestral para inferir sobre las características generales -o poblacionales, como se estila en el medio- de estas dos variables; estamos haciendo inducción.

Pues bien, en este proceso de la prueba de hipótesis estadística se puede ser muy estricto o muy elástico, de acuerdo a la gravedad del error que estemos más inclinados a evitar debido a la presencia de la inducción. Si aceptamos como falsa una proposición que en realidad es verdadera, estamos cayendo en el denominado error tipo I; en cambio, si asumimos como verdadera una aseveración que en realidad es falsa, incurriríamos en el error tipo II. En la mecánica inferencial tales tipos de error siempre están presentes y no hay forma de aislar estos dos efectos pues su traslape nos indica que los dos riesgos están involucrados simultáneamente. Si creemos que el riesgo más peligroso es el segundo -y por lo común este es el mayor temor-, la idea en la prueba de hipótesis es la de disminuir la probabilidad de caer en el error de tipo II, es decir, se intenta reducir al mínimo la 'posibilidad' de aceptar como verdadero algo que en realidad es falso.

----

7. Los valores de las variables para encontrar esa correlación particular por regla general son obtenidos a través de información muestral, mediante la utilización de los principios del muestreo estadístico.

8. Adelantándonos un poco a la próxima sección, definimos el coeficiente de correlación -R- como la razón entre la variabilidad conjunta -covarianza- de dos variables con respecto al producto de sus variabilidades internas -es decir, el producto de sus desviaciones estandard-. **R=Cov[x,y]/Sx.Sy.**

Por estas razones se plantean dos hipótesis. Una, que expresa que la correlación poblacional y verdadera entre las dos variables es inexistente a la que llamamos hipótesis de nulidad y la presentamos como:

$$Ho.: R = 0$$

La otra es la hipótesis de investigación que queremos comprobar de que la correlación es distinta de cero. A ésta la llamamos hipótesis alternativa:

$$Ha.: R \neq 0$$

Así, expresado en este par de hipótesis estamos descargando la aversión a cometer error tipo II: porque queremos comprobar la existencia de correlación, intentamos probar lo contrario, la ausencia de correlación. Afirmar que la hipótesis nula es cierta quiere decir que si partimos de la consideración de que es inexistente la correlación entre las variables, llegamos a que la distribución de todos los posibles valores del **R** particular encontrado tiene características aleatorias: estamos afirmando que la verdadera correlación es cero. Por este motivo si la magnitud específica de la correlación muestral encontrada cae fuera del rango de los posibles valores aleatorios cuando **R = 0**, entonces rechazaríamos la hipótesis nula de que no existe correlación y podríamos aceptar -al menos provisionalmente- que las variables están correlacionadas. La definición de ese rango de valores nos la suministra el estadístico de prueba (Z, T, F,..).

Esto es lo que llamamos la inferencia estadística: se afirma la ausencia de correlación en términos probabilísticos y no con una certeza definida. Por esto cuando se afirma que con un 95 % de confianza podemos rechazar la hipótesis nula estamos suponiendo que de cada 100 casos que examinemos con un procedimiento igual al anterior, sólo cinco de estos contradicen lo afirmado. De la misma manera, si el nivel de confianza es del 99 % , significa que en 99 de cada cien casos, al encontrar R muestrales, estaremos rechazando que no existe correlación; el aumentar el nivel de confianza equivale a la disminución de la probabilidad de aceptar como verdadero algo que realmente no lo es.

Dos comentarios es necesario mencionar a continuación:

1.) No es posible obviar los problemas que están implícitos en los riesgos de aceptar determinado tipo de error. Por regla general la salida que se intenta es la de disminuir el peligro de cometer error tipo II: hay mayor disposición para aceptar como falso algo verdadero, que aceptar como verdadero algo falso. La analogía se puede hacer con el ya trajinado ejemplo de que es preferible salvar a un culpable que condenar a un inocente.

2.) Se construye la teoría de una forma tal para que pueda ser manipulada y operada en lenguaje matemático y así mismo sea susceptible de verificación a través de la cuantificación y estimación de los parámetros. El planteamiento teórico tiene una traducción inmediata en términos estadísticos y su dinámica tiene un derrotero como el siguiente:

CONCEPTOS ----› VARIABLES ----› FUNCIONES ----› ESTIMACIONES ----› PRUEBAS DE HIPOTESIS ----› VALIDACION ----› CONCEPTOS ....

Cabría aquí alguna pregunta: es siempre posible expresar un planteamiento económico mediante un modelo susceptible de comprobación empírica ? La respuesta es no. Aceptamos que la economía postkeynesiana intenta arreglárselas para poder ofrecer un contenido empírico que pueda ser demostrable, pero sus logros no han sido tan acertados como parece. El criterio de la práctica social así lo ha insinuado. El aparente enriquecimiento teórico de la economía postkeynesiana se ha llevado a cabo merced a los debates modelístico-estadísticos tendientes a negar o afirmar una hipótesis o un conjunto de hipótesis.

Este no ha sido el camino para la economía marxista; el criterio último de validez para ella es el que encierra el concepto de la práctica social, no la mera contrastación de hipótesis; y esto se debe no simplemente a que su espacio teórico carece de un aparato de medición y contrastación como el positivismo neoclásico sino porque el camino para analizar la realidad es diametralmente distinto.

En primera instancia porque las preguntas que se hace el maraxismo sobre la economía van más allá de las aparenciales relaciones entre variables, su búsqueda es la causalidad, mientras que la economía postkeynesiana tiene que efectuar verdaderos malabarismos teóricos para que el juego TEORIA --› MODELO --› MEDICION --› ESTIMACION --› PRUEBA DE HIPOTESIS --› TEORIA.... pueda alcanzar las metas causales que se plantea la economía marxista.

Buena parte de la gigantesca literatura existente sobre los debates entre postkeynesianos tienen esa mecánica del conocimiento. La economía marxista se ha interesado tradicionalmente en los cambios cualitativos y para esto no siempre la matemática ni la estadística pueden dar cuenta de este cambio substancial. En su esencia, la estadística y la matemática intentan dar respuesta a los cambios de magnitud y al juego de interrelaciones entre variables y son éstas precisamente las preguntas que se plantea el positivismo postkeynesiano. Y ésas son las preguntas porque ése es su interés: todos los planteamientos de la política económica tienen que ver con las transformaciones del régimen económico **dentro** del régimen económico, y , en éste, al nivel del proceso de circulación.

Para el marxismo los planteamientos de la teoría económica neoclásica pre y postkeynesiana formarían parte de ciertos niveles de abstracción como momentos horizontales dentro de todo el edificio teórico; no responden a las principales inquietudes causales, cualitativas, sino que se quedan en el aspecto cuantitativo al interior de un cierto nivel abstracto; no darían cuenta de ciertos cambios en el nivel de abstracción, o , para matizar, el sentido de los cambios de cualidad no se eleva hasta plantear la transformación de la sociedad. Cuando se requieren hacer comprobaciones sobre los cambios cuantitativos los problemas son muchísimo menores y bien pueden utilizarse técnicas estadísticas de contrastación con las limitaciones que hemos comentado.

A manera de ilustración de lo que se quiere afirmar, mencionemos la diferente forma de entender lo que conocemos como el proceso de la predicción. Esta se concibe de manera distinta, según el planteamiento teórico adoptado: podemos afirmar que con la teoría postkeynesiana la predicción es la magnitud que tendrá la variable en un futuro -y breve- período. Con la teoría marxista la predicción toma la forma de anticipación de los cambios de cualidad de los fenómenos. Recordemos que la teoría marxista descubrió el proceso de concentración y monopolización de la economía en la etapa de la libre competencia a través del análisis de las contradicciones que se generaban en ella misma.

Así las cosas, cabría preguntarnos: Dónde tendría la estadística campo de acción dentro de la concepción totalizadora marxista ? En principio, y al igual que otras escuelas, dentro de un mismo nivel de abstracción en el estudio del comportamiento cuantitativo de los fenómenos y en ese derrotero se podrían ampliar los análisis económicos de naturaleza marxista. Pero el salto de ciertas cualidades, el paso a niveles de abstracción más elevados no puede ser fácilmente resuelto por la matemática, por lo menos con la matemática actual.

Por estos motivos son injustas ciertas críticas a la teoría económica marxista que pretenden explicar su estancamiento a una suerte de descuido del análisis cuantitativo de la realidad circundante.[9] Se le critica precisamente por la cualidad que lo diferencia, se le critica su carencia de positivismo. No se puede asumir que la falla garrafal que explique un estancamiento del marxismo sea originado por esa desatención de los problemas cuantitativos. Lo que sí se puede aceptar es que las

----

9. Críticas como las señaladas pueden verse en los artículos de Loyola y Martín. Ver bibliografía.

investigaciones de corte marxista no han recogido y utilizado el importante arsenal matemático-estadístico. Desde luego, el estudio de interrelaciones cuantitativas, la organización de la información y, más en general, el uso de las modernas técnicas estadísticas y matemáticas permiten de una manera si se quiere, más sofisticada, vislumbrar derroteros en el proceso de conformación de nuevas categorías y conceptos, contribuyendo a la tarea de abstracción de la realidad. Simultáneamente, pueden desarrollarse análisis matemático-estadísticos que estén formulados para dar respuesta a cuestiones teóricas que estén engendradas desde la economía. Al menos éste parece ser el ideal.

La diferencia fundamental que queremos dejar en claro es entre aquellos que apelan a la matemática y estadística como **EL METODO** de aprehensión y descubrimiento de la realidad y aquellos que hacen uso de la matemática y estadística dentro de una concepción metodológica propia diferente, mediante la cual pueden apropiarse de la lógica estadístico-matemática en ciertas esferas de análisis en las que previamente la teoría las ha ubicado.

Adicionalmente, para estar al corriente de los debates teóricos entre postkeynesianos es obligado aprender a manejar el instrumental matemático-estadístico. Además, según las inquietudes investigativas, según la pregunta o preguntas que nos formulemos, podemos considerar la utilidad de una cierta técnica. Esto obliga, por supuesto, a tener un panorama teórico de la estadística que permita discernir sobre el procedimiento a utilizar, sus supuestos, sus objetivos, sus ventajas y sus limitaciones.

1.2 CORRELACION Y CAUSALIDAD.

En el campo de la investigación estadística [10] en muchas ocasiones y a pesar de las advertencias, se utiliza la correlación como explicación causal. Nada más lejano de esta afirmación. La correlación simplemente expresa la variación de alguna medida con respecto a la variación en otra. Si se constata que los movimientos de los dos fenómenos que estamos asociando tienen la misma dirección, entonces consideramos que la correlación es positiva; si la dirección es contraria, afirmamos que la correlación es negativa. El coeficiente de Pearson -el **R** ya visto en la sección anterior- es la medida generalizada de la correlación. Varía entre +1.0 y -1.0 evidenciando perfecta correlación positiva o negativa, según el caso.

Veamos su significado en forma más intuitiva: si en un par de ejes colocamos dos variables: la tasa de inflación (**P**) en el eje de las ordenadas y la tasa de crecimiento en la oferta monetaria (**M**) en el eje de las abscisas para diferentes períodos de tiempo, la representación visual de tal sistema podría tomar la forma de alguna de las gráficas que se ven en la página siguiente.

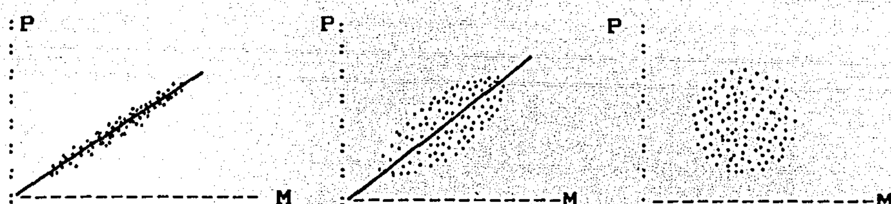
Cada punto es una pareja de valores de **P** y de **M** en un momento dado. La nube de puntos nos representa el conjunto de parejas obtenido de la observación en un buen período de tiempo.

----

10. Algunas de las ideas de este apartado se han inspirado en el interesante libro de S.J. Gould," The Mismeasure of Man". Cfr. Bibliografía.

En términos amplios podemos considerar el coeficiente de correlación **R** como una medida de la figura generada por la aglomeración de las parejas de puntos. Si la figura encontrada es una elipse angosta o delgada, decimos que la correlación es alta. La máxima correlación se obtendría con una elipse tan pero tan delgada que se vería como una línea recta (R = 1.0). En cambio a medida que la elipse se va 'engordando', la correlación tiende a disminuir. La más gordita posible corresponderá a un círculo con una correlación nula (R = 0.0).[11]

Gráfica No. 1.1



Una implicación inmediata de este análisis es el que se puede predecir cómo se comporta una variable ante el movimiento de la otra si la correlación entre ellas es alta. Ya habíamos mencionado que la correlación no es más que la razón de la variabilidad conjunta entre dos variables con respecto al

----

11. El R de Pearson mide la intensidad de la asociación **lineal** entre dos variables. Trata de cuantificar qué tan fuerte es la tendencia en que la nube de parejas de puntos pueda representarse como una línea recta. Relaciones no lineales de variables cuya representación de puntos arroja formas curvas -hipérbola, parábola- no mostrarían un **R** particularmente alto.

producto de sus desviaciones típicas. Dada la estructura propia de la definición de la correlación, en ocasiones se pretende encontrar causalidad entre variables cuando éstas están altamente correlacionadas. Dentro de lo que hemos visto hasta el momento no hemos establecido ningún tipo de relación causal entre la **P** y la **M** que estamos midiendo.

Hemos colocado a propósito a la oferta monetaria en el eje de las X, como se acostumbra a representar matemáticamente a una variable independiente. (El asignar a uno u otro eje tal o cual variable no cambia el sentido de la correlación). Parecería y sólo parecería que la variable **P** se mueve en función de la variable **M**.; más aún, si constatamos un coeficiente de correlación alto entre las dos, podríamos predecir -con una cierta probabilidad- cuánta sería la variación en los precios ante un cambio en la oferta monetaria. No obstante esto no significa relación de causalidad entre los fenómenos que estamos midiendo, ambos pueden obedecer a una estructura subyacente que les genera esa dinámica común.

Una versión más desarrollada de esta pretensión de querer encontrar causalidad se encuentra cuando la correlación no es significativamente alta en el período actual, pero sí lo es cuando se corre con alguna variable rezagada. Consideremos una variable en el tiempo corriente -por ejemplo la variación de precios en el período actual- en relación a los valores anteriores o rezagados de la oferta monetaria. De nuevo, si la correlación es alta, se arguye que la variable rezagada determina cambios en la otra variable: es como si el movimento de una originara los cambios en la otra.[12]

De nuevo ésta no sería afirmación de causalidad puesto que las dos variables pueden estarse moviendo correlacionadamente porque existe(n) una(s) tercera(s) variables que determinan su trayectoria, las cuales pueden estar o no rezagadas en el

----

12. Particularmente esto tiene que ver con el planteamiento de Friedman respecto a que los cambios seculares en el stock monetario real per cápita están altamente correlacionados con los cambios seculares en el ingreso real per cápita, planteamiento que expresado más sofisticadamente constituye puntal importante en sus formulaciones teóricas.

tiempo. A este tipo de correlación se le acostumbra llamar espuria, definición que en el ámbito de la interpretación abre las puertas a soterrados reclamos de causalidad, cuando se pierde ese carácter espurio. Pero, quiérase o no, la correlación no es sino la constatación del continuo movimiento del mundo real, sencillamente.

Habíamos mencionado que la predicción era posible si las variables tienen una fuerte correlación. No obstante, el acertar en las predicciones no significa que las variables guarden entre sí relaciones causales. El sentido común con facilidad puede conjeturar una conclusión de esa naturaleza. La historia de la astronomía nos ofrece un buen ejemplo de lo que queremos decir.

Desde tiempos remotos viajeros y navegantes utilizaban como guías para su orientación a las estrellas. Las épocas de siembra y de lluvias podían predecirse gracias a las fases de la luna y el recorrido del sol. En la Antigua Mesopotamia se trazaba la trayectoria solar y era posible predecir los eclipses de luna. En Grecia con Hiparco -centuria y media antes de cristo- la concepción celeste consideraba a la tierra como fija con planetas, luna y sol girando alrededor en una variedad de curvas.

Con el sistema de Hiparco, perfeccionado luego por Ptolomeo, era también posible la predicción, como acontecía con los mesopotámicos. Pero eso no significaba que esta gran y compleja cosmovisión fuera correcta: sufría de garrafales fallas de concepción originadas en su misma base pero que les permitió a griegos y mesopotámicos atinar en la predicción. Esta última, por tanto, no garantiza el establecimiento de relaciones de causalidad.

En la situación de correlación que nos ocupa, **M** y **P** pueden expresar sólo parte de una compleja estructura, cuyos fundamentos no pueden agotarse a través de la verificación de una muy buena correlación entre ellas. La única situación donde correlación y causalidad pueden identificarse es cuando la primera es el reflejo de una afirmación teórica ya discutida, esclarecida y establecida. En otras palabras, dada ya una teoría que establezca relaciones jerarquizadas entre conceptos, es posible a través de la cuantificación de estas conceptualizaciones que la correlación no sea espuria. Pero el proceso inverso no tiene la misma connotación. No se trata, por consiguiente, de descubrir la teoría a partir de la estadística, se trata por lo pronto, de ilustrar y medir ciertas relaciones teóricas por medio de técnicas estadísticas.

En la vida cotidiana encontramos una gran variedad de correlaciones entre distintos fenómenos: si correlacionamos la edad de cualquier mexicano con el precio de la gasolina o con la devaluación de un tiempo para acá encontraremos coeficientes **R** bastante altos. Nadie diría que hay asociación causal entre estos fenómenos. En general obtendremos correlaciones altas de variables que hayan estado creciendo y/o decreciendo, y esto no es sino una manifestación de la dinámica del mundo real, pero que como tal no nos permite elaborar teorías a partir de su simple constatación. Lo único que podemos aseverar con una correlación alta es que las variables están linealmente asociadas, pero la naturaleza de esa asociación es un problema para la teoría, forma parte del proceso de explicación global del conjunto de fenómenos al que pertenecen y en cuya jerarquía se insertan. El juego de categorías y conceptos debe dar cuenta de estas variables y sus asociaciones y no al contrario.

1.2.1 CORRELACION MULTIDIMENSIONAL Y LA REDUCCION DE LA DIMENSIONALIDAD

En dos dimensiones no se presentan dificultades especiales para entender intuitivamente la correlación. Hagamos constar, simplemente que de dos dimensiones con que empezamos (**P** y **M**) las hemos simplificado a una sola: la línea recta nos representa la información suministrada, siempre que la correlación entre las dos entidades sea relativamente alta. Ahora bien, si añadimos una dimensión más a nuestro ejemplo inicial -podría ser la evolución del tipo de cambio (**E**)-, para los mismos lapsos de tiempo considerados, podemos representar la correlación por una matriz simétrica de coeficientes de correlación en la cual la diagonal principal nos representa la correlación de cada variable consigo misma -$R(i,i) = 1.0$- y fuera de la diagonal las correlaciones de las variables entre sí:

M A T R I Z DE CORRELACIONES

	M	P	E
M	1.0	R(m,p)	R(m,e)
P	R(p,m)	1.0	R(p,e)
E	R(e,m)	R(e,p)	1.0

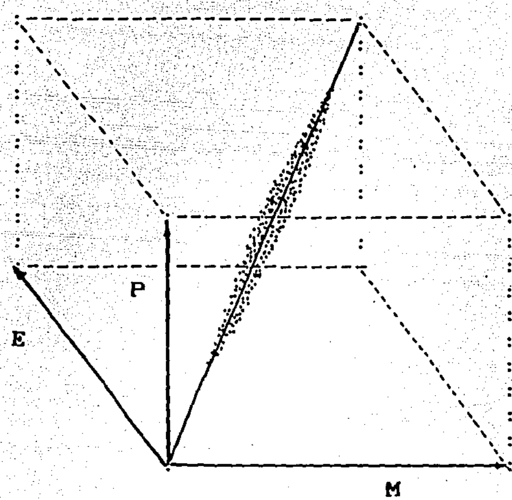
Donde:

M : Tasa de crecimiento de la oferta monetaria.
P : Tasa de Inflación.
E : Variación en el tipo de cambio.

Otra forma de representar la correlación es utilizando una gráfica que represente tres dimensiones -una para cada variable-, en la que se ubiquen puntos 'tridimensionales'. Si suponemos correlaciones altas y positivas entre estas variables, la nube de puntos aparece organizada como un elipsoide y la línea que lo recorre a través de sus ejes nos representa esa fuerte correlación positiva que asumimos inicialmente. (Ver Gráfica No. 1.2)

De pasada se menciona que hemos expresado las variables originales en una línea recta, hemos logrado reducir dimensionalidad sin gran pérdida de la información inicial. Podemos aumentar el número de variables (10, 20, 100...) y expresarlas en una matriz o bien 'imaginar' una representación de ellas en un espacio multidimensional en el que aparezcan tantos ejes perpendiculares entre sí como variables queramos incluir...

Gráfica No. 1.2



Si bien la representación matemática de este sistema es perfectamente factible, gráficamente no podemos visualizar más de tres dimensiones. La simplificación que hagamos puede chocar con el sentido e interpretación teóricos atribuídos a una combinación particular de variables -como la nueva línea recta obtenida- que ha permitido disminuir dimensionalidad.

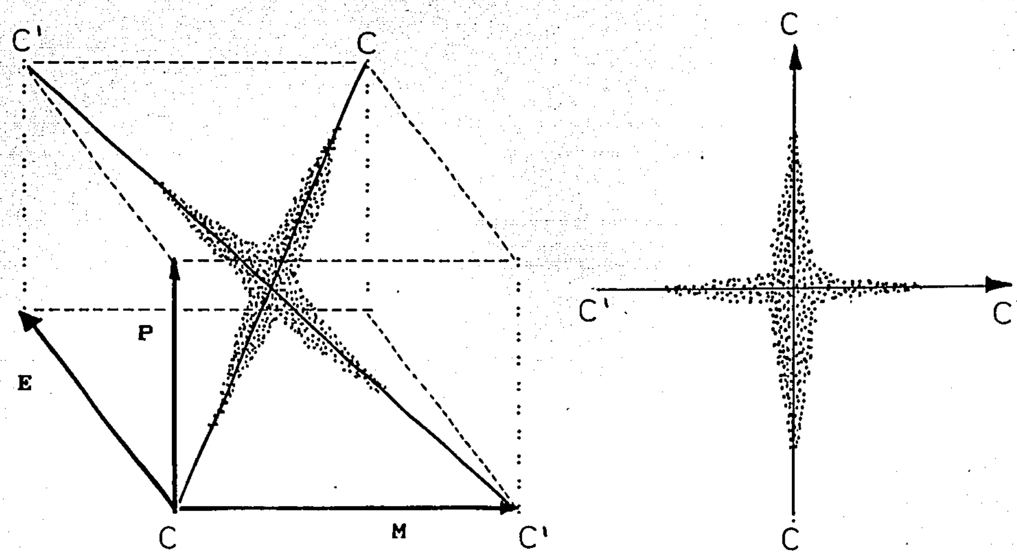
La pérdida de información de las variables originales puede ser compensada por la simplificación a que dio lugar, esto por lo menos en un sentido estadístico. Desde el punto de vista de la economía lo que interesa es si esta simplificación tiene sentido teórico; si es así, la responsabilidad no es de la técnica estadística como tal, sino del enfoque categorial teórico que orienta la investigación económica. En nuestro caso, podríamos aventurarnos a afirmar que una simplificación de las tres variables podría representar una nueva variable de índole monetaria: sería una combinación especial de tres fenómenos económicos de corte monetario: precios, oferta monetaria y tipo de cambio. Desde luego la interpretación de una combinación como la anterior no es fácil, y la más de las veces, es desafortunada.

El proceso de reducir la dimensionalidad, común a todos los métodos multivariados, puede ejemplificarse con ayuda de la Gráfica No. 1.2 : un eje nuevo que pasa a lo largo de la nubecilla de puntos del elipsoide. El significado geométrico para casos multidimensionales difícilmente se puede resolver con un solo y nuevo eje que atraviese el hiperelipsoide; de todos modos al obtener un primer eje consideramos -porque se construye de cierta manera particular- que en su trayectoria recoge la mayor variabilidad posible sobre cualquier otro eje análogamente construido. A este primer eje lo llamamos primer componente principal.

Se puede representar un segundo eje -o una segunda nueva dimensión o segundo componente- por una perpendicular al primero con la propiedad de retomar la mayor parte de la variabilidad que no fue absorbida por el primero sobre cualquier otra de las perpendiculares que pueden llegar a este primer eje. La gráfica tres nos ayudará a aclarar lo que queremos decir.

Las líneas CC y **C'C'** del lado izquierdo de la Gráfica No. 1.3 son perpendiculares entre sí: la recta **CC** es perpendicular a la recta **C'C'**. En términos vectoriales estaríamo hablando de ortogonalidad. Sobre **CC** puede pasar un infinito número de perpendiculares y se selecciona aquélla que recoge la mayor variabilidad dejada por el primer componente (**CC**). Si proyectamos el espacio tridimensional a dos dimensiones -o si aplastásemos sobre un plano la nube de puntos- encontraríamos algo como lo representado en la parte derecha de la gráfica. El primer componente recorre por el medio y de arriba a abajo el plano de puntos. El segundo componente también pasa por la mitad pero de lado a lado.

Gráfica No. 1.3



En **n** dimensiones pueden presentarse subsecuentes perpendiculares a los ejes previamente encontrados; cada nuevo componente obtenido recoge la mayor parte de la variabilidad no absorbida por los anteriores. Cuántos nuevos ejes tendríamos que buscar ? Si con -por nombrar un número- 6 ejes acumulamos la mayor parte de la variabilidad total de las dimensiones originales (80 a 90 %), hemos logrado una supersimplificación con una pérdida relativamente pequeña de la información original.

Concluimos: partimos de las variables originales y ajustamos nuevos ejes de tal forma que el primero de ellos explique más información que cualquier otro eje que pueda pasar por el hiperenjambre de puntos. Cada eje subsecuentemente encontrado además de explicar el máximo de la variabilidad restante de la información, debe ser perpendicular a los anteriores, de esta manera cada nuevo componente va explicando una parte proporcionalmente menor de la información.

Si la explicación anterior ha dejado lagunas, veamos si éstas se resuelven presentando el proceso en términos de vectores. Las variables iniciales pueden expresarse por vectores que tienen un punto común de partida.[13] La cercanía de los vectores puede medirse por el ángulo formado entre ellos y el coseno de ese ángulo nos mide la correlación entre los vectores.

Por ejemplo, la correlación entre un vector consigo mismo sería igual a la unidad, porque el ángulo formado sería de cero grados y el coseno de cero grados es igual a uno. Si dos vectores son perpendiculares su correlación será nula -y se dice que son independientes-, el coseno de 90 grados es igual a cero. Si los vectores parten del mismo punto pero en direcciones opuestas, su correlación será perfecta pero negativa -el $\cos(180^\circ) = -1.0$-.

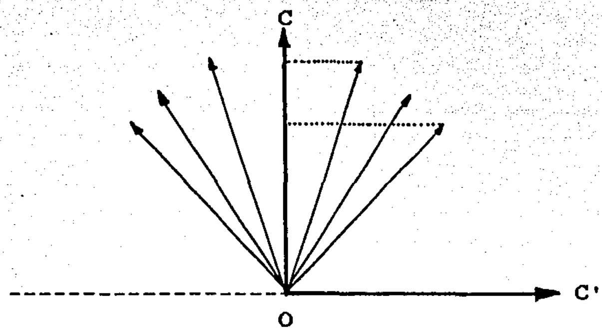
Conclusión: entre más 'cercanos' estén los vectores, midiendo su cercanía por el ángulo generado entre ellos, mayor será su correlación.

Con el fin de simplificar la explicación supondremos que tenemos una matriz de correlaciones en la cual éstas sean altas y positivas para que su representación gráfica sea un haz de vectores separados entre sí por su correlación. (Ver Gráfica No. 1.4).

----

13. En este caso por vectores de longitud unitaria, es decir los escalamos de una forma tal que su longitud sea igual a la unidad.

Gráfica No. 1.4



Nuestro primer componente corresponde al eje **OC** en la Gráfica No. 1.4, el cual puede verse como una combinación especial entre todos los vectores -por ejemplo un promedio de ellos- y que tiene la propiedad de resumir la información de todos los demás vectores. Qué tanto es la información que se resume ? Esto depende de la razón entre la longitud de la proyección del vector sobre la longitud del mismo vector.[14]

Si el vector está sobre el eje mismo, entonces este último absorberá toda la información. En tanto el vector se aleje del eje -en el sentido de que rota hacia algún lado teniendo el mismo origen-, menor será la proporción de información que puede 'guardar' el eje. Cuando el vector forme 90 grados con el eje -de nuevo, si hay independencia- la proyección será sólo un punto de modo que el eje es incapaz de conservar cualquier información sobre el vector.

En la Gráfica No. 1.4 el eje **OC** agrupa más información que cualquier otro que se pueda trazar. Buscando un segundo eje perpendicular al primero -en este caso **OC'**- se observa gráficamente que sólo da cuenta de una mínima parte de la información no contenida en el primer componente. Subsecuentes ejes nada tendrían que hacer al respecto. Es así como

----

14. La distancia del origen a las líneas punteadas que perpendicularmente llegan al eje son la mencionada proyección.

gráficamente se ilustra el número de ejes que deben mantenerse o, lo que es lo mismo, hasta dónde se considera que la información original ha sido rescatada en forma simplificada por varios vectores que son combinaciones de los originales.

En el libro ya mencionado de Gould se presenta el resultado de una investigación por él realizada en el campo de la biología:

> "Los sistemas de alta correlación positiva son encontrados frecuentemente en la naturaleza. En mi primer estudio de análisis de factores, por ejemplo, consideré catorce mediciones de huesos de 22 especies de reptiles pelicosaurios...Mi primer componente principal explicaba el 97.1 % de la información de todos los catorce vectores, dejando solamente un 2.9 % para ejes subsecuentes. Mis catorce vectores formaban un racimo extremadamente hermético -prácticamente traslapándose-; el primer eje pasaba a través de la mitad de este racimo. El rango de la longitud del cuerpo de mis pelicosaurios iba de un poco menos de dos a más de once pies... Todos los coeficientes de correlación entre los huesos son muy altos, de hecho, el más bajo de ellos es aún un sorprendente 0.912. Pero muy poco sorprendente, pues grandes animales tienen huesos largos y animales pequeños, huesos pequeños. Puedo interpretar mi primer componente principal como una abstracción del factor tamaño, reduciendo de esta manera -con mínima pérdida de información- mis catorce mediciones originales en una sola dimensión interpretada como el crecimiento del tamaño del cuerpo. En este caso el análisis ha logrado tanto la **simplificación** por la reducción de las dimensiones -de catorce a efectivamente una-, como la **explicación** por una razonable interpretación biológica del primer eje como el factor tamaño." [15]

A pesar de estos interesantes resultados el mismo autor llama la atención sobre los peligros que conllevan los problemas de agrupamiento de nuevos ejes y la interpretación teórica que los asiste. Volveremos sobre esto en capítulo posterior.

----

15. Gould. p. 249, 250. Enfatizado en el original.

1.3 NOTAS SOBRE LA METODOLOGIA.

1.3.1 LAS FORMAS DE PENSAMIENTO.

En esta sección nos interesa elaborar una especial síntesis de las corrientes metodológicas de una manera tal que se acomode al interés de la presente investigación. No se trata de efectuar un resumen sobre el nivel actual de la discusión sino de organizar y ordenar ciertas notas con la mira puesta en los problemas que se quieren abordar y que de alguna manera se esbozaron en las secciones anteriores. La intención es una reconstrucción personal de los principales planteamientos que nos sean de utilidad para entender la lógica de la contrastación y comprobación de teorías. Esto mismo nos servirá como orientación general sobre el papel de la estadística.

Como punto de partida podemos considerar que la metodología se refiere a la forma en la que las teorías son formuladas, a la manera en la cual el conocimiento se genera bajo condiciones de incertidumbre. Los problemas metodológicos están enlazados a los procedimientos por los cuales los economistas generan sus teorías, a la formulación de planteamientos acerca de cómo la propia teoría debe gestarse y a los criterios de comparación y aproximación de teorías. Se considera la metodología como las etapas del estudio de la formulación teórica al nivel de la construcción de modelos y a nivel de la subyacente y a menudo implícita visión del mundo. Dichos niveles son interdependientes, una particular visión del mundo está generalmente asociada al propio esbozo teórico de un sujeto. Se trata de estudiar a la metodología como algo que trasciende todo contenido teórico y que provee un criterio universal de aproximación que se puede utilizar como instrumento de clasificación de los cuerpos teóricos.

Para hacer un bosquejo sobre la cuestión metodológica asumiremos el enfoque desarrollado por algunos autores [16] el cual va a ser especialmente útil para los propósitos que nos interesan. Evidentemente pueden existir -y en general de hecho los hay- mejores percepciones del asunto, pero por las características que tiene nuestro trabajo y dados los aspectos que queremos subrayar es suficiente con el enfoque mencionado.

Se parte de considerar dos tipos o formas fundamentales de pensamiento, intentando subrayar tanto las aproximaciones metodológicas tradicionales como las modernas que hayan tenido eco en la teoría económica. Tales tipos corresponden a lo que llamaremos el método axiomático y el método babilónico.

Por formas de pensamiento estamos expresando, sencillamente, la manera en la cual se establece y presenta la argumentación, la forma por la cual se intenta convencer a otros de la validez o verdad de esa argumentación. Estos dos modos de pensamiento encierran dos diferentes vías de construir argumentos y de desarrollar teorías. Como tal no son todas la posibles formas de pensamiento. La agrupación que se pretende hacer simplemente subraya dos instancias que consideramos de utilidad para precisar algunas características que aceptamos como las más importantes dados los objetivos de nuestro estudio.

----

16. Dow, Feynman, Wimsatt, Stohs, Macfie. Cfr. bibliografía.

1.3.1.1 EL METODO AXIOMATICO.

A grandes rasgos podría considerarse el método axiomático como uno de los procedimientos mas acostumbrados en el terreno económico. Su atractivo inmediato deviene por virtud de permitir la construcción de un sistema lógico completo, lo cual siempre aparece como muy deseable. Dentro de este enfoque la matemática se concibe como el ápice de la pureza científica. Algunos autores [17] emplean el término "método Euclidiano/Cartesiano" (E/C) para referirse a todo el pensamiento científico influenciado por el ideal de sistemas cerrados de lógica axiomática; la mención de Euclides y Descartes se hace por referencia al propio método matemático. Por ser de popular aceptación dentro de ciertas corrientes de la economía -a veces no explícitamente- lo retomaremos en este trabajo para referirnos a aquellos que lo consideran un ideal dentro de la metodología tradicional.

Como un ejemplo, solamente, en economía los axiomas de la racionalidad del consumidor permiten la construcción de una gran cantidad de teoremas gracias a la lógica deductiva, aunque la axiomatización en este caso tiene que ver más con los procesos de introspección que con la constatación inicial de la realidad empírica.

La aplicación de esta aproximación sistémico-axiomática a la generación del conocimiento ha fomentado dentro de las diferentes disciplinas algunos rasgos característicos. Ellos son la tendencia al reduccionismo y la presencia constante del dualismo.

----

17. Sheyla C. Dow intenta sintetizar los troncos metodológicos más comunes en la economía a partir de un estudio de las aproximaciones metodológicas en la teoría macroeconómica. Cfr.bibliografía.

En cuanto a la primera se justifica dentro del propio enfoque por cuanto la estructura lógica completa va a depender de los axiomas básicos, por esta razón su aceptabilidad debe ser lo más amplia posible. De esta manera se obtiene como resultado que las proposiciones sean subdivididas en sus componentes más pequeños de una manera tal que todas puedan derivarse por medio de la deducción. En el terreno de la teoria económica los capitulos sobre la conducta del consumidor individual es de los mejores ejemplos que se puedan encontrar al respecto.

El segundo rasgo característico es tal vez el más importante. El dualismo es la pretensión de clasificar todos los conceptos, proposiciones y eventos en forma dual como perteneciente a una de dos categorías: falso o verdadero, lógico o ilógico, hecho u opinión, positivo o normativo, científico o no científico, espíritu o materia... Adelantándonos un poco, debemos mencionar que la estrechez de esta visión lo ha emparentado con modos metafísicos de pensamiento que son incapaces de concebir la riqueza contradictoria y compleja del mundo real, motivo por el cual el dualismo se enfrenta a serias dificultades para trasladar la lógica de lo real a la lógica del pensamiento.

En términos amplios podemos considerar que los planteamientos de los teóricos monetaristas, de los teóricos del equilibrio y en general de los neoclásicos de estirpe similar están estrechamente ligados a este tipo de argumentación.

1.3.1.2 OTRO MODO DE PENSAMIENTO.

El segundo modo de pensamiento que llamaremos babilónico, para continuar con la lógica de los autores ya mencionados, no constituye como tal un ideal de la ciencia dentro de la filosofía occidental, y su origen, desarrollo y características no están definidos con claridad.

Más que usar un sistema lineal de deducción lógica tomando como base axiomas básicos, este modo de pensamiento comienza con la convicción de que es imposible, en general, establecer axiomas y puntos nodales inconmovibles y precisos (estancos) en el camino en el cual el error axiomático está compuesto por cada eslabón en la cadena axiomática de la lógica. La aproximación alternativa es el emplear varias líneas de argumentación las cuales tienen diferentes puntos de partida que dentro de una teoría exitosa sirven de refuerzo mutuo. Aquí no hay una única fuente de verdad, el proceso de comprensión de la realidad escapa a una exclusiva forma de percepción.

Cualquier argumento no se erige o se derrumba sobre la base de la aceptabilidad de un grupo concreto de axiomas. El conocimiento se genera por aplicaciones prácticas de teorías, usando una variedad de métodos y no jugándose la verdad por la defensa y apología de uno en particular. A este conjunto inicial de planteamientos es lo que se ha dado en llamar método babilónico de pensamiento.

Para el sistema babilónico es básico la inexistencia de axiomas como punto de partida, pero algunos teoremas pueden ser axiomas en otras partes del sistema. Si el mismo teorema sale como resultado de la aplicación de la lógica a los axiomas derivados de diferentes partes del sistema, entonces la aceptabilidad de tal sistema se incrementa porque no depende para su validación de cualquier grupo particular de axiomas. Si una conjunción de eventos persiste a pesar de un cambio en la estructura económica entonces la probabilidad de una relación causal se incrementa.

En este enfoque pueden incluirse la corriente postkeynesiana (en su sentido estricto), las corrientes marxistas, historiadores económicos y demás teóricos de planteamiento similar.

1.3.2 CORRIENTES METODOLOGICAS TRADICIONALES.

Para otear las corrientes metodológicas tradicionales tendríamos que referirnos al debate sobre los méritos relativos de la deducción y la inducción.

En la tradición E/C la deducción es el método clásico por excelencia: se trata de aplicar la lógica a alguna ley general o axioma, lo cual, conjugado con algunas condiciones iniciales, nos lleva a derivar teoremas particulares. El punto de partida de la inducción, en cambio, es la observación de la conjunción particular de eventos los cuales son susceptibles de conectarse causalmente -por ejemplo, expresando estas conexiones como teoremas- caso en el cual el mecanismo lógico operararía hacia atrás para obtener verdades o axiomas más generales. El paso crucial e inicial en la inducción es la formulación de hipótesis que intentan reflejar relaciones de causalidad.

El problema de la deducción se deriva de su dependencia de la validez de las leyes generales o axiomas de los cuales el teorema se dedujo. La validez de la lógica deductiva por sí misma no inmuniza al argumento de la necesidad de demostrar su validez empírica. De otro modo, el problema de la inducción es que el conjunto de eventos asociados puede no reflejar causalidad de tal manera que de los diferentes grupos de observaciones podrían seguirse diferentes conclusiones. El ejemplo típico del problema de la inducción es la afirmación de que todos los cisnes son blancos a partir de la observación de que algunos de ellos tienen ese color.

En el proceso de aplicación a campos científicos determinados, dentro del modo de pensamiento E/C, la inducción y la deducción generalmente se combinan para solventar las limitaciones de cada uno de ellos: se realizan las observaciones con un cierto sistema de lógica deductiva para que el conjunto observado de eventos se explique en términos de ese sistema, obteniendo las conclusiones a través de la inducción. No obstante, dentro del campo de la filosofía de la ciencia, la gran carga de dualismo ha generado una gran tendencia por optar entre estos dos métodos, considerando a la alternativa seleccionada como el ideal del método de aproximación científica. La historia de la filosofía de la ciencia refleja la lucha entre los "deductivistas" y los "inductivistas". La discrepancias, en cambio, son menos agudas en los campos aplicados, es algo así como que en la práctica de la investigación nos tenemos que convertir en babilónicos, pero espiritualmente debemos sustentar una posición ortodoxamente axiomática.

Estos métodos también tienen que ver con la clasificación de las proposiciones entre analíticas y sintéticas. Para las proposiciones analíticas el criterio de verdad está asociado a la estructura lógica de la definición de sus términos. Todos los cisnes blancos son blancos es una proposición de naturaleza analítica, igualmente lo sería la ecuación de la teoría cuantitativa del dinero. Para las proposiciones sintéticas el criterio de verdad depende de la consistencia de la proposición no en sí misma, sino con los hechos observados. De esta manera, las proposiciones significativas pueden ser analíticas o sintéticas y capaces de probarse contra los hechos, a través de la verificación. El positivismo lógico con su criterio de lo que puede ser significativo no es más que la extrapolación del modo de pensamiento E/C a campos distintos del matemático.

Es a partir de lo anterior que se desarrolla el "principio de la verificabilidad" o "principio de la verificación" dentro del positivismo lógico como un criterio para diferenciar la ciencia de la no ciencia a través de las llamadas proposiciones significativas. Estas son proposiciones generadas por medio del método hipotético-deductivo: se construyen hipótesis a partir de las observaciones para derivar teoremas que serán probados contra posteriores observaciones.

Al interior de la filosofía de la ciencia, si la verificación es el criterio para aceptar teorías, su ausencia sugiere el rechazar las teorías. Si las teorías tienen la forma simple de 'si A => B' la verificación es un problema también muy simple. No obstante, las teorías son mucho más complejas que lo que enuncia la proposición anterior o, puede ser también, que A (o B) encierre un contenido mucho más denso en términos de hipótesis o teoremas. En esta situación la comprobación deja lugar a que no podamos discernir si el mecanismo de verificación afecta o comprueba una sola proposición de todo el conjunto de proposiciones, a algunas de ellas, o a todo el grupo. Análogamente igual podria decirse en el caso de que rechazáramos dicha verificación.

1.3.3 CORRIENTES METODOLOGICAS ACTUALES.

1.3.3.1 LA "FALSACION" DE POPPER.

La capacidad de probar hipótesis fue más tarde cuestionada por el problema lógico generado por la inducción, por lo tanto la verificación no podía, en este nuevo terreno, ser la garantía de las relaciones de causalidad. Si la teoría es de la forma 'si A => B' y si B era verdadero, no necesariamente A tiene también que serlo; podría existir un tercer evento C verdadero y por lo tanto 'si A => C' sería el caso. Este cuestionamiento del positivismo lógico lo intentó rescatar uno de sus principales críticos, Popper, quien lideró los nuevos desarrollos en la metodología.

Considerando que la aproximación convencional prevaleciente estaba adoptando una serie de preceptos para inferir proposiciones extraídas a partir de la inducción en la forma más satisfactoria posible dado que se reconoce el problema de la inducción, Popper pensó probar una solución alternativa la cual evitaba el problema.

Popper arguyó que la prueba empírica no implicaba verificación, pero sí falsación. Mientras las observaciones son tomadas como aseveraciones verdaderas, el único conocimiento real que nosotros tenemos es que ciertas teorías no son verdaderas. Si encontramos que B no es verdadera en la proposición obtenida por lógica deductiva 'si A => B', concluimos que A tampoco es verdadera.

La anterior es la forma por la cual el criterio de Popper para identificar las proposiciones científicas es el criterio de la falsación. Establece una serie de reglas procedimentales que permiten construir teorías que, aunque no son verificadas, por lo menos no son falseadas por el mecanismo de la comprobación.

En este marco se considera que los teoremas deben ser lanzados sobre la base de una conjetura audaz, antes que sobre la base de la observación; las reglas normativas restantes se concentran en evitar estrategias defensivas para que el desarrollo de las teorías por parte de sus creadores no intente asegurarse resultados no falseables. El desarrollo de la ciencia lo explica Popper -en tanto el objetivo de las teorías son la explicación y la predicción- como el proceso por el cual una falsación puede provocar una modificación de la teoría al incorporar la nueva evidencia, de esta manera el progreso de la ciencia está marcado por el conflicto entre las viejas y las nuevas teorías. Si el mecanismo de prueba arroja una ausencia de falsación, esto podría ser pensado como una corroboración -no una comprobación- de la validez de una teoría. Dentro de este marco el único conocimiento verdadero que se tiene es que ciertas teorías **no** son verdaderas. Teorías no falseadas no pueden considerarse verdaderas -sólo corroboradas- porque puede aparecer posteriormente un hecho empírico que las falsee.

Dentro de la economía los criterios anteriores para el comportamiento científico pueden tener una aparente validez surgida del hecho de que rara vez se disponen de observaciones bajo circunstancias controladas. Algunos rasgos de la economía son susceptibles de cambios imprevistos de una forma tal que el rango de posibilidades es difícil de anticipar antes de que una teoría sea probada. No obstante, variedades del principio falsacional de Popper son ampliamente usados en economía como método preferido de prueba. [18]

Pero el proceso de falsación no es fácil. La tesis de Duhem y Quine, la cual cuestiona la facultad de los procedimientos falsacionales para trabajar con un bloque complejo de teoría, plantea que sólo los teoremas más simples son posibles de

----

18. Al respecto, las llamadas pruebas de hipótesis en estadística tienen rasgos muy parecidos con el principio falsacional de Popper. Remitimos al lector a la sección 1.1.2 para hacer una comparación con la visión de este autor.

verificar, por cuanto planteamientos más complejos no tienen la facultad de poderse comprobar de una manera conjunta; siempre se están contrastando hipótesis en conjunción de una serie de condiciones auxiliares, por tanto, es imposible de precisar donde está el **quid** de la refutación. Esto ha sido motivo lo suficientemente fuerte para limitar aplicaciones posteriores de los principios de Popper a la economía:

> "Popper ofrece una guia inadecuada más allá del reino de las proposiciones falsacionales. En cambio, como un resultado, la práctica económica ha divergido marcadamente de lo que se **profesa** a lo que se prefiere metodológicamente."[19]

En Popper es clara una demarcación dualista entre lo que es y no es la ciencia; su filosofía de la ciencia es marcadamente normativa estableciendo un código de conducta de lo que es una buena práctica científica, no sólo en términos del mecanismo de prueba de una teoría, sino también en términos de la formulación de la teoría misma.

1.3.3.2 LOS "PARADIGMAS" DE KUHN.

Se considera que Popper heredó la visión "tradicionalista" dentro de la filosofía de la ciencia. Kuhn, al contrario, rompe con esta cadena de pensamiento. En su libro llamado LA ESTRUCTURA DE LAS REVOLUCIONES CIENTIFICAS, publicado en 1962, Kuhn cuestiona el "buen" recetario normativo de lo que debe ser una ciencia y en su lugar plantea que el criterio de la falsación no conduce en la práctica a rechazar una teoría. Considera que cualquier teoría individual forma parte de toda una estructura teórica y no puede plantearse como una entidad aparte susceptible de comprobación. Más aún, históricamente las estructuras teóricas se han derrumbado y su vez reemplazado por otras estructuras, por razones diferentes a la falsación.

----

19. Dow, p. 24. Enfasis en el original.

Kuhn no se interesa exclusivamente en la "racionalidad interna" de la actividad cientifica, explora el contexto histórico en el cual se obtienen los descubrimientos científicos y cuestiona la visión dualista como procedimiento adecuado para abordar el estudio del desarrollo científico. Para él, el proceso del descubrimiento científico es el cambio de estructuras teóricas y forma parte del medio ambiente histórico que va tomando lugar. Por estas razones se considera que Kuhn antes que un filósofo de la ciencia es un historiador de la ciencia: enfatiza la particularidad histórica del desarrollo científico y concluye que no es posible identificar una tendencia necesaria -y general- de pensamiento a través del tiempo sino sobre la base de una particular escuela de pensamiento.

El concepto central que Kuhn emplea es el concepto de "paradigma" o "matriz de disciplinas" y es lo suficientemente amplio para incluir todos los aspectos de una estructura teórica, recorriendo desde las técnicas prácticas de análisis hasta la visión del mundo y modo de pensamiento subyacentes de la comunidad científica que forma parte de una estructura teórica determinada y que los hace identificarse dentro de ésta. El modo común de pensamiento y la estructura teórica permiten la comunicación entre los miembros del grupo como un medio de avanzar en el descubrimiento científico dentro de un paradigma y esta actividad es lo que Kuhn llama la "ciencia normal", la cual involucra los nuevos descubrimientos que son compatibles con el paradigma dominante.

Para Kuhn la historia de la ciencia se caracteriza por períodos en los cuales un paradigma es dominante y su desarrollo se ve interrumpido cuando se cuestiona la base de su estructura teórica haciendo que el paradigma existente entre en crisis y pueda ser reemplazado por uno nuevo como resultado de una revolución científica. La crisis ocurre, no por la falsación de alguna parte del viejo paradigma, sino como un resultado de una percepción de que el viejo aparato es insuficiente para abordar y explicar un nuevo y crucial problema. El nuevo paradigma, por definición, es una estructura compleja formada sobre la base de una diferente visión del mundo respecto al viejo paradigma y está dotado de diferentes técnicas y lenguaje. No se cuenta con una base sobre la cual puedan ser comparados y no existen criterios analógicos que nos informen si ha habido algún progreso científico. El proceso de revolución no es discontinuo por el hecho de que las anomalías dentro del anterior paradigma deben haber sido el sujeto de comunicación en términos del lenguaje del nuevo paradigma, sino que la discontinuidad se refiere al hecho de que no existe un paradigma neutral con el cual comparar el paradigma anterior con el que lo reemplaza.

Los científicos que "comulgan" con una ciencia normal están de acuerdo tanto en los problemas que requieren solución como en la forma general que esta solución toma lugar; no obstante, sólo el juicio de los propios colegas se considera como el relevante en la definición de los problemas y en las soluciones; en consecuencia la ciencia normal se sostiene a sí misma mediante el proceso acumulativo de resolución de los interrogantes dentro del contexto de un marco analítico común. El rompimiento de una ciencia normal, cuando ello ocurre, es abanderado por una proliferación de teorías y la aparición de la controversia metodológica. El nuevo marco ofrece un camino alternativo a los problemas sin resolver abandonados por la ciencia normal en crisis y comienza a configurarse un nuevo paradigma que puede dar lugar a la ciencia normal imperante en el futuro.

A pesar de algunas críticas de las que ha sido objeto, vale la pena destacar que la teoría de Kuhn sugiere que la práctica científica ha destacado una multiplicidad de métodos científicos. Unos reemplazan a otros por razones diferentes a lo que tradicionalmente se conoce como racionalidad científica, y es imposible identificar objetivamente el cambio como progreso científico. Esto último le ha valido a Kuhn la crítica de relativista o, más aún, de nihilista.

1.3.3.3 LOS "PROGRAMAS" DE LAKATOS.

Continuando con nuestra línea de discusión es necesario mencionar los planteamientos de uno de los discípulos de Popper, Imre Lakatos.

Lakatos desarrolla como aspecto esencial de su planteamiento la noción de "Programas de Investigación" (P.I.) como la estructura teórica definida por dos elementos, el núcleo de proposiciones, irrebabtible por su naturaleza, y el elemento "heurístico":

> "La unidad básica de evaluación no debe ser una aislada teoría o conjunción de teorías, sino, antes bien, un '**programa de investigación**', con un '**duro núcleo**' convencionalmente aceptado (y así, por decisiones convencionales, '**irrefutable**') y con un '**heurístico positivo**' que define problemas, esboza la construcción de un cinturón de hipótesis auxiliares, prevé anomalías y las convierte victoriosamente en ejemplos, todo ello de acuerdo con un plan preconcebido."[20]

La actividad dentro de un P.I. es dirigida por su heurística positiva, entendiendo ésta cómo la agenda de problemas para ser resueltos y los métodos para resolverlos. Para el hombre de ciencia, "es básicamente el heurístico positivo de su programa, no las anomalías, el que dicta la elección de sus problemas."[21] En su opinión, la falsación en el sentido de Popper no tiene por qué implicar rechazo de una teoría, las falsaciones son anomalías que deben tenerse en cuenta y ubicarse dentro de la perspectiva de un P.I.: las anomalías por sí mismas no dan lugar a rechazar una teoría porque se puede "defender '**progresivamente**' cualquier teoría durante largo tiempo, aun si es falsa." [22].

El significado de los P.I. progresivos y regresivos los explica Popper de la siguiente manera:

----

20. Lakatos, p. 221. En bastardillas en el original.

21. O.C., p. 221.

22. Lakatos, p. 223. En bastardillas en el original.

> " Se dice que un programa de investigación está **progresando** mientras su desarrollo teórico se anticipa a su crecimiento empírico, es decir, mientras continúa prediciendo hechos nuevos con algún éxito; es regresivo o estancado si el desarrollo teórico se queda atrás de su desarrollo empírico, es decir, mientras sólo dé explicaciones post-doc, sea de los descubrimientos causales o de los hechos provistos por un programa rival."[23]

En esta situación no se trata de que una teoría reemplace a otra por el criterio de la falsación de Popper. Cuando un P.I. explica progresivamente más que un P.I. rival, lo supera y por tanto puede eliminarse el programa rival.

Para Lakatos, como para Popper, la metodología de la ciencia tiene que ver con la "lógica de la aproximación", esto es, con el problema normativo de proveer criterios de progreso científico. Se acerca más a la noción tradicionalista dentro de la discusión metodológica. Lakatos tiende hacia el planteamiento de Kuhn en el sentido de que puede persistir un programa de investigación a pesar de sus anomalías, pero se acerca más a Popper cuando especifica los criterios para rechazar un programa de investigación determinado. Sus planteamientos serían una versión más sofisticada del enfoque falsacional de Popper, intentan configurar un conjunto de criterios neutrales y universales los cuales trascienden los cuerpos individuales de investigación.

----

23. Lakatos, p. 223, 224. Enfasis en el original.

Apéndice A

COMENTARIOS SOBRE LOS DATOS PRIMARIOS.

Una de las tareas importantes a las que se enfrenta el economista en la actividad investigativa es el manejo de los datos primarios. La teoria por sí misma es incapaz de generar información cuantitativa y gran parte de la labor del investigador es dedicarse a la búsqueda de datos primarios. Como regla más común los proveedores de estos datos son los organismos estatales. En muy contados casos el investigador posee el control sobre alguna de las etapas de la obtención y presentación de la información estadística de estos organismos.

En la información manejada por las instituciones estatales se presentan problemas como son las cuestiones de cobertura y calidad, tiempo de recolección y elaboración, objetivos de la investigación, métodos y técnicas empleados, costos involucrados etc. No obstante el problema más agudo está en que los objetivos para los cuales se diseΣ$ó un proyecto de medición no tiene por qué responder a las necesidades de una investigación particular. A lo sumo esto se vincula a los objetivos de planeación y política económicas -en los casos más optimistas-, y a los objetivos de justificación de una determinada gestión -en la situación más pesimista-.

Al mismo tiempo, y como parte de este problema, los sistemas de contabilidad social tienen un soporte teórico determinado, son una contrapartida empírica a un enfoque de la economía. Sus fines y sus objetivos están definidos -y limitados, evidentemente- por el conjunto de hipótesis a las cuales responden.

Las apariencias de imparcialidad y universalidad con que se presentan los sistemas de cuentas nacionales, los modelos de insumo-producto y la balanza de pagos, no deben hacernos olvidar que responden a la formulación y organización de un marco general que va a condicionar los resultados. Danilo Astori hace el siguiente comentario al respecto:

> "En particular, el modelo mencionado (refiriéndose al Sistema de Cuentas Nacionales de las Naciones Unidas) ha sido construido para ser aplicado en economías con organización capitalista y se apoya en un conjunto de hipótesis que interpretan de determinada manera el funcionamiento de aquélla. O sea que no puede ser universal por dos razones esenciales: por referirse a una sola forma posible de organización económica y, adicionalmente, por estar sustentado en un solo tipo de percepción de dicha organización. En lo fundamental esta percepción es la que corresponde a las categorías de análisis de la economía keynesiana y, en lo que se refiere particularmente al problema central de la generación del valor, a la interpretación neoclásica que, como se sabe, es de naturaleza subjetiva."[24]

A pesar de que el economista comparta esta concepción teórica, no necesariamente los problemas de que se ocupa tienen su contraprestación en términos de mediciones económicas. Nuevas categorías de análisis económico no son cobijadas por los organismos encargados de la recolección de la información. Más aún, en el caso de presentar una cierta o cercana similitud, el tiempo de procesamiento y publicación de la información puede retrasar años la búsqueda de respaldo empírico a un planteamiento teórico.

Como alternativa está el intentar aproximar -y a veces, transformar para ello- los datos existentes a mediciones de naturaleza económica que adquieran sentido para los propósitos de la investigación. Recordemos que dentro de este marco el interés central es acopiar información, traducirla o acomodarla a un lenguaje económico y así dotar de una base empírica al problema

----

24. Astori, p. 82.

analizado. Por esta razón Kuznets plantea que hay que transformar los datos e información en mediciones económicas con las cuales se obtengan mediciones con un claro sentido en la conceptualización económica:

> "Se trata de convertir dicha materia prima (refiriéndose a la información bruta) en una perfeccionada medición económica."[25]

Y, en tanto el correlato empírico no puede ceder lugar a la conjetura teórica en el proceso de contrastación, el mismo autor privilegia la información existente, a pesar de su confiabilidad:

> "Por deplorablemente endeble que sea la contabilidad económica en general, en un momento determinado proporciona el único fundamento empírico para orientar un problema."[26]

Como ejemplo de lo que se está mencionando podría ser el conjunto de análisis que se propone estudiar la capacidad subutilizada en la industria. Los criterios que se establecen para definir la capacidad ociosa no están en consonancia con la información disponible al respecto. Se acude a medición indirecta a través del manejo de variables relacionadas.

Las verdaderas dificultades se crean cuando la investigación empírica se efectúa con un marco teórico diferente -por ejemplo con el del aparato clásico-. Dada la imposibilidad física de construir un sistema de contabilización de distinta naturaleza al institucionalmente establecido, se tiene que acudir al existente, lo cual exige una ardua tarea de transformación de los datos que en su origen responden a problemas distintos y que en ocasiones mistifican el asunto que se quiere investigar. La alternativa de aproximación, además de requerir gran ingenio y buen manejo conceptual-estadístico, la más de las veces difícilmente perderá ese carácter de aproximación. No obstante los criterios teóricos de selección y definición, así como la habilidad en la operatividad estadístico-conceptual, pueden concederle un cierto contenido específico teórico a la medición económica.

----

25. Kuznets. 1979, p. 32.

26. Kuznets. 1979, p. 31.

Graves problemas se presentan en el tratamiento cuantitativo, si queremos, por ejemplo, estudiar el nivel de ingresos de las clases sociales según la conceptualización clásica. El mayor inconveniente es el que la información con que se cuenta entorpece tanto la categorización en clases sociales, como su propia medición. En este caso el papel mistificador de la información nacida de otra teoría económica no es tan simple de romper. La labor de tamización no siempre resulta fácil.[27]

Por último, se debe mencionar que la calidad de la información disponible es función -términos más, términos menos- del nivel de desarrollo del país de que se trate. En situaciones de bajo nivel de desarrollo, los sistemas de contabilización que tienen su fuente en censos, encuestas, muestreos..., por lo general adolecen de la calidad necesaria para ser considerados confiables.

----

27. Como ejemplo, véase al respecto el interesante trabajo de Anwar Shaikh.

Capitulo 2.

COMPONENTES Y FACTORES

En el capitulo anterior -sección 1.2- habiamos introducido el tema de los componentes principales. En aquella parte nuestro interés era explicar gráfica e intuitivamente el proceso de obtención de nuevos ejes que nos simplificaran un mayor conjunto de información. Se justificaba esa presentación porque las técnicas multivariadas van a tener en común ese método de simplificación. En lo que sigue se pretende desarrollar con una mayor amplitud los análisis de componentes principales y factores. Se incluyen en el mismo capitulo por la cercana familiaridad que tienen entre si esas dos técnicas.

Iniciamos la presentación con el análisis de componentes principales (C.P.) por un doble motivo: de las técnicas multivariadas es la más sencilla de explicar y por ende permite acostumbrar al lector con la notación matricial. [1] En la segunda sección del capitulo abordaremos el tema de los factores para luego comentar en la tercera sección sobre algunas aplicaciones que se han hecho en México en el campo socioeconómico. Terminamos con una parte dedicada a la notación y el tratamiento puramente matemáticos, exclusivamente de los temas que en este capitulo se tocan. Lo incluimos como sección especial para agilizar la presentación de los aspectos principales.

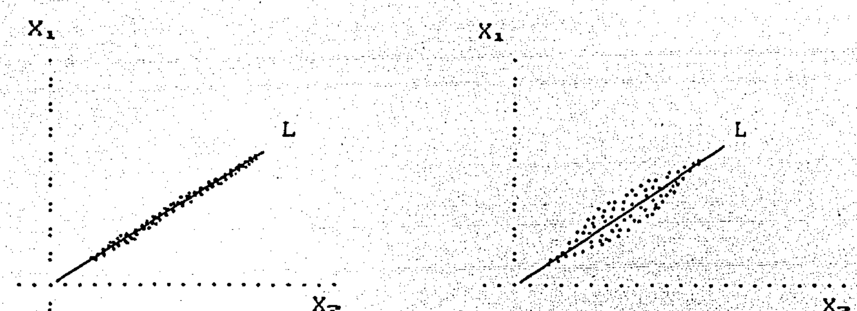
----

1. En este trabajo se incluye un apéndice con algunos elementos de álgebra matricial que tiene por objetivo presentar lo que a nuestro juicio se considera como mínimo indispensable para entender la matemática multivariada.

2.1 ANALISIS DE COMPONENTES PRINCIPALES.

Retomemos lo mencionado en la sección 1.2: si contamos con información a lo largo de un buen período de tiempo de dos variables económicas como la tasa de inflación $-X_1-$ y la tasa de devaluación $-X_2-$, la representación geométrica de las dos variables si están positiva y perfectamente correlacionadas sería una recta como la línea **L** en la izquierda de la gráfica 2.1.

GRAFICA 2.1



Esta línea puede ser descrita como una combinación lineal de las variables mencionadas:

$$y=a_1X_1+a_2X_2 \quad (2.1.1)$$

Esta última combinación lineal puede ser usada alternativamente como representación de la distribución de X_1 y X_2 en lugar de utilizar a las variables originales. Los coeficientes a_1 y a_2 son ponderaciones de ajuste.

Si la correlación positiva entre las variables no es tan perfecta -ver la parte derecha de la Gráfica 2.1 - la combinación lineal representada por **L** ofrece sólo una aproximación a la distribución de las dos variables en las que se captura una porción importante de la variabilidad conjunta entre ellas. Como es obvio, a medida que la asociación entre X_1 y X_2 se incrementa, el elipsoide tiende a concentrarse más cerca de la línea recta y se mejora el grado de aproximación de esta última a la distribución de las dos variables.

2.1.1 OBTENCION DE LOS COMPONENTES.

Ahora, en lugar de contar con sólo dos, consideremos p variables:

$$\mathbf{X_1, X_2, X_3, X_4, \ldots, X_p}$$

El análisis de componentes principales busca obtener **p** combinaciones lineales de las variables, análogas a las expresadas en (2.1.1), de tal manera que cada una de estas combinaciones lineales capte la mayor parte posible de la variación en las variables originales y, adicionalmente, sean linealmente independientes entre sí.

De esta manera un componente principal cualquiera -por ejemplo $\mathbf{y_j}$- lo escribimos como una combinación lineal de las **p** variables iniciales:

$$y_j = a_1 X_{1j} + a_2 X_{2j} + \ldots + a_p X_{pj}.$$

$j=1,2\ldots n.$ (2.1.2)
$i=1,2\ldots p.$

Donde:

y_j : es el j-ésimo componente principal.

X_{ij} : representa la i-ésima variable para la j-ésima observación.

a_i : representa el coeficiente desconocido.

Hacemos notar que de cada variable poseemos n observaciones.

En forma matricial podemos expresar las n observaciones de cada una de las p variables mediante la matriz X:

$$\begin{bmatrix} X_{11} & X_{21} & X_{31} & \ldots & X_{p1} \\ X_{12} & X_{22} & X_{32} & \ldots & X_{p2} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{1n} & X_{2n} & X_{3n} & \ldots & X_{pn} \end{bmatrix} = X .$$

Cada uno de los componentes principales y_j los podemos juntar en un vector columna de componentes, Y:

$$Y = [y_1, y_2, y_3, \ldots, y_n]' .$$

Y los coeficientes desconocidos a_i, los resumimos en un vector columna de coeficientes, A:

$$A=[a_1,a_2,a_3,...,a_p]'$$

Así, nuestro sistema completo lo podemos anotar de la siguiente manera:

$$Y_{nx1}=X_{nxp}A_{px1} \quad (2.1.3)$$

Nuestro propósito es estimar el vector de coeficientes -también llamado de ponderaciones- **A**, considerando que nos interesa capturar la máxima variabiliad de cada una de las combinaciones lineales y_j. Se trata entonces de maximizar:

$$VAR^E[Y]=A'SA \quad (2.1.4)$$

Donde $S=VAR^E[X]$

que es la expresión de la varianza estimada (VAR^E)[2] del vector de componentes la cual puede tener solución indeterminada si no incluímos restricciones. La usual en estos casos es normalizar el vector **A** de tal forma que:

$$A'A=[a_1]^2+[a_2]^2+...+[a_p]^2=1 \quad (2.1.5)$$

Lo que solamente significa que el vector que define la combinación de ponderaciones debe ser de longitud unitaria.

Nuestro problema es de maximizar **A'SA** sujeto a la restricción

----

2. En la sección 2.3 de notación matemática demostramos lo mencionado y desarrollamos otros aspectos relacionados.

de que **A'A=1**, que es un caso sencillo de multiplicadores de Lagrange. La función a maximizar es:

$$A'SA-\gamma[A'A-1] \quad (2.1.6)$$

Sujeta a la restricción mencionada, en la cual γ es el multiplicador de Lagrange. El vector de derivadas parciales de esta función con respecto al vector **A** es:

$$2SA-2\gamma A \quad (2.1.7)$$

El cual, después de igualar a cero, se reduce a

$$[S-\gamma I]A=0 \quad (2.1.8)$$

y que tiene una solución diferente de cero sólo si

$$DET[S-\gamma I]=0 \quad (2.1.9)$$

Esta última expresión no es sino una ecuación polinomial en γ y corresponde precisamente al problema de encontrar las p raíces características de la matriz de covarianzas S. Estas raíces cumplen con:

$$\gamma_1 > \gamma_2 > \gamma_3 > \ldots > \gamma_p \quad .$$

Para determinar cuál de estas raíces se usa para encontrar el vector característico que maximiza **A'SA** premultiplicamos a (2.1.8) por **A'**:

$$A'[S-\gamma I]A=A'SA-A'\gamma A=0 \quad (2.1.10)$$

De donde $A'SA=\gamma$ (2.1.11)

por tanto $VAR[Y]=\gamma$ (2.1.12)

En consecuencia, para maximizar la varianza de **Y**, escogemos la mayor raiz característica de la matriz de covarianzas S. LLamamos a esta raiz γ_1 y a_1 al vector característico normalizado asociado a este eigenvalor encontrado.

El primer componente principal está dado por:

$y_1=x_1A$, con varianza igual a γ_1 .

En general, con **p** variables, el primer componente principal y_1 es una combinación lineal de las **p** variables con coeficientes dados por el vector característico normalizado asociado con el valor latente más grande de la matriz S.

El segundo componente principal y_2, es la combinación lineal -independiente de la primera- de las **p** variables con coeficientes iguales al vector latente normalizado asociado a la segunda mayor raiz característica de la matriz de covarianzas S y así se sigue sucesivamente hasta conformar el p-ésimo c. principal.

En forma matricial los **p** componentes principales están dados por:

$$Y=XA \quad (2.1.13)$$

Donde **A** es una matriz cuyas columnas son cada uno de los vectores característicos asociados a cada una de los valores propios de la matriz de covarianzas, con la propiedad particular de que:

$$Y'Y=A'X'XA=\text{DIAG.}[\gamma_1, \gamma_2, \ldots, \gamma_p] \; .$$

y cada uno de los elementos:

$$y_i'y_i=\gamma_i=\text{VAR}[y_i] \; .$$

$$y_i'y_j=0=\text{COV}[y_i, y_j] \; .$$

Si el rango de la matriz **X** fuera menor que **p**, $(r < p)$, entonces $(p - r)$ valores propios serían iguales a cero -existirían en la matriz **X** variables que son combinaciones lineales de otras- y la variación en las variables originales se puede expresar con los **r** componentes, lo que equivale a decir que se consideran sólo las **r** variables independientes. De manera semejante, a pesar de que el rango de la matriz **X** sea **p**, algunos de los eigenvalores pueden ser muy cercanos a cero y por ende un número más pequeño que **p** de componentes podría explicar la mayor parte de la variabilidad de las X originales.

La contribución relativa de cada uno de los C.P. se evalúa por la proporción que tiene la varianza de un componente sobre la variabilidad total:

$$\frac{\text{VAR}[y_i]}{\text{Var. Total}} = \frac{\gamma_i}{\Sigma\gamma_i}$$

2.1.2 COMENTARIOS.

A través del breve esbozo que hemos presentado podemos concluir que el análisis de C.P. es una técnica descriptiva alternativa a la presentación de los datos originales que tiene como ventajas sobre estos últimos el que se puede reducir el número inicial de variables, el que los componentes no son correlacionados entre sí y el que nos muestre la proporción de varianza explicada por cada uno de los componentes. Lo anterior se traduce en que un conjunto más pequeño de componentes es susceptible de representar la mayor parte de la variabilidad de la información original.

No hemos explicitado supuestos especiales -por ejemplo, distribucionales- sobre el grupo de variables a las que se les aplica un análisis de C.P. Ya se mencionaba en el capítulo anterior un problema delicado en el A.C.P.: la interpretación y significado. Veamos algo en referencia a estos dos aspectos: supuestos e interpretación.

Una complicación inicial surge cuando las variables no están medidas en las mismas unidades. La salida que intenta darse a este asunto es mediante el proceso de estandardización de las variables. En la sección que trata el manejo matemático se desarrollan los procedimientos. Nos interesa señalar acá que en los casos en que se obtengan componentes estandardizados el significado de cada una de las variables es teóricamente indeterminado.

El problema es similar al que acontece cuando en un análisis de regresión los coeficientes estimados obtenidos pueden dar una indicación del comportamiento de las variables cuando éstas se miden en sus escalas originales. Con la estandardización desaparece el término independiente y cualquier indicación del comportamiento de los nuevos coeficientes obtenidos de variables 'puras' carece de significado.

Utilicemos un ejemplo simple y bastante usual en estos casos para ilustrar lo que deseamos decir: la estimación de los parámetros en una regresión de ingresos y consumo familiares para encontrar mediciones del consumo autónomo y de la propensión al consumo dentro del enfoque keynesiano. Al correr la regresión podemos encontrar algo como lo siguiente:

$$[c_i]^E=[B_0]^E+[B_1]^E y_i$$

Donde:
y: representa el ingreso familiar medido en pesos.
c: representa el consumo familiar medido en pesos.
$[B_0]^E$: estima el consumo autónomo.
$[B_1]^E$: estima la propensión al consumo.

En este modelo los dos coeficientes de regresión tienen una interpretación fácil e inmediata. Si estandardizamos el ingreso y el consumo obtendríamos:

$$[c_i]^* = \frac{c_{ij}-c_i}{S_c}$$

$$[z_i]^* = \frac{y_{ij} - y_i}{S_y}$$

Donde
S_k : es la desviación estándard de la variable (k).

Al correr esta nueva regresión los resultados serían:

$$[c_i]^{*E}=z_i{}^*[A_i]^E$$

En esta situación la interpretación teórica de los coeficientes no es fácil. Primero que nada, B_0 desaparece, lo que podría llevar a pensar que el consumo auntónomo es inexistente en la nueva formulacióon. Antes B_1 era la propensión al consumo, explicaba la proporción en que variaba el consumo ante cambios en el ingreso. Aunque funcionalmente representa la proporción con la que responde c^* a las variaciones en el z^* no podemos decir nada más. Se ha perdido sentido teórico, las nuevas variables han dejado de ser consumo e ingreso, a pesar de que los representan de manera estandardizada. No obstante, por la simplicidad del ejemplo podemos darle cierto significado al nuevo coeficiente.

Con componentes el cambio es mucho más brusco. Si las variables con las cuales contamos son:

X_1 : productividad agrícola medida en ton/ha.
X_2 : valor de producción agrícola (pesos constantes).
X_3 : número de trabajadores por hectárea.
X_4 : régimen pluviométrico (mm cúbicos por año).
X_5 : distancia en kilómetros al centro de mercadeo.
X_6 : número de predios agrícolas.
X_7 : costo de abonos, fungicidas e insecticidas (pesos cttes.).
X_8 : régimen de explotación (variable indicadora).

0: pequeña propiedad.
1: mediana propiedad.
2: propiedad ejidal.
3: arrendamiento.

Como estas variables están medidas en diferentes unidades lo aconsejable según la técnica en referencia es efectuar una estandardización de ellas. Si la intención es encontrar unos nuevos componentes que nos simplifiquen el estudio agrícola, la interpretación de los resultados es sin duda alguna un oficio de prestidigitadores.

Aun si pudiéramos uniformizar las mediciones de las variables mediante, por ejemplo, medir la variable X_3 en salarios pagados por hectárea, o la variable X_6 en costos de transporte por tonelada/kilómetro etc.., los resultados que se obtienen diferirán de los encontrados cuando se utilicen las variables originales, las estandardizadas o las homogeneizadas por algún otro criterio: las ponderaciones de cada una de las variables -los a_{ij}- en cada componente y las varianzas de éstos van a ser distintas. Si alteramos la escala en alguna(s) variable(s), se obtiene una diferente ecuación polinomial en (2.1.9) y, por ende, distintas soluciones para los eigenvalores. No hay entonces un camino fijo.

A pesar de que se puede contar con un elemento estadístico adicional -las correlaciones entre las variables originales y los componentes- como se menciona en la sección última de este capitulo, las conclusiones resultantes no son de inmediata obtención y, más aún, no necesariamente son únicas.

Existe otro obstáculo en la técnica misma. Al escribir un componente como:

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p \ .$$

y sabiendo que

$\Sigma[a_{1j}]^2 = 1$ y que $VAR[y_1] = \gamma_1$ -->

Tenemos que las a_{1j} forman el primer vector propio relacionado con el mayor valor característico y dado que la varianza de cada componente es igual a ese valor, desde el punto de vista de las unidades de medición, variables con escalas y varianzas grandes arrojan coeficientes chicos y variables medidas en escalas pequeñas y varianzas pequeñas generan coeficientes grandes. Por este motivo las ponderaciones -coeficientes- van a destacar con diferente importancia a unas y a otras. De ahí que sea recomendable hacer análisis de C.P. cuando las variables estén medidas en rangos no muy dispares y se aconseja acudir a las correlaciones de las variables con los componentes para guiar la definición de las distintas X que entran en la composición de los nuevos ejes encontrados.

2.1.3 ILUSTRACIONES.

Por considerarlo de interés por la discusión que suscita la intérpretación y el análisis de los resultados, resumiremos un trabajo de economía monetaria desarrollado por Bolch y Huang [3]. Ellos están interesados en desentrañar los elementos monetarios y reales de un grupo reducido de variables económicas. Se plantean la posibilidad de construir algún indicador monetario que pueda reflejar si la política monetaria es expansionista o contraccionista durante un semestre. Como hay discrepancias al responsabilizar a una sola variable -por ejemplo, las existencias monetarias- como la única indicadora de las condiciones de la política monetaria -por cuanto un incremento en las existencias no necesariamente es síntoma de expansión monetaria puesto que sus cambios pueden ser endógenos al sistema económico y no una respuesta automática a los cambios en la política monetaria-, consideran si es posible encontrar una nueva variable que pueda precisar la trayectoria futura de la política.

Comienzan su estudio suponiendo que todas las series monetarias tienen un doble componente: uno monetario y otro real, independientes entre sí. En el caso de la tasa de interés plantean que un indicador de política que podría estar libre del fenómeno no monetario es el componente principal asociado con el comportamiento monetario. Escogieron tres series con información mensual durante tres años de las siguientes variables en E.E.U.U.:

X_1 : porcentaje de cambio en las reservas libres de todos los bancos miembros del Sistema de Reserva Federal.

X_2 : porcentaje de cambio en la tasa prima de interés.

X_3 : porcentaje de cambio en las existencias monetarias.

----

3. Bolch, p. 240 y ss.

Encontraron que las raíces características y sus vectores asociados son:

CUADRO No. 2.1

C.P.	Raíz	Vectores			% Varianza acumulado.
y_1	1.76	.6059	-.6391	.4737	58.7
y_2	.79	-.4516	.2139	.8662	85.0
y_3	.45	.6549	.7388	.1590	100.0 %

De acuerdo a los resultados anteriores, cada uno de los componentes estaría formado por:

$$y_1 = .606\ x_1 - .639\ x_2 + .474\ x_3$$

$$y_2 = -.452\ x_1 + .214\ x_2 + .866\ x_3$$

$$y_3 = .655\ x_1 + .739\ x_2 + .159\ x_3$$

Tratando de descifrar el C.P. monetario, los autores enfatizan que el primer componente explica la mayor parte de la variabilidad -59 %- del sistema y subrayan además los signos de los coeficientes. Ilustremos algunas conclusiones: en consonancia con los signos de los parámetros del primer componente se observa un incremento en las reservas libres y en las existencias monetarias y un decrecimiento en las tasas de interés, lo cual puede interpretarse como un relajamiento de las restricciones monetarias, de esta manera con ponderación negativa en la tasa de interés y signos positivos en las otras dos variables conducirían a pensar a y_1 como el componente monetario.

El segundo componente $-y_2-$ tiene coeficientes con signos tales que podrían concebirlo como el componente real: un aumento en la demanda real de crédito se espera que esté asociado con una elevación en las existencias monetarias y en la tasa de interés y una disminución en las reservas libres. Estos tres aspectos son los que inducen a considerar a y_2 como el componente real.

La anterior no es más que una interpretación que no tiene por qué tener validez general. Los intentos de interpretar los componentes principales como las características primitivas o latentes de las variables inspirados en las raíces y vectores latentes es extremadamente riesgoso:

Primero, porque puede haber un desacuerdo entre teóricos sobre el significado del comportamiento de los signos.

Segundo, porque el análisis supone que existen características primitivas de la información y que son linealmente independientes, lo cual puede no ser cierto: pueden no existir esas características primitivas y de hacerlo no tienen por qué ser independientes.

Tercero, y aquí cito a los autores cuando se refieren a este tipo de estudios:

> " El análisis de componentes principales es a menudo usado como una técnica 'de pesca'. Hay una mala broma entre estadísticos: 'cuando todo falle, trate análisis de componentes principales o su extensión, análisis de factores."[4]

----

4. Bolch, p. 247.

Y Bolch y Huang culminan su ejercicio con una sana advertencia:

> " Se podría recordar que la inferencia estadística clásica no consiente hipótesis a ser desarrolladas de la información y que la pesca de información para hipótesis reales, siempre conduce a un fracaso."[5]

Consideremos ahora otra situación: el Cuadro No. 2.2 presenta tres diferentes índices de costo de vida durante el transcurso de once años [6] y se quiere determinar cual de ellos adoptar para un estudio económico de largo plazo.

La solución puede ser formar una combinación lineal de estos índices y usarla como una representación de todos ellos. Un promedio simple de las series sería una alternativa satisfactoria; otra, más afín a la variabilidad del conjunto de datos, sería encontrar un componente principal que nos recoja lo fundamental de las tres. No hay obstáculos especiales con la medición puesto que las tres variables están expresadas en las mismas unidades y el rango de valores no difiere de manera particular. Los resultados de un A.C.P. se presentan en el Cuadro No. 2.3, en la página siguiente.

----

5. Bolch, p. 249.

6. Siguiendo con la lógica de Bolch y Huang, se incluye otro ejemplo presentado por ellos con fines eminentemente ilustrativos.

C U A D R O No. 2.2

Indices de Costo de Vida

Año	F.R.B. N Y	HANSEN	BURGUESS
0	80.0	76.0	67.7
1	82.0	75.0	70.6
2	84.0	78.0	74.8
3	88.0	81.0	74.8
4	87.0	81.0	76.1
5	87.0	81.0	76.0
6	90.0	85.0	78.2
7	95.0	90.0	82.0
8	91.0	87.0	84.4
9	91.0	91.0	88.6
10	96.0	94.0	93.1

C U A D R O No. 2.3

Componentes Principales.

C.P.	Raiz	V e c t o r e s			% Varianza por cpte.
y_1	118.44	.4398	.5762	.6889	96.07
y_2	4.00	-.7170	-.2365	.6557	3.24
y_3	.85	.5408	-.7823	.3091	.69

El primer componente y_1 -que absorbe el 96 % de la variabilidad de los indices- se puede calcular mediante:

$$y_{1j} = .4398\ x_{1j} + .5762\ x_{2j} + .6889\ x_{3j}$$

$$y_{11} = .4398\ (80.0) + .5762\ (76.0) + .6889\ (67.7) = 125.61$$

Se consiguió obtener un nuevo índice de costo de vida para los once períodos anuales; los resultados se resumen en el Cuadro No. 2.4. Los tres índices de precios están altamente correlacionados lo que hizo posible combinarlos en uno solo. Si hubiésemos tomado un promedio simple de los indicadores, los coeficientes para cada una de las series serían iguales a .3333, bastante diferentes de los obtenidos por medio del A.C.P.: las ponderaciones son distintas.

CUADRO No. 2.4

Indice Resultante.

Período	Indice: y_{1j}
0	125.61
1	127.91
2	133.42
3	136.90
4	137.36
5	137.29
6	142.43
7	150.13
8	148.30
9	153.50
10	160.52

En economía es muy frecuente sintetizar una gran complejidad de agregados numéricos convirtiendo un sistema de p dimensiones en una sola dimensión, tal es el caso de los números índice -de precios, de salarios, de costo de vida, de actividad comercial...-. En la elaboración de estos índices se ponderan los ítems por las cantidades constitutivas para recoger la importancia relativa de cada uno de ellos. Una buena posibilidad que tendría en cuenta de una manera más cercana la variabilidad de los ítems puede ser el primer componente principal. Por lo menos dos buenas razones apuntan en su favor: por su obtención absorben la mayor variación posible en la información y porque suministra sus propias ponderaciones. La comparación entre los índices usuales basados en ponderaciones fijas y un índice constituido por componentes permitiría sopesar las varianzas del C.P. y la del índice típico. Kendall[7] considera que se puede obtener, a través del componente principal, alguna medida del alcance en el cual los índices usuales pierden eficiencia.

Dentro de este espíritu se realizó un intento de esta naturaleza utilizando información mensual del I.N.P.C. de Méexico de 1982 clasificado por durabilidad de los bienes.[8] Los datos se presentan en el cuadro No. 2.5. Utilizando la matriz de covarianzas de los tres tipos de iíndices, se encontró que el primer componente principal explicaba el 91.63 % de la varianza del conjunto, proporción superior a la encontrada en el trabajo de Bolch y Huang del cuadro No.2.3. Resumimos los resultados en el cuadro 2.6 en el que se observa la composición de los vectores propios y la proporción de varianza por componente.

La gráfica No. 2.2 se refiere a la representación visual del nuevo componente derivado de los tres índices. La alineación de puntos nos ilustra fehacientemente la posibilidad de concebir al primer componente como una particular combinación de los tres tipos de bienes. Esa alineación de los puntos coordenados de los tres bienes en la diagonal de la gráfica responde al alto porcentaje de varianza explicada por ese primer componente.

----

7. Kendall, p. 25.

8. Nafinsa, p. 217.

C U A D R O No. 2.5

México: Indice Nacional de Precios al Consumidor.
Clasificación por Durabilidad de Bienes.
- información mensual de 1982 -

período	índice general	Durabilidad de Bienes: bienes no duraderos	bienes duraderos	servicios
enero	223.7	219.7	203.8	232.6
febrero	232.5	225.4	220.6	244.9
marzo	241.0	232.9	231.7	254.7
abril	254.4	245.3	252.5	266.8
mayo	268.4	259.7	266.7	280.6
junio	281.3	271.1	275.7	296.9
julio	295.8	285.8	291.3	310.7
agosto	329.0	323.0	318.8	337.7
septiembre	346.5	340.6	345.8	352.7
octubre	364.5	360.3	365.8	366.9
noviembre	382.9	378.7	395.8	382.5
diciembre	423.8	427.4	422.3	412.5

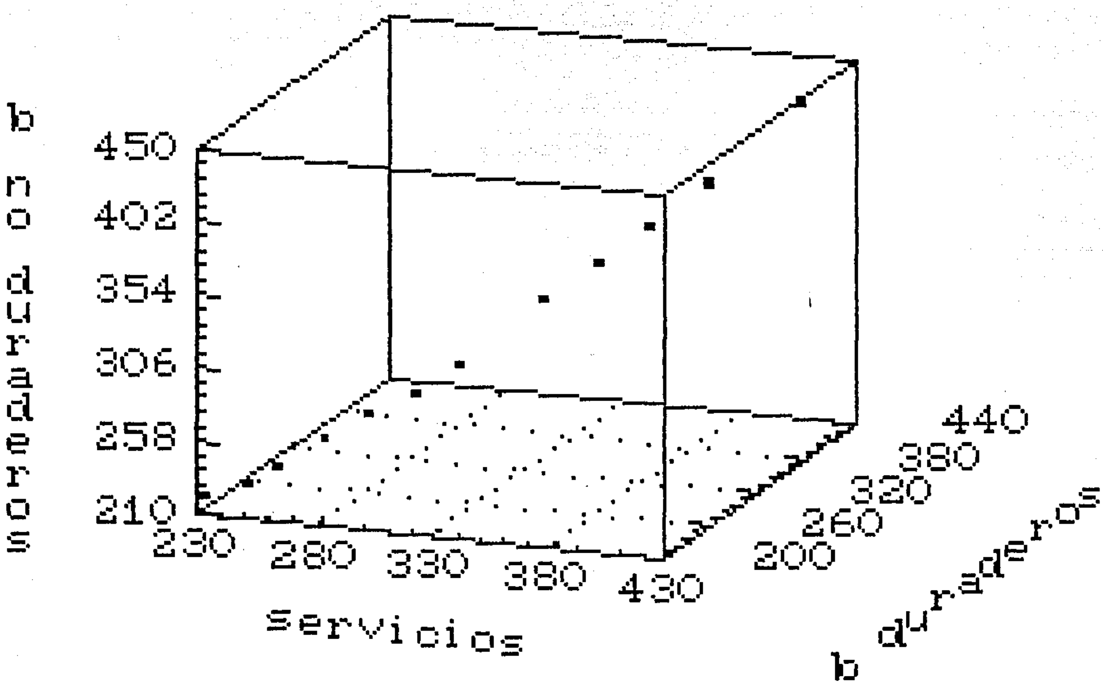
C U A D R O No. 2.6

México: componentes principales de los indices
por durabilidad de bienes.

C.P.	raiz	v e c t o r e s			% var. por componente.
y_1	12974.000	.59521	.79879	-.10423	99.63
y_2	32.373	.62082	-.53524	-.57280	.25
y_3	15.889	.51333	-.27480	.81304	.12

GRAFICA 2.2

EXPOSICION VISUAL DE LA COMBINACION DE BIENES.
MEXICO: INDICE RESULTANTE. 1982



Finalmente en el cuadro No. 2.7 se anotan los elementos del primer componente principal, es decir, las variaciones de precios a que da lugar el nuevo índice encontrado, el cual se obtuvo por la suma del producto de cada uno de los respectivos índices por el vector característico asociado al primer eigenvalor obtenido.

C U A D R O No. 2.7.

México: Indice resultante. [1982]

periodo	Indice C.P.
enero	376.1
febrero	396.2
marzo	412.6
abril	439.1
mayo	463.5
junio	484.2
julio	509.7
agosto	562.6
septiembre	597.5
octubre	628.9
noviembre	666.5
diciembre	727.2

Como se observa del cuadro No. 2.6 las ponderaciones son diferentes para cada uno de los tres índices de tipos de bienes: 0.59 para los bienes no duraderos, 0.62 para los bienes duraderos y 0.51 para servicios. Los resultados del nuevo índice construido se encuentran en el cuadro No. 2.7 y lo más llamativo es que su crecimiento es mayor que el reseñado inicialmente como índice general, lo que reflejaría un incremento inflacionario superior. Una anotación importante es que el indice obtenido por componentes varía al tenor de los sectores que están integrándolo. He ahí uno de los graves problemas de su generalización. Desde luego los índices usuales también se ven afectados por la composición de los sectores. Alternativamente la construcción de índices puede partir directamente de los precios originales detectando sus movimientos en periodos determinados y la obtención del índice obviaría la necesidad de ponderar por las cantidades.

Queremos enfatizar, por último, que en el análisis de componentes principales no hay supuestos distribucionales sobre el comportamiento de las variables X. Si adicionalmente podemos suponer distribución normal multivariada, los C.P. tendrán propiedades adicionales y podremos realizar pruebas de hipótesis sobre el número de componentes, pero no es nuestro interés profundizar en este aspecto, particularmente porque sería demasiada temeridad suponer que por lo general las variables económicas tienen distribución normal.

Las limitaciones del A.C.P. se han presentado en las páginas anteriores. Sólo nos resta decir que la técnica en mención no es sino una forma diferente de enfocar una misma información. Es un proceso descriptivo alternativo que puede ser ventajoso según el tema tratado y los objetivos que se persiguen. La idea que está implícita en el análisis de componentes principales es la de observar el comportamiento de las variables desde un punto de vista que permite apreciar la mayor variabilidad del conjunto de información. La metamorfosis de variables a componentes no es sino una rotación de los ejes que ofrece una ventajosa y distinta visión de la misma realidad. Más aún, si suponemos distribución normal multivariada, el A.C.P. no es más que la determinación de los ejes de un elipsoide de concentración.

2.2 ANALISIS DE FACTORES.

A comienzos del presente siglo el sicólogo y estadistico Charles Spearman publicó un trabajo denominado **"LA INTELIGENCIA GENERAL OBJETIVAMENTE MEDIDA Y DETERMINADA"**. Lo ambicioso del título obedecía a las conclusiones por él obtenidas en la aplicación de una técnica matemática que habia estado elaborando desde tiempo atrás y que pretendía encontrar los elementos subyacentes e implícitos en un conjunto de información sobre pruebas mentales.

Spearman estudió las correlaciones entre estas pruebas en diferentes individuos y notó que los coeficientes de correlación de los puntajes obtenidos eran positivos y bastante altos. Si todas las correlaciones eran positivas, supuso que estos resultados estaban indicando algo que iba más allá de las simples mediciones encontradas. Las fuertes correlaciones positivas claramente le insinuaban que cada prueba no media sólo un atributo independiente del funcionamiento mental: debía existir una estructura básica que estaba detrás de todo y que generaba las respuestas en las altas correlaciones.

Se sugerían algunas alternativas que explicaban esta estructura subyacente: o bien las correlaciones positivas podían reducirse a un conjunto más pequeño de atributos independientes, o bien se reducirían a un factor único y general. Evidentemente cualquiera que fuera la situación, Spearman entendía que no toda la información en la matriz de correlaciones obedecía a una u otra estructura. Debia existir un elemento residual que respondía a la información especifica de cada prueba que no estaba relacionado con ninguna otra. A esta parte la llamó varianza residual de cada prueba y se explicaba por la cualidad especifica que por sí mismo debía tener cada uno de las exámenes.

Sinteticemos las dos proposiciones:

1.) Cada prueba se explicaría por una parte específica -lo que Spearman llamó **s**- y por un factor único e implícito, la inteligencia general, llamado **g** en la terminología de Spearman.

2.) Cada examen, además del factor específico **s**, obedece a un **CONJUNTO** de factores independientes - y no a un único factor - que subyace en la información allegada.

Si la primera alternativa fuera la cierta, Spearman había encontrado no sólo una definición de la inteligencia, sino que además era capaz de medirla objetivamente.

La técnica desarrollada por Charles Spearman se planteó como un procedimiento para decidir si la varianza común de la matriz de correlaciones entre las características iniciales se explicaba por ese factor único o por ese conjunto de factores independientes. Por el título de su trabajo se deduce cual fue la conclusión a la que finalmente llegó. Estábamos en el nacimiento de una técnica estadística que podía encontrar modelos de causalidad, que era capaz de descubrir los factores determinantes que subyacen en la realidad y que podía aplicarse en campos distintos de la sicología. La sola mención y definición de factores subyacents da lugar a todo tipo de especulaciones causales sobre el comportamiento de la realidad. El argumento parece seductor: la información por sí misma no deja entrever la estructura interna explicativa de los fenómenos, ésta hay que desentrañarla, descubrirla y hacerla explícita, el modelo de análisis de factores (M.A.F.) tiene la capacidad de realizar esa tarea. Es una técnica cuya finalidad es ésa precisamente: descubrir la estructura elemental oculta y escondida. Por si fuera poco, los resultados obtenidos son susceptibles de someterse a pruebas de hipótesis estadísticas capaces de probar si obedecen a factores objetivamente determinados o si son expresión de un comportamiento aleatorio en las variables que estamos estudiando.

2.2.1 EL DESARROLLO DEL MODELO.

A **n** individuos se les hicieron **p** diferentes pruebas o exámenes. Para explicar por qué los resultados están altamente correlacionados Spearman sugirió el modelo siguiente:

$$X_{ij} = \gamma_i g_j + s_{ij} \text{ con } i=1..p;\ j=1..n \quad (2.2.1)$$

Donde:

X_{ij}: representa el puntaje del j-ésimo estudiante en el i-ésimo examen.

γ_i : es el coeficiente que representa la ponderación o importancia del factor inteligencia general. En otras palabras, representa el grado en el cual la inteligencia determina el rendimiento del i-ésimo examen.

s_{ij}: es la parte específica atribuída a la prueba; es el residual correspondiente a la información específica de cada examen no relacionada con ningún otro.

Por regla general se supone que **g** se distribuye normalmente y que los residuales s_{ij} están normal e independientemente distribuidos. Es usual escalar las variables X para que tengan varianza unitaria, además se supone que también están normalmente distribuidas. Subrayamos que la estructura del modelo depende sólo de las correlaciones entre las X.

Dentro del campo mismo de la sicología las críticas abundaron, en especial se cuestionaba al modelo por ser poco realista y demasiado restrictivo.[9] Particularmente, porque sólo podía pensarse en un factor único -la inteligencia general-. Acaso no era posible que el proceso se explicara por otro tipo de factores como la habilidad linguística, la memoria, la velocidad, -la segunda alternativa de Spearman- que influyeran en unos exámenes, pero no en otros ?

Aceptadas estas razones la formulación inicial se transformó en:

$$X_{ij} = {}_{j}\Sigma^{k} \gamma_{ik} f_{jk} + e_{ij} \quad (2.2.2)$$

En el cual se introducen **k** distintos factores, llamados factores comunes, donde la **f** representa cada factor común y la **e**, el término residual o factor especifico. La γ_{ik} indica la importancia de los factores comunes en los diferentes exámenes.

Mencionemos los supuestos básicos del M.A.F. Primero que nada tenemos que enfatizar el hecho de que el comportamiento de los factores con respecto a los atributos iniciales es de carácter lineal. De por sí es un supuesto muy fuerte, no obstante sería el menos cuestionable entre todos. El M.A.F. considera que las variables pueden expresarse por la combinación de **k** factores, fijado este **k** de antemano. De ahí algunas criticas dentro del campo mismo de la sicometría como de poco realista.

Se supone también que cada uno de los factores comunes se distribuye normalmente con media cero y varianza igual a la unidad y que, además, no tienen ningún patrón de asociación lineal entre sí. Los factores especificos se suponen con distribución normal con media cero y varianza dada -Φ_i-, y se supone que la covarianza de los factores comunes con los

----

9. En la sección siguiente expondremos algunas criticas al modelo de factores, los comentarios que en esta parte nos interesa subrayar son los que dieron lugar al perfeccionamiento del modelo.

específicos y al interior de ellos es inexistente: los factores son independientes.

Dados estos supuestos encontramos [10] que la varianza de la i-ésima variable se descompone en:

$$VAR[X_i] = {}_j\Sigma^k[\gamma_{ij}]^2 + \Phi_i \quad (2.2.3)$$

Donde:

Φ_i: se refiere a la varianza del factor específico.

$\Sigma[\gamma_{ij}]^2 = h_i$: es la varianza de los factores comunes.

A h_i se le acostumbra a llamar **comunalidad** de la variable X_i. Recordemos que cada uno de los γ es la ponderación de cada factor, también se les denomina la carga del factor en la variable asociada.

De (2.2.3) la variabilidad de cada X se descompone en una parte debida a los factores comunes, llamada **comunalidad** de la variable X_i -h_i- y otra atribuida a la variabilidad específica de la X_i -Φ_i-. Si la mayor parte de la varianza de esta última corresponde a la varianza de los factores comunes significaría que el modelo propuesto sí es el adecuado para explicar a las variables iniciales: la variable X_i se puede expresar como una combinación lineal de los γ factores comunes. Lo anterior quiere decir que la varianza del factor específico no tiene peso importante en la varianza de las X_i. Por el contrario si en la varianza del factor específico se concentra la mayor parte de la variabilidad del modelo, el comportamiento de las X_i no podría representarse por (2.2.2): el modelo propuesto no tiene utilidad, los factores no son combinación lineal de las variables primitivas.

----

10. Ver la sección del tratamiento matemático para una formalización de lo mencionado y de lo que se presenta a continuación.

Otro resultado importante es el que la covariación entre dos de las variables iniciales se puede expresar como la sumatoria del producto de las respectivas cargas de los factores a ellas asociadas:

$$COV[X_i, X_j] = {}_m\Sigma^k \gamma_{im}\gamma_{jm} \quad (2.2.4)$$

De otra parte, al considerar la covariación entre cada variable con cada factor encontramos que la podemos indicar por la carga del factor sobre la variable:

$$COV[X_i, f_j] = \gamma_{ij} \quad (2.2.5)$$

Es decir, gracias a los supuestos, la estructura de covarianzas del M.A.F. se resuelve en términos de los γ. Cuando se quiere interpretar los factores el interés se centra en la magnitud y en el signo de las cargas para cada uno de los factores comunes.

Matricialmente el modelo lo escribimos como:

$$\mathbf{X} = \mathbf{\Gamma F} + \mathbf{E} \quad (2.2.6)$$

Donde:

$\mathbf{\Gamma}$: matriz (pxq) llamada matriz de cargas.

$\mathbf{F}$: vector (kx1) de factores comunes.

$\mathbf{E}$: vector (qx1) de factores específicos.

De acuerdo a los supuestos encontramos que:

$$VAR[\mathbf{X}] = \mathbf{S} = \mathbf{\Gamma\Gamma}' + \mathbf{\Phi} \quad (2.2.7)$$

Donde: $\boldsymbol{\Phi} = \mathrm{DIAG}[\Phi_{11}, \ldots, \Phi_{pp}]$

Cuando las variables X_i se estandardizan, se utiliza la matriz de correlaciones $\mathbf{R}$ que se define como:

$$\mathbf{R} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}' + \boldsymbol{\Phi} \quad (2.2.8)$$

La cuestión a la que nos enfrentamos es el que no hay una única solución a la descomposición de la matriz de covarianzas de **X** en las matrices de cargas y de factores específicos. Notemos que si la estimación de la matriz de cargas la encontramos como:

$$\boldsymbol{\Gamma}^{*} = \boldsymbol{\Gamma}\mathbf{B} \quad (2.2.9)$$

Donde $\mathbf{B}$: es una matriz (k x k) ortogonal.

No hay una única matriz **B** ortogonal. En vez de una solución única, tenemos un espacio de soluciones. Esta indeterminación en la estimación de la matriz de cargas se puede resolver mediante la rotación de los factores de tal manera que se cumplan restricciones [11] como:

$\boldsymbol{\Gamma}'\boldsymbol{\Phi}^{-1}\boldsymbol{\Gamma}$ sea diagonal. (2.2.10)

$\boldsymbol{\Gamma}'\mathbf{D}^{-1}\boldsymbol{\Gamma}$ sea diagonal. (2.2.11)

----

11. Mardia, p. 58.

Donde $\mathbf{D} = DIAG[s_{11}, \ldots, s_{pp}]$

Encontrada la matriz de cargas estimamos la matriz de varianzas de los factores específicos mediante:

$$\Phi^{*} = S^{*} - \Gamma^{*}\Gamma^{*\prime} \quad (2.2.12)$$

$$[\Phi_{ii}]^{*} = s_{ii} - {}_{j}\Sigma^{k}[\gamma_{ij}]^{2*} \quad (2.2.13)$$

Donde s_{ii}: es el elemento referenciado por los subíndices de la matriz de covarianzas de **X**. (Con i=1..p) Esto significa que Φ es una función de Γ. Si usamos la matriz de correlaciones la expresión (2.2.13) se convierte en:

$$[\Phi_{ij}]^{*} = 1 - \Sigma[\gamma_{ij}]^{2} \quad (2.2.14)$$

2.2.2 METODOS DE ESTIMACION.

2.2.2.1 Método del Factor Principal.

Se trabaja con la matriz de correlación estandardizando las variables originales. Intuitivamente la idea es la siguiente: como el M.A.F. trata de buscar una explicación a la fuerte estructura de correlación, se espera que las **comunalidades** sean más altas entre mayor sea la correlación de las variables de la matriz **X**. Por esta razón una estimación preliminar de la i-ésima **comunalidad** puede encontrarse por el cuadrado del coeficiente de correlación múltiple de la i-ésima variable contra todas las demás, o bien, por el coeficiente de correlación más alto encontrado en la matriz de correlaciones. En la matriz:

$[R-\Phi]$ (2.2.15)

los elementos de la diagonal están constituidos por las **comunalidades** estimadas:

$$[h_i]^2 = 1 - \Phi_{ii}$$

Si obtenemos los eigenvalores y eigenvectores ortonormales relacionados $-a_i$ y t_i, respectivamente- podemos estimar la i-ésima columna de la matriz de factores mediante:

$$[\gamma_i]^2 = [a_i]^{1/2}\, t_i \text{ con } i=1..k \quad (2.2.16)$$

o en forma matricial:

$$\Gamma^2 = T_1 [A_1]^{1/2} \quad (2.2.17)$$

Donde:

$T_1 = [t_1, t_2, \ldots, t_k]$

$A_1 = \mathrm{DIAG}[a_1, a_2, \ldots, a_k]$

2.2.2.2 Método de Máxima Verosimilitud.

En esta aproximación se supone además que las **X** están normalmente distribuidas. Mencionaremos que se requiere una serie de restricciones y que la estimación se encuentra para γ factores distintos [12]

La ventaja de este método sobre el anterior estriba en que permite una prueba de hipótesis basada en las varianzas y covarianzas de la matriz **X**. La mecánica de la prueba es la siguiente: se ajusta un factor y se prueba su significancia. Si el estadístico no resulta significativo quiere decir que ese único factor explica satisfactoriamente los resultados; si no es así, se ajustan dos factores, se prueba de nuevo y se sigue con esta dinámica hasta que el estadístico no resulte significativo. Antes de pasar a la siguiente sección debemos destacar que los resultados por uno u otro método no dan ponderaciones coincidentes.

2.2.2.3 Rotación de los Factores.

Por regla general se utiliza la rotación de los factores con fines básicamente interpretativos. La idea central es tratar de encontrar dentro del conjunto de factores específicos aquellos que más peso tengan en la explicación de las características iniciales. Habíamos anotado antes en (2.2.9) que en la solución a los factores teníamos que encontrar una matriz $B_{k \times k}$ ortogonal:

$$\Gamma^{*} = \Gamma B \quad (2.2.9)$$

----

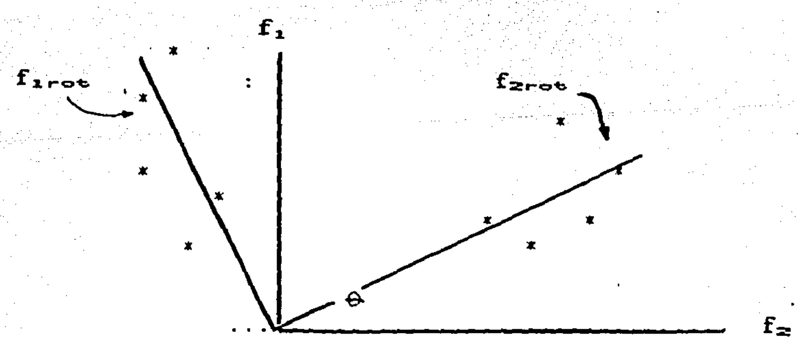
12. En la sección del tratamiento matemático se anotan las expresiones citadas.

Donde Γ^{*}: es la matriz de cargas rotadas.

Cada elemento de esta última matriz $-[\gamma_{ij}]^{*}-$ representa la carga rotada de la i-ésima variable sobre el j-ésimo factor. El objetivo es encontrar nuevos ejes ortogonales que permitan subdividir a las cargas en dos grupos: uno, de pocos factores con cargas altas, y el otro, del resto de los factores, con ponderaciones cercanas o iguales a cero. Además se trata de buscar, en la medida en que esto sea posible, que cada variable pueda asociarse con un sólo factor y que cada factor se relacione con un número mínimo de variables.

Para lograr este propósito se maximiza una función cuadrática de las cargas, esto es, se maximiza la suma de las varianzas de las cargas al cuadrado dentro de cada columna de la matriz de cargas, en las cuales cada fila se ajusta por su comunalidad. Es un procedimiento iterativo para encontrar la B particular que logra la rotación deseada. Esta técnica se conoce como método de rotación varimax. La gráfica No. 2.4 ilustra la situación.

GRAFICA No. 2.4.



Los asteriscos representan puntos -coordenadas que corresponden a los valores de las cargas de los factores uno y dos-. θ expresa el ángulo de rotación. Al llevar a cabo la rotación cada uno de los asteriscos se aproxima más a los nuevos ejes de tal manera que aquéllos que estén a lo largo del eje f_{2rot} y que tienen un valor 'alto' para él, representan una magnitud muy baja para el eje f_{1rot} Lo contrario, desde luego, también se cumple.

Otras rotaciones son la cuartimax y la oblicua. La primera sigue siendo ortogonal y busca minimizar la suma del producto del cuadrado de las cargas. En la segunda se da el ángulo de separación y se maximiza la varianza de todo el conjunto, perdiendo los nuevos ejes rotados la ortogonalidad que estaba presente en las otras situaciones. No hay un criterio general para definir el ángulo pues se considera que depende de la investigación particular.

2.2.3 RELACION ENTRE A.F. Y C.P.

Es muy común la confusión entre estas dos técnicas estadísticas y el traslape va más allá de la aplicación que de ellas se hacen en diferentes disciplinas científicas. En el campo mismo de los estudiosos del tema hay una tendencia a identificarlas.

Desde el punto de vista de los resultados, pueden encontrarse similitudes cuando las varianzas específicas son siginificativamente muy pequeñas o iguales a cero, situación que se puede presentar en el caso más puro de A.F. Desde el punto de vista de la estimación a través del método del factor principal, el M.A.F. se apoya en la técnica de componentes principales para descubrir los factores. Pero los valores y vectores característicos se derivan de una matriz que sólo de lejos se parece a la matriz de covarianzas de las **X**, como es el caso en C.P. En A.F. los valores y vectores propios se obtienen de la matriz dada en (2.2.5) en la cual se efectuó una estimación preliminar de las comunalidades tomándolas como el coeficiente de determinación múltiple de la regresión de cada variable $\mathbf{X}_i$ contra las demás variables. Hasta aquí las similitudes.

Los componentes no son más que una transformación de la información, se carece de supuestos sobre la distribución de los datos iniciales. En cambio, con A.F. estamos asumiendo un comportamiento específico de la información tal como se presenta en la formulación (2.2.2), además consideramos una serie de supuestos importantes sobre la estructura interna de (2.2.2) que nos permite llegar a conclusiones como las de (2.2.3). Mardia comenta que si estos supuestos mencionados no se cumplen los resultados obtenidos del análisis de factores son espurios.[13] En los C.P. el énfasis está puesto en un camino que va de las variables observadas a los componentes principales. En A.F. la ruta es distinta: es una transformación que parte de los factores subyacentes hacia las variables observadas.

Sin duda alguna la diferencia sustancial radica en los propósitos de uno y otro. C.P. es una descripción alternativa. A.F. es la propuesta de un muy definido y estricto modelo matemático de explicación del fenómeno estudiado. Puede concebirse que mientras componentes es una ojeada al problema desde una cierta perspectiva, el análisis de factores interpreta la realidad con un esquema de comportamiento análogo a la vieja fábula de quien se recorta los dedos para que le quepan los zapatos.

Un comentario de interés del M.A.F. en el campo de la sicología aplicada se refiere a los signos de las cargas; si éstos son menores que cero, significaría que una habilidad específica afecta negativamente el rendimiento de una prueba, aspecto que es, evidentemente, muy poco realista: si el coeficiente del factor inteligencia tiene signo negativo querría decir que la capacidad intelectiva del sujeto que presentó un examen es poco menos que la de un inválido mental. Esto es suficiente razón para creer que deben incluirse restricciones en el modelo sobre el signo y magnitud de los coeficentes con el fin de evitar que las interpretaciones sean tan disparatadas. No obstante, un mayor número de restricciones haría mucho más artificioso al M.A.F.

----

13. Mardia, p. 275.

2.3 TRATAMIENTO MATEMATICO.

2.3.1 DE LOS COMPONENTES.

2.3.1.1 La varianza del componente principal.

Si tenemos: $\mathbf{Y_1 = XA_1}$ (2.4.1)

Donde:

$\mathbf{Y_1}$: vector de n x 1.

$\mathbf{X}$: matriz (n x p) expresada como desviación de las medias.

$\mathbf{A_1}$: vector (p x 1).

Nos interesa una expresión para la varianza de $\mathbf{Y_1}$:

$$\mathbf{VAR[Y_1] = VAR[XA_1] = A'_1 VAR[X] A_1} \Rightarrow$$

$$\mathbf{VAR[Y_1] = A'_1 S A_1} \quad (2.4.2)$$

porque $\mathbf{VAR[X]} = \mathbf{S}$.

Como estimamos a la matriz de varianzas-covarianzas de **X**, es decir, la matriz S, por medio de observaciones muestrales, llamamos $S^{=}$ a la varianza estimada de S. Una estimación de la **VAR[Y_1]** es:

$$VAR^{=}[Y_1]=A'_1 S^{=} A_1 \quad (2.4.3)$$

2.3.1.2 Derivación de los componentes.

El primer componente lo podemos expresar como:

$$Y_1=XA_1$$

Como lo que interesa es transformar las **X** en un nuevo grupo de variables **Y**, no correlacionadas entre si, buscando que la primera de ellas tenga la máxima varianza posible; la segunda la mayor varianza entre aquellas no correlacionadas con la primera y así sucesivamente, se escoge A_1 de tal manera que maximice la **VAR[Y_1]** con la restricción de que $A'_1A_1=1$ para evitar la solución indeterminada en el sistema.

Definimos entonces que la función a maximizar utilizando multiplicadores de Lagrange es:

$$A'_1 S^{=} A_1-\gamma_1[A'_1A_1-1] \quad (2.4.4)$$

Donde γ_1: es el multiplicador de Lagrange.

Cuya derivada parcial con respecto a A_1 es:

$2S^{=}A_1 - 2\gamma_1 A_1$ que

tendrá un valor máximo cuando se iguale a cero, lo que da:

$$S^{=}A_1 = \gamma_1 A_1 \quad (2.4.5)$$

De esta última expresión se concluye que A_1 es el vector característico de la matriz $S^{=}$ asociada al valor característico γ_1.

Si a (2.4.5) la premultiplicamos por A_1 encontramos que:

$$A\gamma_1 S^{=}A_1 = \gamma_1 A'_1 A_1 = \gamma_1 = VAR[Y_1] \quad (2.4.6)$$

Y concluimos que A_1 es el vector característico que hace máxima la varianza del primer componente principal y por lo tanto γ_1 es la mayor raíz característica de la matriz $S^{=}$.

El segundo componente principal lo expresamos como:

$$Y_2 = XA_2 \quad (2.4.7)$$

Con la particularidad de que este componente no debe estar correlacionado con el primero, es decir:

$$COV[Y_2, Y_1] = A'_2 S A_1 = A'_2 \gamma_1 A_1 = \gamma_1 A'_2 A_1 = 0 \quad .$$

como $\gamma_1 \neq 0$, se sigue que $A'_2A_1=0$ (2.4.8)

Otra forma de ver este mismo resultado es la siguiente:

Si $Y_1=XA_1$ y $Y_2=XA_2$ entonces:

$$Y'_2=[XA_2]'=A'_2X' \quad .$$

La suma de cuadrados de Y_1 y Y_2 la anotamos como:

$Y'_2Y_1=A'_2X'XA_1$ y dado que

$X'XA_1=\gamma_1A_1$ entonces

$$Y'_2Y_1=A'_2\gamma_1A_1=\gamma_1A'_2A_1=0.$$

si y sólo si $A'_2A_1=0$.

Lo cual signfica que los vectores A_2 y A_1 son ortogonales.

En tanto deseamos escoger el vector de coeficientes del segundo C.P. de tal forma que maximice :

$$VAR[Y_2]=A'_2SA_2 \quad .$$

sujeto a las dos restricciones:

1.) $A'_2A_2=1$;

2.) $A'_1A_2=0$.

La segunda restricción tiene sentido para que el segundo componente no esté correlacionado con el primero. Ahora la tarea es maximizar:

$$A'_2SA_2-\gamma_2[A'_2A_2-1]-t[A'_1A_2].$$

donde γ_2 y t son los multiplicadores de Lagrange.

La derivada parcial de esta función con respecto a A_2 e igualada a cero es:

$$2SA_2-2\gamma_2A_2-tA_1=0 \quad (2.4.10)$$

Si premultiplicamos a (2.4.10) por A_1 tenemos que:

$$2A'_1SA_2-2\gamma_2A'_1A_2-tA'_1A_1=0.$$

siendo consecuentes con las restricciones del modelo, obtenemos:

$$2A'_1SA_2-t=0 \quad (2.4.11)$$

Lo cual significa que el coeficiente **t** vale cero porque la parte de la izquierda de (2.4.11) es la covarianza entre Y_1 y Y_2 que es igual a cero. En consecuencia, volviendo a (2.4.10) planteamos que:

$$S^{*}A_2=\gamma_2 A_2 \quad \text{y que}$$

$$VAR[Y_2]=\gamma_2=A'_2 S A_2 \quad .$$

Y ratificamos a γ_2 como la segunda raíz propia de la matriz S^{*} y a A_2 como el segundo vector característico asociado a γ_2. Los (p - 2) valores y vectores propios restantes se encuentran de manera análoga. Cada uno de los eigenvalores refleja sucesivamente de mayor a menor las respectivas varianzas de los p componentes principales.

Si organizamos en la matriz **A** a los vectores característicos obtenemos:

$$A=[A_1,\ A_2,\ ...,A_p] \quad \text{y}$$

por lo tanto, podemos representar a **Y'Y** como:

$$Y'Y=A'S^{*}A=DIAG[\gamma_1,\gamma_2,...,\gamma_p] \quad .$$

donde la varianza de cada uno de los componentes es:

$$Y'_1Y_1=\gamma_1 \quad \text{y}$$

las covarianzas:

$$COV[Y_i, Y_j] = 0 \quad .$$

De otra parte, tenemos que la varianza total de la matriz **X**, o sea **A'S^{x}A**, la podemos representar por la $\Sigma\gamma_i$; es decir, tenemos que:

$$\text{Traza}[S^{x}] = \Sigma\gamma_i = VAR[Y_1] + \ldots + VAR[Y_p] \quad (2.4.12)$$

por lo tanto, cada una de las expresiones siguientes son consecutivamente equivalentes:

$$\frac{\gamma_1}{\Sigma\gamma_i}, \ldots, \frac{\gamma_p}{\Sigma\gamma_i} \quad \text{con} \quad \frac{VAR[Y_1]}{\text{Var.total}}, \ldots, \frac{VAR[Y_p]}{\text{Var.total}} \quad (2.4.13)$$

y representan la contribución porcentual de la varianza de cada uno de los componentes con respecto a la variación total de las variables originales. En general:

$$\frac{VAR[Y_i]}{\Sigma var[Y_i]} = \frac{\gamma_i}{\Sigma\gamma_i} \quad (2.4.14)$$

representa la importancia relativa del i-ésimo componente principal en la descripción del sistema.

2.3.1.3 Correlación de variables con componentes.

Para encontrar la correlación entre un componente $-Y_1-$ y las variables $\mathbf{X}$, consideremos lo siguiente:

$\mathbf{X'Y}$: arroja los productos cruzados entre Y_1 y cada variable X.

$$X'Y_1 = X'XA_1 = \gamma_1 A_1 .$$

por tanto la correlación entre X_i y Y_1 es:

$$R_{i1} = \frac{COV[X_i, Y_1]}{[VAR[X_i]\ VAR[Y_1]]^{1/2}}$$

$$= \frac{[\gamma_1]^{1/2} a_{i1}}{[[X_{ij}]^2 \gamma_1]^{1/2}} \qquad (2.4.15)$$

$$= \frac{\gamma_1 a_{i1}}{[{}_j\Sigma^n [X_{ij}]^2]^{1/2}}$$

Donde: a_{i1} es el i-ésimo elemento en el vector A_1.

X_{ij} es la desviación respecto a la media.

En general:

$$Rij = \frac{a_{ij}[\gamma_j]^{1/2}}{[\Sigma[X_{ij}]^2]^{1/2}} \qquad (2.4.16)$$

y estos resultados son útiles en el proceso de interpretación de los componentes.

2.3.1.4 Los C.P. y la estandardización de variables.

En general puede decirse que si las variables están medidas en las mismas unidades los problemas de interpretación no son tan agudos en comparación a situaciones en que las unidades de medida son diferentes. En estos casos usualmente se estandardizan las variables antes de calcular los coeficientes a_{ij} -en el caso de estandardización los llamaremos b_{ij}- y el análisis e interpretación se realizan en términos de las puntuaciones estándard.

Definimos:

$$Z_{ij} = \frac{X_{ij}-X_{i.}}{S_i}$$

Si obteníamos la matriz S^2 para las variables medidas en las unidades iniciales, obtenemos en este caso la matriz R -matriz de correlaciones- para las puntuaciones Z. En estos términos el i-ésimo componente principal sería:

$$Y_{ij}= b_{i1}Z_{1j} + b_{i2}Z_{2j} +\ldots+ b_{ip}Z_{pp} \quad i=1..n.$$

que se puede escribir matricialmente como:

$$\mathbf{Y}_i = \mathbf{Z}\mathbf{B}_i \quad (2.4.17)$$

Donde $\mathbf{B}_i$: vector de coeficientes.

$\mathbf{Z}$: matriz de puntajes estándard.

$\mathbf{Y}_i$: vector de componentes.

Los B_i son los vectores característicos normalizados asociados con la i-ésima raiz caracteristica de la matriz de correlaciones. En este caso la traza de la matriz **R** es:

$$\text{TRAZA}[R] = \Sigma \gamma_i = p \quad (2.4.18)$$

y la importancia relativa de cada uno de los componentes se mediría por:

$$\frac{\gamma_i}{p} \quad (2.4.19)$$

2.3.2 DE LOS FACTORES.

2.3.2.1 El propósito central.

Consideremos al vector **X** en el cual el subíndice p nos indica el número de variables:

$$X = [X_1, X_2, \ldots, X_p]'$$

Supondremos que el vector **X** se distribuye normal multivariado con un vector de medias dado y una matriz de covarianzas S_{cov}. Las variables X_i las representamos como una combinación lineal de los γ factores comunes más un factor único, considerado el término residual:

$$X_i = \gamma_{i1} f_1 + \gamma_{i2} f_2 + \ldots + \gamma_{ik} f_k + e_i \quad (2.4.20)$$

Donde:

γ_{ij}: ponderación de la i-ésima variable en el j-ésimo factor.

f_j: j-ésimo factor común.

e_i: i-ésimo factor único específico.

2.3.2.2 Supuestos.

Adicionales al supuesto anterior, el M.A.F. implica una serie de incisos sin los cuales es imposible obtener resultados coherentes. Los supuestos son los siguientes:

$f_j \rightarrow N[0,1]$ (2.4.21)

$e_i \rightarrow N[0,\Phi]$ (2.4.22)

$COV[f_i,f_j] = 0$: los factores comunes son independientes. (2.4.23)

$COV[f_i,e_j] = 0$: los factores comunes y el factor único son independientes. (2.4.24)

$COV[e_i,e_j] = 0$: los factores únicos para cada una de las variables son independientes. (2.4.25)

La varianza para cada una de las X_i la encontramos:

$$VAR[X_i]=VAR[\gamma_{i1}f_1 +...+ \gamma_{ik}f_k + e_i]$$

$$= [\gamma_{i1}]^2 +...+ [\gamma_{ik}]^2 + \Phi_i$$

$$= \Sigma_j[\gamma_{ij}]^2 + \Phi_i \quad (2.4.26)$$

= varianza debida a factores comunes o **comunalidad** + varianza específica de la variable X_i

Es decir, la variabilidad de X_i se descompone en dos partes: una atribuida a los factores comunes y que la llamaremos h_i y otra que responde a la parte residual. La primera $-h_i-$ recibe el nombre de **comunalidad** y mide la información que la variable X tiene en común con otras variables en estudio a través de los factores comunes. Sirve para evaluar la magnitud de la varianza de la variable que es explicada por los factores. Como se ve, la $VAR[X_i]$ es una función lineal de las varianzas de los factores.

Si Φ_i es 'chica', la VAR[X_i] es absorbida por los factores, lo que significaría que el modelo propuesto es el adecuado para explicar el comportamiento de la información original. Por el contrario, si su magnitud es 'grande' quiere decir que la varianza de la variable se explica más por la de su factor específico que por la variabilidad del modelo de factores. En este último caso habría que descartar el A.F. como el apropiado para modelar el comportamiento de las X_i

Veamos a continuación la estructura de covariación de las X.

$$COV[X_i,X_j] = COV[[\gamma_{i1}f_1 +..+ \gamma_{ik}f_r + e_i],[\gamma_{j1}f_1 +..+ \gamma_{jk}f_k + e_j]]$$

$$= \gamma_{i1}\gamma_{j1} + \gamma_{i2}\gamma_{j2} +...+ \gamma_{ik}\gamma_{jk}$$

$$= {}_m\Sigma^k \gamma_{im}\gamma_{ik} \qquad (2.4.27)$$

De lo cual resulta que en la matriz de covarianzas de X, la diagonal principal está compuesta de las comunalidades $-h_i-$, y fuera de la diagonal cada uno de los elementos acabados de encontrar en (2.4.27). Las cargas de los factores están presentes en dicha matriz.

Observemos ahora la covariación entre la i-ésima variable y el j-ésimo factor:

$$COV[X_i,f_j] = COV[[\gamma_{i1}f_1 +...+ \gamma_{ik}f_k + e_i], f_j]$$

$$= COV[\gamma_{i1}f_1,f_j] +...+ COV[\gamma_{ij}f_j,f_j]+...$$

$$..+ COV[\gamma_{ij}f_k,f_j] + COV[e_i,f_j]$$

$$= \gamma_{ij} \qquad (2.4.28)$$

Encontramos que los coeficientes γ_{ij} son la covarianza de la variable con el respectivo factor. Si trabajamos con la matriz de correlaciones en lugar de la de covarianzas, el resultado en (2.4.28) es la correlación de la i-ésima variable con el j-ésimo factor. En la práctica, este último procedimiento es el acostumbrado. Registremos que los coeficientes obtenidos difieren según se utilice R o $S^{=}$.

2.3.2.3 Formulación matricial.

Si las variables originales las podemos representar por:

$$
\begin{aligned}
X_1 &= \gamma_{11} f_1 + \ldots + \gamma_{1k} f_k + e_1 \\
X_2 &= \gamma_{21} f_1 + \ldots + \gamma_{2k} f_k + e_2 \\
\cdot \quad & \quad \cdot \qquad \cdot \qquad \cdot \\
\cdot \quad & \quad \cdot \qquad \cdot \qquad \cdot \\
X_p &= \gamma_{p1} f_1 + \ldots + \gamma_{pk} f_k + e_p
\end{aligned}
$$

Entonces el modelo matricialmente quedaría así:

$$\mathbf{X} = \mathbf{\Gamma F} + \mathbf{E} \quad (2.4.29)$$

Donde:

- $\mathbf{X}$: vector (p x 1) de variables originales.
- $\mathbf{\Gamma}$: matriz (p x k) de cargas.
- $\mathbf{F}$: vector (k x 1) de factores comunes.
- $\mathbf{E}$: vector (p x 1) de factores especificos.

Por los resultados obtenidos de (2.4.21) a (2.4.28) tenemos entonces que:

$$\mathrm{VAR}[\mathbf{X}] = \mathrm{VAR}[\mathbf{\Gamma F}] + \mathrm{VAR}[\mathbf{E}] \quad \text{o sea}$$

$$\mathbf{S} = \mathbf{\Gamma\Gamma}' + \mathbf{\Phi} \quad (2.4.30)$$

Donde Φ es la matriz de covarianzas del vector de residuales en el cual la diagonal principal está compuesta por las varianzas de los factores especificos y ceros fuera de la diagonal.

Si estandardizamos las variables **X** podemos utilizar la matriz de correlaciones:

$$\mathbf{R}=\Gamma\Gamma'+\Phi \quad (2.4.31)$$

Con sus elementos conformados como aparece a continuación:

$$\begin{matrix} 1 & \cdots & r_{1j} & \cdots & r_{1p} \\ \vdots & & \vdots & & \vdots \\ r_{j1} & \cdots & 1 & \cdots & r_{jp} \\ \vdots & & \vdots & & \vdots \\ r_{p1} & \cdots & r_{pj} & \cdots & 1 \end{matrix}$$

2.3.2.4 Estimación de los factores por máxima verosimilitud.

Suponiendo distribución normal multivariada para la matriz **X**, se maximiza con respecto a Γ y a Φ la siguiente función[14]:

$$[-1/2]n \ \mathrm{Log.DET}[2Pi S]-[1/2n]\mathrm{traza}[S]^{-1}S_{cov}$$

Donde **S** es la matriz definida por (2.4.30) Se requieren [1/2]k(k-1) restricciones para romper la indeterminación y se

----

14. Ver entre otros: Mardia, p. 263 y ss.; Marriot, p. 46 y ss.; Giri, p. 296 y ss.

obtienen γ factores para hacer posible la estimación por máxima verosimilitud. El número de factores viene dado por :

$$k < p - 1/2\ [[8p + 1]^{1/2} - 1]$$

Si $S^{=}$ es la matriz de dispersión muestral, entonces

$$S^{=} = \Gamma^{=}\Gamma'^{=} + \Phi^{=} \dashrightarrow$$

por lo tanto

$$\frac{-n \quad \mathrm{Ln\ DET}[S_{cov}]}{\mathrm{DET}[S^{=}]} - \mathrm{traza}[S_{cov}S^{=-1}] + P \quad .$$

se aproxima a una distribución X^2 con

$1/2\ [[n - k]^2 - [p+k]$ grados de libertad.

Las pruebas de hipótesis se realizan de acuerdo a lo acabado de mencionar.

Capitulo 3.

ALGUNOS TRABAJOS DE INVESTIGACION.

Nos interesa en el presente capitulo comentar criticamente algunos trabajos que han apelado al M.A.F. como técnica estadística para abordar estudios socio-económicos. Resaltamos en cada uno de ellos diferentes aspectos de acuerdo al contenido mismo de las investigaciones. Pero antes de continuar queremos reconocer el importante esfuerzo teórico, estadístico, analitico e interpretativo que suponen estas investigaciones. Detrás de ellas hay un trabajo enorme cuyos aportes no se demeritan por las criticas que aquí se efectúan. Son investigaciones importantes que en gran medida han contribuido al proceso exploratorio de la realidad y a su discusión científica dentro de sus respectivos campos de aplicación.

3.1 LA MORTALIDAD Y LOS FACTORES ECONOMICOS.

En una interesante investigación cuyos resultados se publican en un artículo llamado "La Mortalidad y su relación con Factores Sociales, Económicos y Culturales."[1], se utiliza el A.F. para encontrar los factores comunes a cinco grupos de variables [2] estudiadas en los municipios de México con el objeto, después de hacer una primera interpretación de los resultados del A.F., de utilizar los índices obtenidos en el A.F. para correr regresiones de las variables de mortalidad con los grupos restantes, estudiar su comportamiento y, posteriormente, encontrar regionalizaciones del país.

El trabajo está dividido en tres partes. En la primera se describen fuentes y confiabilidad de la información, se presentan y justifican las variables, se efectúan análisis de distribuciones de frecuencias de las variables a nivel municipal y se hacen sugerencias respecto a la captación de las estadísticas en México. En la segunda parte, a través del M.A.F. se obtienen los índices -factores comunes- para cada grupo de variables. A partir de esto último, se configuran modelos de regresión tomando a los índices de mortalidad como variables dependientes y a las demás como independientes para cinco estratos de municipios -muy rural, rural, semiurbano, urbano y muy urbano-. Terminan describiendo y comentando algunas características que se presentan en la información como posibles responsables de los resultados poco alentadores de las regresiones para los diferentes estratos. En la última parte se construyeron tres índices para cada tipo de variable buscando clasificar diferentes regiones de acuerdo a los criterios de selección de los cinco grupos definidos.

----

1. Cañedo et al., p. 805 - 849. Cfr. bibliografía.

2. Los grupos de variables seleccionados fueron los siguientes: mortalidad -seis variables-, atención a la salud -siete-, condiciones de vivienda -cinco-, económicas -siete- y, por último, educativas -nueve variables-.

Nuestro interés en el documento en mención se circunscribe al tratamiento e interpretación de los resultados encontrados en lo que respecta al subgrupo de variables económicas. La justificación de la inclusión de estas últimas se plantea en el material reseñado en los siguientes términos:

> "Se ha establecido clara e indudablemente que la diferencia de clases produce una diferencia inherente en cuanto a los niveles de mortalidad, los cuales son mayores en cuanto menor es la capacidad económica de la clase social de que se trate... En México se ha señalado la necesidad de realizar investigaciones en salud pública en las que se incluyan los elementos sociales y económicos, e independientemente se ha descrito la incidencia de los factores sociales sobre la mortalidad de dos poblaciones con diferente grado de desarrollo socioeconómico. También se ha reconocido que los factores socioculturales y económicos son parcialmente responsables de la mortalidad y se ha enfatizado la diferencia de patologia que existe entre diversas clases sociales mexicanas."[3]

Con estos antecedentes teóricos se definieron aquellas variables que se espera que directa o indirectamente estén más asociadas con los problemas de salud. Se escogieron los porcentajes de la P.E.A. en la agricultura, en la industria, en servicios y en sectores no especificados con respecto a la P.E.A. total. Se incluyen además, las productividades industrial y agropecuaria y el ingreso per cápita. Se supone que son datos de 1970.[4] En cuanto a las primeras cuatro variables los investigadores advierten su complementariedad y por eso reconocen correlación espuria, no obstante consideran en la investigación que no había información que pudiese reemplazar estos rubros.[5] Adicionalmente, no se contempló el D.F. en los resultados finales.

----

3. Cañedo, p. 806.

4. Como fuentes de información los autores mencionan el IX Censo General de Población y la Comisión de Salarios Mínimos.

5. Cañedo, p. 813.

Después de estas aclaraciones y de un análisis de la información de las variables no económicas el equipo de investigadores se planteó como objetivos:

"a) construir índices que permitan reducir el número de variables, reteniendo lo que cada una de ellas aporta a la explicación del fenómeno

b) encontrar aquellos factores socioeconómicos y culturales que están asociados con las diferentes causas de la mortalidad, dependiendo de sus grados de ruralidad y urbanismo

c) detectar algunas fallas en el proceso de captación de la información, comparando las relaciones obtenidas con las que se espera obtener en las diferentes regiones del país."[6]

Mediante el SPSS [7] se obtuvieron las ponderaciones de los factores rotados, resultados que se muestran en el cuadro No. 3.1. Sólo reproducimos lo referente al conjunto de variables económicas, pues, como ya lo habíamos anotado, nos interesa la particular interpretación que se hace de los factores económicos.

----

6. O.C., p. 822.

7. El programa SPSS usa el procedimiento del factor principal para la estimación de las cargas de los factores.

C U A D R O No. 3.1 [8]

COEFICIENTES DE LOS FACTORES ROTADOS.

VARIABLES	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
% P.E.A. Agrop.	-.96009 *	-.01633	-.21336	-.02855
% P.E.A. Ind.	.84882 *	.04806	-.04999	.06352
% P.E.A. Servs.	.88024 *	.07624	.09668	-.00523
% P.E.A. no espec.	.10987	.01440	.98571 *	-.02865
Prodtvdd. Ind.	.10039	.98147 *	.01680	.00536
Prodtvdd. Agrop.	.00996	.00359	-.02768	.99650*
Ingreso per cáp.	.67136 *	.20443	-.08016	-.06501

*: coeficientes usados en la interpretación del factor.

De las siete variables iniciales de naturaleza económica llegamos a sólo cuatro factores -los "factores económicos" a los que responde el título de la investigación- que explican el 84.9 % de ellos. Veamos el significado que en la investigación se le asigna a cada uno de los factores:

> "Factor 1: este factor es bipolar y representa un contraste entre la variable P.E.A. en el sector aagropecuario [9] -con coeficiente de signo negativo-, y las variables P.E.A. en servicios, P.E.A. en industria e ingreso per-cápita. Se le interpretó como un factor del **grado de actividad industrial**. Explica el 43 % de la variación de los datos."[10]

----

8. Fuente: Cañedo, p. 825.

9. Queremos hacer hincapié que en el material reseñado se utilizan alternativamente descripciones diferentes para las variables económicas. Aquí la referencia es a la participación de la P.E.A. de la agricultura con respecto al total, y no a la P.E.A. en la agricultura como dice textualmente la cita. Igual comentario vale para la P.E.A. en industria, servicios y actividad económica no especificada.

10. O.C., p. 823. De no indicar lo contrario, hemos enfatizado.

Cabría otra interpretación para este factor ? En la explicación anterior se presume que el ingreso per-cápita es parte integrante y fundamental del "grado de actividad industrial", con lo cual, tomado así no más, muchos analistas no estarían de acuerdo aduciendo que es una variable explicativa demasidado burda: es un promedio de los ingresos en una distribución muy sesgada de considerable asimetría positiva y muy sensible a los valores extremos, por lo cual no hay mucha seguridad de que en un corte transversal como el que se realiza tenga un buen contenido indicador. Cuestionada esta variable en la composición del primer factor, podríamos interpretarlo como la distribución de la P.E.A. por actividad económica y sería una interpretación que nada tiene que envidiar a la presunción dada de "grado de actividad industrial". Lo que se quiere remarcar es el que pueden coexistir interpretaciones diferentes de un mismo factor, de acuerdo a los objetivos del análisis y a las cualidades interpretativas e intereses del investigador. A modo de moraleja: el factor subyacente, del que se esperaba un papel esclarecedor, no tiene un sentido único, desafortunadamente.

Veamos la interpretación de los otros factores:

> "Factor 2. La variable de valor agregado en la industria por persona activa en este sector es la única con coeficiente cercano a uno. Se interpretó como un factor de productividad industrial, representando el 14.8 % de la variación de los datos.
>
> Factor 3. Este es un factor de actividad económica no definida por ser la variable de P.E.A. la única con coeficiente significativo para la interpretación. Este factor explica el 14.2 % de la variación de los datos.
>
> Factor 4. La variable de valor agregado del sector agropecuario por persona activa en este sector es la única con peso significativo para la interpretación, por esta razón se le consideró como el factor de

productividad agropecuaria. Explica el 13 % de la variación de los datos."[11]

Cuáles fueron los logros obtenidos ? En la interpretación de los tres últimos factores se acudió exclusivamente a las mismas variables iniciales, con una agravante adicional en el caso del factor tres: el expresarlo como "actividad económica no definida" introduce mayor confusión respecto a su definición inicial de proporción de la P.E.A. en sectores económicos no especificados. Desde luego, la actividad económica no especificada puede ofrecer un gran contenido analítico de acuerdo al desglose que hagamos de ella. Pero los autores pretenden dotar a variable de un contenido muy particular: la trastocan, primero en factor y luego en factor de actividad económica no definida, y al hacerlo, permítase el juego de palabras, definen a un factor por su indefinición.

Conclusión: al buscar los factores latentes del conjunto de la información, el A.F. en la mejor de las situaciones dejó intactas a las variables iniciales -los factores dos y cuatro-, y en el peor rebautizó no muy venturosamente a otra -es el caso del factor uno- y despojó de nombre a la restante -el factor tres-. No es tarea fácil interpretar factores en un modelo explicativo del comportamiento tan direccionado como lo es el M.A.F.: es un verdadero duende de la estadística que hace jugarretas que nos dejan mal parados en un análisis teórico.

----

11. O.C., p. 823.

3.2 LOS FACTORES EN LA ESTRUCTURA AGRARIA: VERSION I.

Otro interesante estudio titulado "Construcción de zonas agrícolas para una análisis de la estructura agraria en México: una aplicación del análisis factorial."[12] publicado en 1977 se plantea:

> "Este trabajo tiene como objetivo presentar una clasificación de los municipios según el tipo de agricultura que predomina en cada municipio, con el fin de estudiar las regiones agrícolas de México."[13]

El estudio considera como unidad de análisis el municipio y con información del Censo Agrícola, Ganadero y Ejidal de 1960 se construyen ciertos indicadores con los cuales se proponen:

> "...llegar a una clasificación de los municipios que permita apreciar zonas diferenciadas que reflejen el desarrollo desigual de la actividad agrícola."[14]

Los indicadores utilizados para "captar los rasgos generales de la estructura agraria" [15] son de cinco tipos que en total dan lugar a 24 variables -que se presentan en el cuadro 3.2-. Los cinco grupos son:

----

12. Appendini et al., (1977). Cfr. bibliografía.

13. O.C., p. 6.

14. O.C., p. 12.

15. O.C., p. 13.

1. Uso de tecnología y disponibilidad de capital.
2. Estructura de la fuerza de trabajo agrícola.
3. Comercialización.
4. Tamaño económico del predio y tipo de predio que predomina en el municipio.
5. Otras variables, particularmente referentes a productividad y a otros recursos y su aprovechamiento.

En las observaciones iniciales del material se recogen algunas críticas de las que habían sido objeto, relacionadas, entre otras, con el uso indiscriminado de los conceptos de agricultura capitalista y agricultura moderna, al manejo de la categoría polarización de la agricultura -en dos tipos, la moderna y la campesina-, a la utilización del predio y no a la superficie de labor como unidad de estudio y a la utilización del M.A.F. Respecto a este último el documento afirma:

> "Es de gran interés también la observación que se refiere a que la técnica del análisis factorial no es adecuada para el estudio de procesos, de lo que ya estábamos advertidas. Sería interesante enfatizar que si el análisis factorial parte de una observación puntual en el tiempo, difícilmente podría ser aplicado como instrumento para el estudio de un proceso o de la dinámica de un fenómeno social; esto se debe a que los diferentes aspectos de un fenómeno tienen ritmos desiguales y específicos de desarrollo, por esto un corte puntualizado en el tiempo no necesariamente rendirá cuenta de la relación compleja entre los diferentes aspectos del fenómeno en examen."[16]

Después de hacer algunas consideraciones generales sobre la agricultura mexicana y de presentar los indicadores y las variables, se pasa en el trabajo a la justificación y descripción de la técnica estadística utilizada. A pesar de lo extenso de la cita, consideramos de importancia transcribirla para entender los propósitos de las investigadoras:

> "Para hacer la caracterización del tipo de agricultura que predomina en el municipio con base en el nivel cuantitativo que presenten los indicadores seleccionados, es posible elegir entre varios métodos

----

16. O.C., p. 4.

> estadísticos que pueden conducir al mismo resultado, con la diferencia de que con algunos de éstos se obtendrá de manera más directa. En un trabajo anterior se analizaron, uno a uno, los indicadores que desde el punto de vista teórico se consideraron más importantes. Este análisis permitió construir una escala para clasificar a los municipios en cada indicador, pero no resultó adecuado para caracterizar el tipo de agricultura predominante en cada municipio debido a que este concepto está definido por la interrelación entre los indicadores ya que considerándolos simultáneamente con respecto al conjunto de municipios permiten distinguir los diferentes tipos de agricultura. Esta limitación llevó a emplear un método que permite estudiar las interrelaciones empíricas entre los indicadores, destacando a través de las regularidades existentes en los datos, grupos de indicadores **cuyas pautas de relaciones** con los demás sean muy parecidas y puedan ser tratados como **factores**; de ahí que el nombre de esta técnica sea 'análisis factorial'."[17].

Reforcemos el planteamiento anterior con lo que se afirma algunas páginas más adelante:

> "Es muy importante hacer hincapié en que la técnica utilizada, en este caso el análisis factorial, debe servir para proporcionar evidencias empíricas que apoyen algunas de las proposiciones o hipótesis del trabajo, pero nunca se ha pensado derivar o inferir de los resultados del método estadístico la construcción de la teoria; el procedimiento de investigación debe ser exactamente al contrario, esto es, partiendo de una postura teórico-metodológica apoyar las principales hipótesis a través de técnicas específicas aplicables a la problemática."[18]

----

17. O.C., p. 18 - 19.

18. O.C., p. 24 - 25.

Y a esta altura hay una nota de pie de página que es imprescindible de conocer:

> "Es bastante común utilizar este método como técnica exploratoria (para describir pautas) pero esta utilización en ausencia de una teoría es arriesgada y hasta peligrosa.."[19]

Un comentario especial ameritan las citas anteriores. El rasgo principal que se quiere subrayar es la ausencia de una consideración explícita sobre los objetivos que se propone en sí mismo el M.A.F. El análisis de factores intenta teorizar un tipo de comportamiento en el cual se considera que las variables en estudio son susceptibles de reconstruirse de una manera tal que permitan descubrir las principales y diferentes causas comunes que le dan razón de ser y que están implícitas en la información cuantitativa: se trata de encontrar las pautas básicas inmersas en los datos iniciales.

Si retrocedemos a los orígenes del M.A.F.,la posición a este respecto es indiscutible. Los cambios que experimentó la formulación misma del M.A.F. se refieren básicamente a las discusiones de si la explicación de una serie de datos se debe a si hay un factor único subyacente o, alternativamente, si se debe a un conjunto de factores inmanentes -los factores comunes-. En el fondo la incapacidad de Spearman y seguidores de encontrar relaciones causales en los asuntos en que se interesaban los llevaron a ingeniarse un modelo que respondiera a este interrogante: Qué había detrás de la información cuantitativa que mostraba correlaciones altas entre las variables que manejaban ? Como respuesta, la manipulación matemática inventó los factores implícitos.

----

19. O.C., p. 25.

Tal vez este doble descuido en entender sobretodo el propósito y un poco menos, la coherencia lógica del M.A.F., se presta a que sea una práctica común "derivar o inferir los resultados del método estadístico la construcción de la teoría y utilizarla como técnica exploratoria para describir pautas". La cuestión no es simplemente el que haya o no una teoría en la que se inserte el M.A.F. El problema es que el M.A.F. es una teoría del comportamiento de las variables que ambiciona otorgar y acuñar cientificidad -a despecho de otra u otras teorías aledañas- a partir de la especificación de un tipo de función matemática que asocia a las variables con los factores. De ahí que sea cuestionable afirmar que:

> "Los supuestos que este modelo requiere son mínimos en tanto **NO** se pretende emplearlo para pruebas de hipótesis estadísticas **o para establecer el tipo de función matemática** que relaciona a los indicadores."[20]

Es cuestionable no sólo por los exigentes supuestos, que de no acatarse arrojan resultados espurios, como ya se mencionó en sección anterior, sino fundamentalmente porque el modelo es **contundente** al establecer un tipo de función matemática que es la que le da sentido al patrón explicativo.[21]

Por su propia lógica matemática el M.A.F. arroja unas magnitudes, unos resultados, el que éstos tengan la connotación de las pautas implicitas es dejarse seducir por la pretensión de descubrir factores subyacentes, es otorgarle a una sofisticada técnica estadístico-matemática la facultad de encontrar los susbstratos escondidos e implícitos de la realidad. Por su proceso de construcción el modelo destacará 'algo', el ambicionar

----

20. O.C., p. 19.

21. Estrictamente hablando la función matemática es del tipo dado por la expresión 2.2.2.

que ese 'algo' sea un factor explicativo es un riesgo que corren quienes creen en él y por ello es responsabilidad del investigador dejarse arrastrar por el espejismo de detectar esos substratos por una especial operación matemática.

Es exactamente en este ambiente en el que los esfuerzos interpretativos abren las puertas a los cuestionables intentos de generar teorías tomando como base unas relaciones cuantitativas particulares -sean éstas cuales fueren y a pesar del buen estudio de definición y selección de variables o indicadores-; parecería que el modelo no consintiera escapatoria, es algo así como que los ineluctables designios del M.A.F. impregnaran a la investigación -y a los sujetos de ella- con una fatalidad insoslayable.

En el material en referencia hay prácticamente un cierto reconocimiento -desde luego no mal intencionado- de que problemas como los acabados de expresar, por decirlo de alguna manera, contaminan el espíritu investigativo. Se dice en el documento que se "caracteriza el tipo de agricultura **con base en el nivel cuantitativo** que presenten los indicadores seleccionados", según se afirma en una de las citas presentadas anteriormente. Más todavía, en el trabajo se critica la utilización del método como "técnica exploratoria para **describir** pautas" y, sin embargo, en la justificación de la técnica estadística se reconoce que se quieren destacar "grupos de indicadores cuyas **pautas de relaciones** con los demás sean muy parecidas y puedan ser tratados como 'factores"; conclusión: el resbalón es inevitable.

Ya para finiquitar este comentario, se quiere resaltar que en el material criticado se acepta finalmente esa búsqueda de los factores subyacentes o implícitos. Transcribimos inmediatamente la parte que se refiere al procedimiento de los factores principales para encontrar los factores:

> "Este método (que se utiliza en este trabajo) de las variables se puede explicar por la relación de cada variable con m factores **implícitos** ($m < p$) entre los cuales se reparte la varianza y por tanto la parte de varianza no explicada por estos factores no contribuye a la intercorrelación entre las variables."[22]

----

22. O.C., P. 60.

Un segundo comentario que queremos desarrollar se asocia con los resultados y las conclusiones de la investigación reseñada. En el cuadro No. 3.3 presentamos los coeficientes de los factores rotados y en el anterior -cuadro No. 3.2, página 121- el significado de cada una de las variables. De los cinco grupos de indicadores que generan las 24 variables, se obtienen en el análisis siete factores, de los cuales sólo dos -los factores uno y tres- entran en la construcción de indices para conformar las zonas agrícolas. El nombre dado a cada uno de los factores es el siguiente: [23]

Factor 1: Dotación de capital por predio y uso de tecnologia.
(V_{14}, V_{16}, V_{12}, V_{17} y V_{19}.)

Factor 2: Relación entre producción, tierra y trabajo.
(V_{20}, V_{2}, V_{24} y V_{10}.)

Factor 3: Polarización por el tipo de agricultura dominante.
(V_{21}, V_{22}, V_{6}, V_{4} y V_{9}.)

Factor 4: Disponibilidad de capital por superficie de labor.
(V_{11}, V_{13} y V_{15}.)

Factor 5: Productividad y producción comercial.
(V_{7}, V_{8}, V_{1} y V_{3}.)

Factor 6: Importancia de la producción ejidal.
(V_{5}, V_{6} y V_{4}.)

Factor 7: Aprovechamiento de la tierra.
(V_{23} y V_{18}.)

----

23. Incluimos entre paréntesis las variables que le dan sentido interpretativo a cada factor de acuerdo al valor de las correlaciones ordenadas de mayor a menor.

Según el M.A.F. los coeficientes del cuadro No. 3.3 representan las correlaciones de las variables con los factores; uno de los problemas más comunes es la determinación de la magnitud -en valor absoluto- que debe tener el coeficiente para que la variable examinada pueda formar parte interpretativa de ese factor. Realmente no hay criterios definidos de tipo estadístico que puedan ser útiles para la mencionada clasificación, pues no hay una definición única que nos permita afirmar que una correlación es alta, por esta razón se deja a criterio del investigador la selección de las variables en la constitución del factor.

Se supone, adicionalmente, que el valor del coeficiente que se considera 'alto' para fines interpretativos en cada factor debe ser definido de antemano, sin conocer todavía las cargas de los factores. En la investigación que estamos discutiendo se integraron las variables en la conformación de cada factor si las cargas daban valores alrededor de 0.5 en valor absoluto. Si estuviéramos en un análisis de regresión una correlación como la última entre la variable explicada y las variables independientes, significaría que sólo el 25 % de la regresión estaría explicada por el modelo. Desde luego este criterio no tiene por qué aplicarse al M.A.F. a pesar de que en el procedimiento de los factores principales, las estimaciones iniciales de las cargas se hacen con los más altos coeficientes de regresión múltiple entre cada una de las variables con respecto a todas las demás. Hay investigaciones más 'rigurosas' que se quedan con las variables cuyas correlaciones con los factores sean al menos de 0.7. Desde luego, la 'rigurosidad' es relativa, pues de exagerarla nos quedaríamos con muy pocas variables y factores. Pues bien, si el criterio de selección fuese el 0.7 mencionado, perderíamos una variable en los factores uno, dos, seis y siete; para el factor tres -que junto con el factor uno son trascendentales para la zonificación agrícola- la pérdida sería de cuatro de las cinco variables. A los factores cinco y seis se les restarían dos variables a cada uno y sólo el factor cuatro se quedaría con el grupo asignado inicialmente. Naturalmente, ante estos cambios, la interpretación y el sentido de los factores podría alterarse, aspecto en el cual ya hemos enfatizado en el ejemplo anterior.

C U A D R O No. 3.2.

NOMBRES DE LAS VARIABLES.

V_1 : Valor de la producción agrícola total/superficie de labor. VPAT/SL.
V_2 : VPAT / número de predios.
V_3 : VPAT / personas ocupadas.
V_4 : Valor de la producción agrícola de predios privados mayores que producen más de $25.000.00 / VPAT.
V_5 : Valor de la producción agrícola ejidal / VPTA.
V_6 : Valor de la producción agrícola de predios privados que producen hasta $25.000.00 / VPAT.
V_7 : VPAT menos valor de la producción de maíz y frijol/ VPAT.
V_8 : Superficie cosechada total menos superficie cosechada de maíz y frijol / superficie cosechada total.
V_9 : Valor de las ventas agrícolas / VPAT.
V_{10} : Productor y familiares / predio.
V_{11} : Capital agrícola / superficie de labor.
V_{12} : Capital agrícola / predio.
V_{13} : Vlr. maquinaria / superficie de labor.
V_{14} : Vlr. maquinaria / predio.
V_{15} : Número de tractores / superficie de labor.
V_{16} : Número de tractores / predio.
V_{17} : Superficie con tracción mecánica / sup. labor.
V_{18} : Superficie fertilizada / superficie de labor.
V_{19} : Superficie de riego / superficie de labor.
V_{20} : Personas ocupadas / predio.
V_{21} : Personas ocupadas / superficie de labor.
V_{22} : Personas asalariadas / personas ocupadas.
V_{23} : Superficie cosechada / superficie de labor.
V_{24} : Superficie de labor / predio.

C U A D R O No. 3.3

COEFICIENTES DE LOS FACTORES ROTADOS. [24]

	Fac.1	Fac.2	Fac.3	Fac.4	Fac.5	Fac.6	Fac.7
V_1	-.060	.088	.103	.141	.679*	.124	.501
V_2	-.380	.754*	-.210	-.089	.265	.029	.054
V_3	-.318	.185	-.332	-.129	.494*	-.005	.024
V_4	-.166	.349	-.582*	.059	.226	.494*	.131
V_5	.131	-.124	-.075	-.025	.070	-.956*	.002
V_6	.033	-.198	.617*	-.019	-.281	.555*	-.125
V_7	-.142	-.033	-.200	.059	.888*	-.039	-.142
V_8	-.357	-.069	-.090	.044	.710*	-.217	-.130
V_9	.127	.076	-.564*	.193	.382	.177	-.057
V_{10}	.187	.576*	.514	.124	-.059	.145	.166
V_{11}	-.230	-.012	.075	.910*	.028	.034	-.008
V_{12}	-.812*	.178	-.108	.282	.094	.957?	-.033
V_{13}	-.311	-.002	.024	.909*	.026	-.013	.041
V_{14}	-.885*	.196	-.120	.249	.111	.035	.051
V_{15}	-.550	-.008	-.050	.718*	.122	.029	.184
V_{16}	-.877*	.149	-.117	.120	.137	.060	.106
V_{17}	-.774*	.082	-.120	.321	.135	.116	.126
V_{18}	-.124	-.044	-.115	.013	.022	-.016	.683*
V_{19}	-.550*	-.032	.096	.084	.308	-.083	.336
V_{20}	-.145	.929*	-.011	.071	-.053	.083	.054
V_{21}	.133	.017	.732*	.177	.034	.198	.142
V_{22}	-.395	.405	-.645*	.067	.081	.041	.005
V_{23}	-.125	.025	.373	.068	-.214	.003	.724*
V_{24}	-.149	.733*	-.281	-.099	-.042	-.016	-.268

* Variables significativas en ese factor.
? Por indicación del texto parece que en el original hay un error en la magnitud de este coeficiente. Se supone que es muchísimo menor que .5.

---

24. Fuente: O.C., P. 77.

Así, no hay una regla definitiva para la inclusión de variables en la definición de un factor, de ahí los delicados problemas de interpretación: hay una relativa elasticidad en la composición de cada uno de los factores y por lo tanto en su interpretación teórica, pero no se pone en duda la existencia de los factores por sí mismos.

Esta situación nos crea conflictos; si despreciamos algunas variables por considerarlas irrelevantes y a su vez incluimos otras, al correr de nuevo el M.A.F. los resultados no tienen por qué coincidir y aparecerían nuevos factores y por ende nuevas interpretaciones de ellos. Esto se reconoce en el documento:

> "Cualquier adición o eliminación de indicadores requeriría una nueva aplicación del método y la composición de los nuevos factores puede ser diferente de la obtenida."[25]

Por estos motivos podemos concluir que -y en el mejor de los casos- los resultados de la investigación comentada y la interpretación que lo acompaña es sólo una entre las posibles y cuya base de sustentación reúne dos niveles de análisis: uno teórico -las consideraciones sobre la agricultura- y uno empírico que son los factores y las cargas del modelo. Pero este último aspecto es una arena movediza la cual depende -adicional a todo lo que se ha comentado ya- de la definición y desglose de variables. Cualquier cambio en éstos moverá el andamio empírico en el que se apoya el análisis teórico. El M.A.F. es demasiado relativista para concebirlo como corroboración estadística. Veamos cual es una de las conclusiones a las que llega el trabajo del que nos hemos ocupado:

----

25. O.C., p. 37.

"La dotación y utilización de capital es el elemento con mayor explicación en la diferenciación dentro de la agricultura, seguido por la asociación entre el monto de producción y los elementos tierra y trabajo y las características atribuibles a la organización de la agricultura en el municipio ya sea capitalista o campesina, de manera predominante."[26]

Estas conclusiones fueron las premisas teóricas con las que se partió para el estudio de la regionalización agrícola. Parecería que fue un exagerado trabajo al que se le dedicaron tantos esfuerzos para concluir con lo que se inició, con lo que ya se sabía: los factores económicos en la agricultura son capital, tierra y trabajo; y esto se debe a que a nivel de la interpretación sólo se puede acudir a los elementos teóricos que ya se tienen; es decir hay que acomodar los datos a una explicación, se trata de jerarquizar magnitudes y darles una posible interpretación, lo que ayuda, en el mejor de los casos, a agudizar la imaginación. Qué es lo que se hace en la práctica con el modelo de factores ? Sencillamente se arranca con unos datos inciales, se encuentran unos factores y unas cargas y, la labor final es el acomodarle una explicación coherente a ellos. Esta no es precisamente una vía de desarrollo de categorías, es un camino que lleva a gestar ciertas explicaciones partiendo de lo que tienen en común las variables con mayores cargas para la definición y especificación de un factor y ajustarle una interpretación acorde a la naturaleza de los datos, la cual sería menos desatinada entre más se acuda a la teoría existente para darles coherencia y armonía. De ahí que se llegue a afirmar -que no confirmar- lo ya conocido.

----

26. O.C., p. 36.

2.3 LOS FACTORES Y LA ESTRUCTURA AGRARIA: VERSION II.

Algunos años después de la publicación del artículo criticado, apareció otro de la misma autora, el cual formaba parte de "la investigación sobre estructura agraria en México que se realizó en el Centro de Estudios Económicos y el Centro de Estudios Sociológicos de El Colegio de México."[27]

Pasó, pues, poco más de un lustro para rehacer el mismo trabajo, pero teniendo como mira un año diferente. El material se tituló "La polarización de la agricultura mexicana: un análisis a nivel de zonas agrícolas en 1970.". Lo interesante de este trabajo, que desde luego adolece de los mismos defectos del anterior, radica en la aplicación del M.A.F. en el estudio de la estructura agraria mexicana no sólo para el año mencionado, sino que se modelan datos también para 1960. He aquí una gran oportunidad: la misma autora, el mismo periodo (1960), la misma fuente (censos) y el mismo procedimiento, pero separados por seis años de experiencias. Es realmente una situación muy venturosa.

Después de hacer una introducción a la temática agraria mexicana y de repetir y enfatizar en el trabajo del 83 algunas citas que ya mencionamos previamente, aplican el análisis factorial utilizando la información de los censos agrícolas, ganaderos y ejidales de 1960 y 1970 con el propósito, para este último año, de obtener los factores que sirvan en la zonificación deseada. En la investigación del 83, antes de recomendar la lectura del apéndice del trabajo del 77, por la explicación que allí se encuentra del modelo de análisis factorial, plantea la

----

27. Appendini (1983). Cfr. Bibliografía.

autora lo siguiente:

> "Para hacer la caracterización del tipo de agricultura que predomina en el municipio con base en el nivel cuantitativo que representan los indicadores seleccionados se puede elegir entre varios métodos estadísticos. Para fines de este trabajo se buscó un método que permitiera estudiar las interrelaciones empíricas entre los indicadores, destacando a través de las regularidades existentes en los datos, grupos de indicadores cuyas pautas de relaciones con los demás sean muy parecidas de manera que puedan resumir el fenómeno a explicar; para ello se consideró que el análisis factorial es un método adecuado, ya que cada grupo de variables interrelacionadas se asocia a una sola expresión llamada factor."[28]

Esta cita, la recomendación a leer el apendice metodológico del 77 y otras aseveraciones, manifiestan que desde el punto de vista metodológico no existe absolutamente ninguna diferencia en los dos estudios. No obstante, hay algunas variables diferentes y, lo principal, resultados, factores e interpretación, también diferentes. Los factores que en la investigación realizada en 1977 caracterizaban a la estructura agraria del sesenta, van a diferir de los factores determinantes encontrados seis años después para los mismos sesenta.

Sin más preámbulos presentemos una comparación de los resultados de los dos trabajos que tuvieron en común procedimiento y objeto (1960). En el bloque de la izquierda se aprecian los resultados encontrados por la investigadora en 1977. En el bloque de la derecha se presenta lo obtenido seis años después. La proporción que aparece entre paréntesis en cada factor es la parte de la varianza que ese factor explica. Las cifras entre paréntesis al frente de cada variable son las respectivas cargas del factor en las variables -las correlaciones de las variables con el factor- multiplicadas por cien y aproximadas a dos enteros.

----

28. Appendini (1983), p. 186. Compárese esta cita con la nota 17 de este capitulo.

CUADRO No. 3.4.

ANALISIS FACTORIAL DE LA ESTRUCTURA AGRARIA DE MEXICO. 1960

Análisis del 77.	Análisis del 83.
Factor 1 (17.95%)	**Factor 1** (40.6%)
DOTACION DE CAPITAL POR PREDIO Y USO DE TECNOLOGIA.	**DOTACION DE CAPITAL AGRICOLA.**
-Vlr. maquinaria por predio (-88)	-Vlr. Cap. agric./LABT. (95)
-Nro. tractores por pred. (-88)	-Vlr. maquinaria/LABT. (91)
-Vlr. cap. agric. por pred. (81)	-Superficie tracción mecánica/LABT. (40)
-Superficie mecanizada/LABT (-77)	
-Superficie de riego/LABT (-55)	
Factor 2 (11.65 %)	**Factor 2** (26.2 %)
RELACION ENTRE PRODUCCION TIERRA Y TRABAJO.	**MODERNIZACION AGRICOLA.**
-Personas ocupadas por predio(93)	-Superficie riego / LABT (68)
-Vlr. Pdn. agric. por pred. (75)	-Gastos insum. agric/LABT(62)
-Superf. de labor por pred. (73)	-VPAT / LABT (58)
-Productor y familiares/pred.(58)	-Suprf. sembrada en cultivos distintos a maiz y frijol / LABT. (54)
	-Suprf. tracción mecánica / LABT (49)

Factor 3 (11.57 %)

POLARIZACION POR EL TIPO DE AGRICULTURA PREDOMINANTE.
(campesino - capitalista.)

- -Personas ocupadas / LABT (73)
- -Vlr. Pdn. agric. que proviene de predios privados que producen menos de $25.000/ VPAT.(61)
- -Asalariados por personas ocupadas (-64)
- -VPAT que proviene de predios privados que producen más de $25.000 / VAPT. (-58)
- -Vlr. ventas / VPAT. (-56)

Factor 3 (16.4 %)

IMPORTANCIA DE LA PRODUCCION PECUARIA.

- -Vlr. Pdn. pecuaria /Vlr.tot. de la pdn. agropecuaria (47)
- -Suprf.labor/suprf.tot. (-57)
- -Número de personas ocupadas / LABT. (-41)

Factor 4 (10.83 %)

DISPONIBILIDAD DE CAPITAL POR SUPERFICIE DE LABOR.

- -Vlr. capital agric/LABT (91)
- -Vlr. maquinaria / LABT (91)
- -Nro. de tractores/LABT (71)

Factor 4 (9.6 %)

AGRICULTURA CAPITALISTA.

- -VPAT que proviene de predios que producen $100.000 y más anuales / VAPT. (74)
- -Nro. de asalariados / total de personas ocupadas (62)
- -VPAT / personas ocupadas. (44)

Factor 5 (10.75 %)

PRODUCTIVIDAD Y PRODUCCION COMERCIAL.

- -VPAT menos valor pdn. de maiz y frijol / VPAT (89)
- -Suprf. cosechada tot. menos sprf. cosechada con maiz y frijol/sprf.cosechada total(71)
- -V.P.A.T / LABT (67)
- -VPAT por persona ocupada (49)

Factor 5 (7.1 %)

APROVECHAMIENTO DE LA TIERRA CAMPESINA.

- -Sprf. cosechada / LABT (70)
- -Sprf. con cultivos distintos a maiz y frijol/superficie cultivada. (-45)
- -Vlr.ventas agric./VPAT (-39)

Factor 6 (6.93 %)

IMPORTANCIA DE LA

PRODUCCION EJIDAL.

-VPAT de predios privados que
producen hasta $25.000/VPAT. (55)
-VPAT de predios privados que
producen más de $25.000/VAPT (49)
-VPAT ejidal / VPAT (-96)

Factor 7 (6.74 %)

APROVECHAMIENTO DE LA TIERRA.

-Sprf. cosechada / LABT. (72)
-Sprf. fertilizada / LABT. (68)

LABT : superficie de labor total.
VPAT : valor de la producción agrícola.

La presentación del cuadro No. 3.4 está hecha para que las conclusiones afloren por sí mismas, por esto no se abundará en mayores comentarios. Una ligera reflexión extrae fácilmente las grandes contradicciones entre los dos estudios. No obstante, se comentará lo siguiente:

Lo primero que salta a la vista es la caprichosa caracterización de los factores en los dos análisis por su número, por su interpretación, por su composición, por la varianza explicada por cada factor y por las diferentes cargas de los factores.

El factor No.1 está compuesto en la segunda situación por datos referidos a la superficie de labor, en cambio en la primera están referenciadas las tres primeras variables al predio. La varianza que explica el factor 1 en el análisis del 83 es mucho más del doble que la respectiva del primer estudio. La única variable idéntica en el primer factor es el de la superficie con tracción mecánica en relación a la superficie de labor, pero con una muy baja correlación en el segundo caso, lo que haría cuestionar la presencia de esa variable en la interpretación del factor si nos atenemos a la discusión sobre la magnitud de la correlación que mencionábamos con anterioridad. Pero comparemos el factor 4 de la primera situación -es decir según el análisis de 1977- con el factor 1 del 83: ambos están compuestos por tres variables y coinciden las dos primeras; la tercera se diferencia en que en un caso es número de tractores y en el otro es superficie mecanizada, podemos decir que guardan una ligera familiaridad. Este factor 4 se interpretó como **DISPONIBILIDAD DE CAPITAL POR SUPERFICIE DE LABOR**, con sentido muy diferente al que se le dió al factor 1 en cualquiera de los dos estudios, no obstante guarda más parecido con el primer factor del análisis de 1983, que éste con el pertinente del 77. Pero, cuando menos desde el punto de vista estadístico del M.A.F., el factor cuatro no alcanza a explicar el 11 % de la variabilidad, mientras que el factor 1 del 83 se lleva poco menos del 41 % de la varianza.

A modo de moraleja: no se debe concebir, -ni nos podemos dejar arrastrar por- una interpretación de una serie de datos como el substrato empírico del cual podamos confirmar una teoría. Y una moraleja adicional: menos aún podemos montar construcciones teóricas sobre bases empíricas tan endebles, por no decir veleidosas. Y es importante esta anotación, porque de partida se planteó que la búsqueda inicial era de una comprobación empírica de unos ciertos planteamientos. En el camino se volteó el enfoque y se vislumbraron, se dilucidaron unos factores teóricos como los elementos explicativos del problema agrario. Y es natural este tipo de conclusiones en tanto no se critique desde una perspectiva teórica la utilización de esta técnica estadística. Lo úunico curioso es que con semejantes disparidades en los dos anáalisis era de esperarse, por parte de la autora, un severo cuestionamiento a la aplicacióon del modelo factorial, pero ni siquiera le mereció una nota comparativa.

2.4 EL FACTOR "MODERNIDAD".

Recientemente ha aparecido un artículo en la revista 'Estudios Económicos' de El Colegio de México en el cual se utiliza el M.A.F. en una investigación sobre la estructura manufacturera mexicana.[29]

Una vez que el autor hace referencia a los diferentes estudios sobre el tema, en los cuales de una forma u otra se plantea la hipótesis de dualidad en la industria como una de las características a destacar en los análisis ecoonómicos, señala el objeto del artículo en mención:

> "A partir de la última dicotomía del sector industrial y, restringiendo el estudio a la economía mexicana,el problema propuesto a consideración es una prueba empírica de la hipótesis de la estructura dual de la industria. La hipótesis que será contrastada empíricamente afirma que en el **CONTINUUM** de industrias ordenadas de acuerdo a algún indicador de modernidad, existe una ruptura causada por diferencias estructurales que caracterizan a la industria mexicana como dual, es decir, formada de dos subuniversos: las llamadas industrias moderna y tradicional. Más específicamente, se espera que tal ditribución será de hecho una distribución bimodal. Las modas constituirán los centros de dos, muy probablemente traslapadas, distribuciones normales."[30]

----

29. Martín Moreno (1987), p.p.: 81 - 112. Cfr. bibliografía. Es conveniente anotar que este trabajo es "una síntesis del ensayo de igual título que obtuvo el Premio Nacional de Economía 'Juan F. Noyola' correspondiente a 1984. Esta síntesis privilegia la prueba empírica y sus resultados principales."

30. Martín (1987), p. 82-83. Bastardillas en el original.

Se basa, posteriormente, en los conceptos sobre la economía dual de Averitt en los que la dualidad se refleja en una economía central y otra periférica. Dice Averitt que la economía central:

> "...se compone de empresas de gran tamaño e influencia. Su organización es corporativa y burocrática; sus procesos de producción son integrados verticalmente a través de la propiedad y control de los oferentes de materias primas críticas y de distribuidores de productos; sus actividades están diversificadas en muchas industrias, regiones y países. El apoyo financiero es fácilmente disponible de fuentes tanto internas como externas. Las empresas en la gran economía atienden mercados nacionales e internacionales, usando sistemas de producción progresivos tecnológicamente. Los asuntos de tales empresas son conducidos con la visión de sobrevivir en prosperidad en tanto enfrentan crisis económicas como estrategias sucesivas de expansión empresarial."[31]

La economía periférica, por el contrario, la define Averitt de la siguiente manera:

> "La otra economía está poblada por empresas relativamente pequeñas. Las empresas están usualmente dominadas por un solo individuo o familia. Las ventas de la empresa se realizan en mercados restringidos. Los beneficios y utilidades retenidas están comúnmente por abajo de aquéllos en el centro; el financiamiento de largo plazo es difícil de obtener. Las crisis económicas generalmente provocan la bancarrota o una severa contracción financiera. Las técnicas de producción y mercadeo son raramente modernas como las del centro. Estas empresas pequeñas son frecuentemente, aun cuando no siempre, seguidores tecnológicos..."[32]

----

31. Averitt,R.T.(1968). LA ECONOMIA DUAL. W.W. Norton. Nueva York. Citado por Martín Moreno (1987), p. 85.

32. O.C., p. 85.

Con este criterio de las dos economías de Averitt, Martín Moreno considera más adelante que la variable clave es el tamaño de la planta y que la distribución de las empresas es asimétrica y bimodal. Justificando este enfoque e hipótesis [33] menciona el ensayista varios trabajos que subrayan -a veces explícitamente- la presencia de la economía dual. En este punto afirma, refiriéndose e esos estudios:

> "No obstante, se observa la ausencia o la pobreza de los enfoques empíricos, de tal modo que existe la necesidad de una investigación en ese sentido."[34]

Resalta el autor a continuación un conjunto de "hipótesis heurísticas" que le permiten justificar el problema que se plantea y luego aborda la búsqueda de la "dimensión de la modernidad" en la industria mexicana dada una cierta correspondencia entre modernidad y concentracion industrial. Para esto, aduce Martín M. que:

> "La técnica que mejor se adecúa a este propósito es el análisis factorial, un procedimiento de reducción de la información estadística que permite descubrir los factores o dimensiones fundamentales que subyacen en un fenómeno... En este estudio se espera que sea posible ubicar un factor que concentre las principales variables que denoten modernidad. La desventaja de la técnica es que consiste sólo en una manipulación algebraica a partir de una matriz de coeficientes de correlación y, por tanto, carece de valor estadístico en sí misma."[35]

No comentaremos el anterior párrafo pues bastante se ha hecho al respecto en páginas anteriores. Sólo se quiere recordar, por lo pronto, la última frase sobre el poco valor estadístico en sí mismo del M.A.F.

En el estudio de Martín Moreno se tomaron 225 clases comunes a los censos industriales de México de 1970 y 1975. La descripción de las variables se encuentra en el cuadro No. 3.5 y es tomado del artículo de la revista 'Estudios Económicos'.

----

33. Cfr. p. 85-86.

34. O.C. p. 86.

35. O.C., p. 89.

C U A D R O No. 3.5.

DESCRIPCION DE LAS VARIABLES. [36]

COD.	VARIABLE	DESCRIPCION.
01	Grado de importancia.	Ventas netas por establecimiento.
02	Grado de productividad.	Valor agregado por trabajador productivo.
03	Grado de concentración.	Indice absoluto de concentración.
04	Grado de beneficio.	Tasa de beneficio.
05	Distribución del ingreso.	Beneficios brutos a remuneraciones.
06	Grado de diferenciación en el mercado.	Publicidad por establecimiento.
07	Grado de solvencia.	Gastos en intereses y préstamos a capital (activos fijos brutos).
08	Nivel de remuneraciones.	Remuneraciones a totales por trabajador.
09	Rel. capital-producto.	Activos fijos a producto.
10	Rel. capital-trabajo. (capital-salarios).	Activos fijos brutos a salarios.
11	Grado de expansión.	Inversión fija bruta en maquinaria y equipo por trabajador productivo. Flujo de capital por trabajador añadido en el período.
12	Estructura del empleo.	Trabajadores no productivos a trabajadores productivos.
13	Grado de dependencia externa en innovación.	Gastos en el uso de patentes, etc, a capital (activos fijos brutos).
14	Grado de participación de capital extranjero.	Participación de las empresas transnacionales en el producto total.
15	Grado de participación estatal.	Participación de empresas estatales en el producto total.
16	Estructura productiva A	Clasificación de bienes por tipo.
17	Estructura productiva B	Clasificación de bienes por tipo
18	Grado de participación del capital privado nacional.	Producto total menos producto de capital estatal y extranjero a producto total.
19	Grado de participación del capital nacional.	Producto total menos producto del capital extranjero a producto nacional.

----

36. Tomado de Martín M. 1987, p. 91.

CUADRO No. 3.6.

RESULTADOS DEL ANALISIS FACTORIAL. [37]

		1970			1975		
	VARIABLES	F1	F2	F3	F1	F2	F3
01	Importancia	81	20	27	71	33	34
02	Productividad	90	27	-03*	83	37	-06*
03	Concentración	46	26	09*	17	21	-15*
04	Beneficios	27	-15*	-67	30	07*	-38
05	Dist. ingreso	67	-15*	-43	65	-03*	-45
06	Difer. mercado	67	28	-22	60	36	-17
07	Solvencia	28	-14*	50	24	13*	36
08	Remuneraciones	75	38	28	68	44	32
09	Capital-producto	18	-07*	68	26	-22	55
10	Capital-trabajo	79	-02*	31	83	-02*	21
11	Expansión	58	23	44	68	16*	39
12	Estructura empleo	56	44	01*	37	37	-18
13	Dependencia en innovación	38	28	-36	19	64	-22
14	Cap. extranjero	31	90	11*	17	93	15*
15	Capital estatal	29	-43	43	17	-16*	48
16	Est. Pdtva. A	10*	08*	61	09*	08*	62
17	Est. Pdtva. B	-04*	25	59	-15*	27	57
18	Capital nacional	-32	-89	-11*	-17	-93	-14*
19	Cap. Privado Nal.	-31	-90	-10*	-17	-93	-14*

* No significativa estadísticamente.

N.B.: las cargas se redondearon a dos dígitos y se presentan como enteros. En el original hay contradicción en las definiciones de las variables 18 y 19 en los dos cuadros, pero no es una contradicción tal que afecte las conclusiones.

37. O.C., p. 93.

En el cuadro 3.6 se resumen los resultados del análisis factorial, después de rotar los factores por el método varimax del SPSS. Se debe anotar que las variables fueron transformadas "en general en forma logarítmica."[38]

El investigador -a través del método del factor principal- obtuvo tres factores que explican el 60 % de la varianza. De estos factores, "el primero reunía las características que pueden considerarse como indicadoras de modernidad."[39] Este factor explica el 35 % y el 31.4 % en los dos años de estudio, respectivamente. También realizó pruebas de hipótesis sobre la significancia estadística de las cargas de los factores.

El autor destaca cuatro aspectos del análisis factorial del cuadro No. 3.6. El primero de ellos es el que:

> "a) La presencia de capital extranjero y estatal en el factor central hace la siguiente hipótesis plausible: que tanto el capital extranjero como el estatal contribuyen a la modernización de la estructura industrial mexicana, cada uno con sus características específicas."[40]

Llama la atención ese comentario porque si la conclusión deviene de las cargas de los factores -que son las correlaciones de las variables con el factor-, éstas son bajas para 1970: 0.31 y 0.29 respectivamente para las variables mencionadas. Para 1975 las correlaciones son muchísimo menores: 0.17 para las dos variables. De acuerdo a la magnitud de las cargas, ocuparían en orden de importancia los puestos 14 y 15 sucesivamente en 1970. Para 1975 sólo dos variables tendrían coeficientes menores. Por esta razón, si nos atenemos a los coeficientes, tendríamos que resaltar la importancia de las otras variables que tienen cargas

----

38. O.C., p. 90.

39. O.C., p. 90.

40. O.C., p. 92-93.

mayores. Aún más: estas variables se resaltan en mayor medida en el factor 2 que en el primer factor, pues junto con las otras variables de capital, contribuirían a su interpretación. El autor está reflexionando solamente a partir del signo del coeficiente, pero es una conclusión forzada puesto que desde el punto de vista del mismo M.A.F. no serían las de mayor impacto en la "modernidad".

Veamos el segundo aspecto:

> "b) Por otra parte, el capital nacional total y el privado tienen cargas factoriales negativas, sugiriendo que, en general, estos dos capitales no son identificados con las características centrales o modernas. Esto también confirma la hipótesis del capital nacional en este sentido."[41]

Uno de los problemas del M.A.F. es que al retomar variables complementarias se tiende a inducir cargas con signo negativo. Esto se presenta porque las correlaciones entre variables complementarias son muy altas pero no se pueden reflejar con el mismo sentido dado ese caracter complementario. Esto es lo que sucede con variables como las participaciones del capital extranjero, estatal, privado nacional y nacional en el producto total. Por este motivo se cuestiona -se invalida- la última afirmación del párrafo pasado. Como ilustración del movimiento de los signos de los coeficientes, compárense en el cuadro No. 3.4 los signos resultantes en los coeficientes del primer factor para 1960. En algunos casos se encuentran altas correlaciones con signos positivos y negativos indistintamente.

Leamos la tercera conclusión:

> "c) La **NO** significancia estadística de las variables 16 y 17, que representan dos estructuras productivas diferentes, es de gran interés. La idea de introducirlas es probar la hipótesis de que en el desarrollo económico mexicano las industrias que producían bienes intermedios y durables eran las más dinámicas y relativamente más modernas. El dominio

----

41. O.C., p. 93.

> llamado **ESTRUCTURA PRODUCTIVA A** fue diseñado de tal manera que si hubiera resultado significativo, hubiera confirmado la hipótesis de que tales tipos de industrias eran de hecho modernas. El dominio llamado **ESTRUCTURA PRODUCTIVA B** se diseñó para sugerir que la modernidad debe ser asociada con bienes de capital, durables, intermedios y de consumo, en ese orden precisamente. Los resultados no son definitivos, por lo menos en este factor, ya que no existe significancia estadística; así pues, las cargas factoriales pueden ser en realidad iguales a cero."[42]

De nuevo aquí se intenta probar hipótesis. Sin embargo, para poder hacerlo ias variables originales deben distribuirse normalmente, pero es sumamente dudoso que alguna de las 19 variables empleadas en el análisis pueda cumplir esta propiedad. Aún más, para emplear una prueba de hipótesis se deben calcular los factores a partir del método de máxima verosimilitud, el cual exige distribución normal multivariada de las variables y éste precisamente no fue el caso. Si la prueba se refiere a la significancia estadística de los coeficiente de correlación, también habría que asumir distribuciún normal. Se hicieron una serie de pruebas a pesar de que Martín Moreno había afirmado que la técnica del análisis de factores carece de valor estadístico en sí misma.

Por último, el cuarto aspecto:

> "d) Un resultado interesante es que la variable 'concentración' tiene relativamente bajas cargas factoriales (46 y 17), aun cuando significativas. Esta situación es difícil de explicar, ya que la industria mexicana es considerada como altamente concentrada. Así pues, el dominio **CONCENTRACION** se suponía asociado con otras variables importantes. Para 1970 parece que tal liga existe, aun cuando no extremadamente fuerte. Para 1975 la asociación es muy baja. Al parecer, la concentración, como variable clave, mantiene sin

----

42. O.C., p. 93-94. Bastardillas en el original.

> embargo, su importancia, ya que fue estadísticamente significativa, pero también, en muchos casos, no puede ser usada como la única variable explicativa para entender un fenómeno. En este contexto, este estudio no se ve afectado en sus conclusiones por este hecho, ya que hace uso de un conjunto de variables clave, en lugar de usar un enfoque monovariable."[43]

Pero lo mismo podriamos decir de variables "clave" como lo son la tasa de beneficios, los gastos en intereses y préstamos a capital, la relación capital-producto, el grado de dependencia externa en innovación, el grado de participación del capital privado nacional y del capital nacional e incluso del capital extranjero, este último que tan importante ha resultado en el estudio de Martín M. en la conformación de la "modernidad" de la estructura productiva. Todas las ocho variables referidas tienen común tres cosas: la primera, coeficientes muy bajos -en general, más chicos que el de la variable "concentración"- y por ende habría que hacer salvedades similares para cada una de ellas; la segunda, que la mayoría de ellas experimentó una ponderación menor en 1975, contrariamente a todo lo esperado, lo cual se interpretaría como un **debilitamiento** del factor "modernidad" entre el 70 y el 75; la tercera, que de acuerdo al criterio de la dualidad de Averitt, éstas son las variables más importantes en la definición y diferenciación de la estructura dual en la industria mexicana. Se supone que un elemento distintivo entre la industria "central" y la "periférica" es la diferente tasa de beneficios. Otro, que la relación capital-producto marca una aguda diferencia en la defínción de los dos entornos. Se asume además, que las grandes empresas tienen sistemas de financiación tales que los distinguen abruptamente de las pequeñas empresas, las cuales forman parte del segundo "polo" industrial. Asimismo, por sus mayores vínculos con la tecnología extranjera, en las empresas del "centro" habría un alto grado de dependencia externa en innovación y, al contrario, en el otro extremo de la industria esta línea carecería de importancia. El capital extranjero, que dentro del estudio en referencia ha jugado un rol de protagonista crucial de la modernización por sus vínculos con las empresas transnacionales y la innovación tecnológica, expresa ponderaciones menores que las establecidas para la variable "concentración", pues éstas son 31 y 17 para los dos años de estudio, respectivamente. Este resultado cuestionaría el aserto a) del material criticado rspecto a su contribución en la modernización industrial. Por último, el capital privado

----

43. O.C., p. 94. Bastardillas en el original.

nacional y el capital nacional carecen de importancia en la definición de la estructura productiva manufacturera, lo cual crearía interrogantes respecto al papel del capital privado nacional en la conformación del "polo periférico" de la industria.

El reconocimiento de las contrariedades sufridas por la variable "concentración", debe pues, proyectarse a un conjunto más amplio de variables. Particularmente queremos enfatizar que los cuestionamientos del autor del artículo, detallados en el inciso d), debió aplicarlos también - como mínimo- a la variable "capital extranjero", pues dentro de la ubicación teórica e histórica que el ensayista hace en la tercera de sus "hipótesis heurísticas" [44] hace figurar al capital extranjero como el abanderado de la modernidad. Infortunadamente el análisis de factores no destacó muchas cosas que esperanzadamente se suponía iba a realizar.[45]

Como argumento tangencial, puede cuestionarse la definición y medición de las variables capital nacional y capital privado nacional, así como las fuertes conclusiones sobre el papel jugado por ellas en el proceso de industrialización en el breve período entre 1970 y 1975. Recordemos que los sectores que en ese lapso experimentaron las más altas tasas de crecimiento del PIB en la economía mexicana fueron, por una parte , Transporte, Almacenamiento y Comunicaciones -12.2 % promedio anual en el quinquenio- y, por otra parte, electridad -con 9.9 %-, rubros en los cuales es obvia la participación del capital nacional (privado y estatal), frente a un crecimiento total del producto del orden del 6.6 % en los cinco años mencionados.[46] Las conclusiones de Martín Moreno tomando como fuente empírica la información arrojada por la corrida del análisis factorial del paquete de computación estadístico SPSS, evidencian las grandes y graves limitaciones de la técnica en referencia.

----

44. Cfr. p. 98.

45. Ver la cita No. 35 sobre las expectativas del M.A.F. al comienzo de esta sección, al justificarlo como técnica capaz de lograr el propósito de encontrar la dimensión de la modernidad en la industria mexicana.

46. Nafinsa, p. 63.

A partir de este primer factor encontrado, Martín Moreno encuentra los llamados "puntajes factoriales" para 1970 y 1975 buscando estudiar la forma de sus distribuciones en cada uno de los dos años para pretender demostrar empíricamente la existencia de la dualidad en la industria manufacturera mexicana. Según los resultados por él expuestos, las distribuciones de los valores factoriales son sesgadas a la derecha y leptocúrticas. Utiliza una prueba estadística para corroborar si la forma de las distribuciones obedece a una mezcla de dos distribuciones normales las que, de encontrarlas, darían razón de la dualidad industrial en el caso de México. Nuestro interés no es entrar a discutir esta parte, precisamente; nuestra inquietud es plantear el sentido mismo de los puntajes factoriales. Veamos: la construcción de este índice de modernidad se basó en la estructura factorial ya criticada, esto como una primera parte. Como una segunda, recordemos que los tres factores explicaban el 60 % de la varianza del modelo. Adicionalmente, si incluimos sólo el primer factor, se explica el 35 % y el 31.4% de la varianza entre todos los factores para los dos periodos considerados. No obstante, a pesar de estas bajas proporciones de las varianzas explicadas se utiliza como criterio básico para "mostrar que existen suficientes elementos para considerar que hay consistencia entre la hipótesis del estudio y la información estadística relevante al caso."[47]. Ya nos habíamos referido a lo artificioso del modelo, pero mayor artificio desempeña la utilización de estos puntajes o valores factoriales para construir un índice, en este caso, de modernidad. Por eso nuestra crítica se detiene aquí, pues queda cuestionado seriamente el proceso de obtención de factores, y por ende, la conformación de índices a partir de esta técnica. También por eso, cualquier resultado y análisis posterior tiene un andamiaje teórico estadístico sumamente endeble.

----

47. O.C., p. 94.

Capitulo 4

DISCRIMINACION, AGRUPACION, CORRELACION CANONICA.

En este capitulo se pretende una exposición de las tres técnicas multivariadas mencionadas, privilegiando dos aspectos: el primero es el enfoque más intuitivo que matemático; el segundo, que resaltamos el propósito descriptivo antes que el modelístico propiamente, entendiendo esta distinción entre una exposición alternativa y organizada de la información bruta, aspecto al que responde el espiritu descriptivo, y un encasillamiento a una forma predeterminada de análisis, noción que se acerca a la intención modelizadora. Es por este motivo que intencionalmente se descuida la presentación de estas técnicas con un enfoque probabilistico que exige el cumplimiento de ciertos supuestos distribucionales. Los casos más claros en esta diferenciación corresponden a los análisis de cúmulos o conglomerados y a la correlación canónica.

4.1 ANALISIS DISCRIMINANTE.

Si contamos con información de dos conjuntos de individuos que forman parte, respectivamente, de dos poblaciones diferentes, el análisis discriminante -A.D.- se propone establecer una regla de asignación que nos indique si nuevos elementos no considerados anteriormente pertenecen a una de las dos agrupaciones.[1] Para lograrlo, el A.D. encuentra una combinación lineal de las variables independientes de los dos grupos, buscando que esta especial combinación esté dotada de un gran poder diferenciador o discriminador. A esta combinación lineal se la conoce como función discriminante.

4.1.1 EXPOSICION GRAFICA.

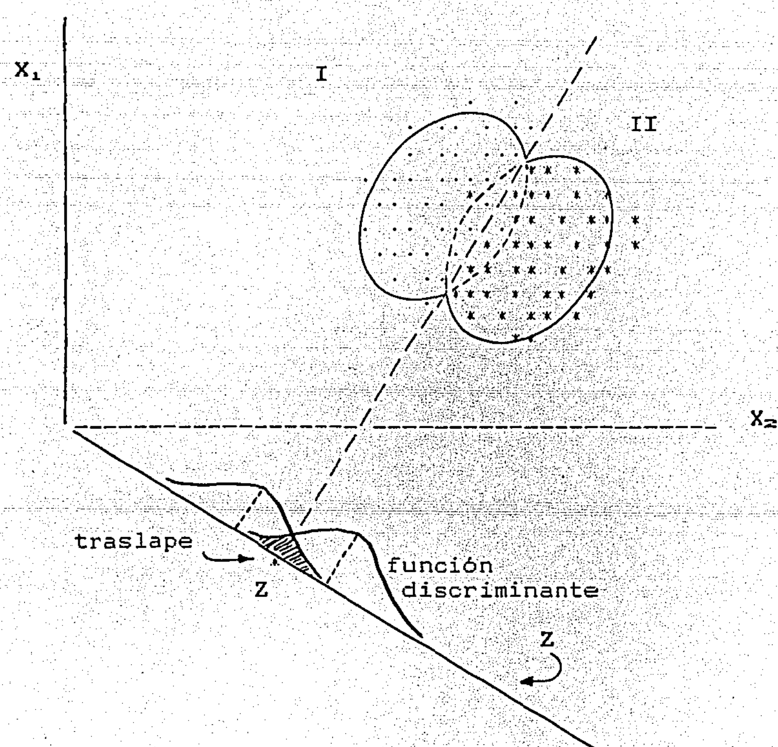
La gráfica 4.1 representa un diagrama de dispersión cuya proyección muestra qué sucede cuando se obtiene una función discriminante. Si tenemos dos grupos -I y II- y un par de mediciones -X_1 y X_2- sobre cada elemento de los dos grupos, podemos representar gráficamente la asociación de las dos variables para cada elemento de I y de II. Las elipses de la gráfica 4.1 encierran una proporción determinada de las asociaciones -usualmente el 95 %-. La línea recta punteada que atraviesa a las dos elipses por sus puntos de intersección se proyecta sobre un nuevo eje Z. El traslape entre la distribución de los dos conjuntos es el más pequeño que podemos obtener sobre cualquier otra línea recta que pase a través de las elipses. El nuevo eje expresa el perfil bivariado de los dos grupos en una

----

1. El análisis puede ser fácilmente ampliado a más de dos poblaciones.

sola dimensión: a ésta la conoceremos como puntaje o índice discriminante, el cual condensa la información acerca de las diferencias entre grupos en un conjunto de puntos sobre un nuevo eje. En resumen, para un análisis discriminante se encuentra una combinación lineal de las variables independientes; de esta manera se obtiene un conjunto de índices o puntuaciones discriminantes para cada elemento en cada uno de los grupos.

Gráfica No. 4.1.



Considerando las variables mencionadas en la gráfica anterior tomadas en un número n de individuos que pertenecen a dos distintos conjuntos la función lineal la escribimos como:

$$Z_{it}=w_1X_{i1t}+w_2X_{i2t} \quad (4.1.1)$$

$$i = 1, 2. \qquad t = 1 \ldots n.$$

Donde:

Z : representa el puntaje o indice discriminante.
X : representa las variables independientes.
w : representa la ponderación discriminante.

A la expresión (4.1.1) la llamamos función discriminante lineal con coeficientes W desconocidos. El subíndice i = 1, 2, nos está diciendo a cuál de las dos poblaciones pertenece la variable respectiva, por lo tanto se refiere al grupo. El subíndice t hace mención a las observaciones en cada grupo.

La ecuación (4.1.1) también puede representarse mediante un plano y la gráfica 4.2 nos ilustra geométricamente la situación. Las proyecciones de X_{i1t} y X_{i2t} transforman los puntajes bidimensionales en uno unidimensional.

En las elipses se concentran los diagramas de dispersión de las dos variables X. El plano corta a las elipses de una manera tal que bajo el plano queda la mayor parte de la elipse, I y sobre él, la mayor parte de la elipse II. Los puntajes que se proyectan sobre el eje Z por encima de Z^* se clasificarán como pertenecientes al grupo I. Este es el caso de Z_{1t}, por ejemplo. Análogamente, puntuaciones cuyas proyecciones caigan por debajo de Z^* se clasificarán como integrantes del grupo II. Una proyección de este tipo es Z_{2t}. De manera similar, las medias de los grupos I y II -llamados los centroides- se proyectan sobre el eje de las zetas como M_{z1} y M_{z2}.

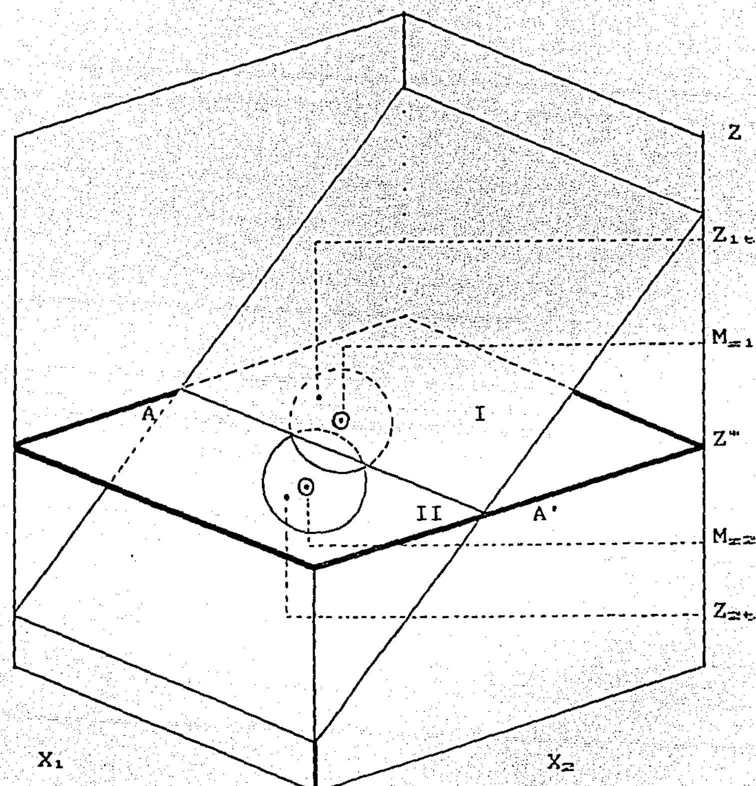
Como podemos analizar de la gráfica 4.2, la clasificación incorrecta ocurre cuando un elemento del grupo I se proyecta por debajo de Z" o viceversa. Si el plano corta a la elipse en los puntos de intersección, el área de clasificación errónea será más pequeña que cualquier otro plano que corte a la elipse en puntos distintos de su intersección. La separación de los grupos en cada uno de los planos discriminantes se puede expresar como el cuadrado de la diferencia -distancia- entre los centroides:

$$[M_{z1} - M_{z2}]^2 \quad (4.1.2)$$

Esta distancia mide la variación entre las medias de los puntajes discriminantes para cada grupo. Es una medida de la variación entre los grupos. Si el plano fuese más inclinado que el de la gráfica 4.2, la proyección de cada centroide sobre el plano sería tal que aumentaría las distancias entre las respectivas medias. Si nuestro interés es separar entre grupos, un plano más inclinado que pase por AA' lograría este objetivo al aumentar las distancias entre los centroides.

Pero, qué sucede al interior de los grupos ? Las distancias al interior de los grupos son más cortas en la situación original que cuando inclinamos más el plano. Esto quiere decir que las diferencias al interior de los grupos también se hacen mayores a medida que el plano gana ángulo.

Gráfica No. 4.2



Por consiguiente las distancias como:

$[Z_{1t}-M_{z1}]^2$ y $[Z_{2t}-M_{z2}]^2$ (4.1.3)

miden la variación al interior de los grupos. Por lo tanto el costo de mayor diferenciación entre los grupos es una mayor variación dentro de éstos. Evidentemente, una gran variación dentro de los grupos es indeseable si nuestro propósito es diferenciar entre ellos: cualquier distancia entre los dos centroides estará afectada por la mayor distancia al interior de cada conjunto. La alternativa es encontrar un plano discriminante óptimo obtenido por la maximización de la siguiente razón:

$$\frac{\text{Variación ENTRE grupos.}}{\text{Variación DENTRO de grupos.}}$$

Formulación que podemos expresar matricialmente como:

$$\frac{w'(\bar{x}_1-\bar{x}_2)(\bar{x}_1-\bar{x}_2)'w}{w'S^*w} \qquad (4.1.4)$$

Donde:

w : matriz de ponderaciones.

$\bar{x}_i$: vector de medias del i-ésimo grupo.

S^*: matriz de covarianzas mancomunadas.

Obteniendo las derivadas parciales e igualando a cero, los valores estimados de los coeficientes vienen dados por:

$$w^* = S^{-1*}[\bar{x}_1-\bar{x}_2] \quad (4.1.5)$$

Si son iguales los costos de clasificación incorrecta para los dos grupos y la probabilidad de que una observación provenga de cualquiera de ellos, entonces Z^* caerá en la mitad de la diferencia entre los centroides. Por esta razón, muestralmente podemos decir que:

$$Z^* = [1/2]W'[\bar{x}_{.1} + \bar{x}_{.2}] \quad (4.1.6)$$

y en consecuencia una nueva observación se clasificará en el grupo I si

$Z_{1t} > Z^*$ y en el grupo II si $Z_{1t} < Z^*$

Si definimos el vector Z' como:

$Z' = [Z_{11}, Z_{12}, \ldots, Z_{1n_1}, Z_{21}, Z_{22}, \ldots, Z_{2n_2}]$.

entonces la función discriminante la podemos expresar así:

$$Z = W' X' \quad (4.1.7)$$

Y la regla de discriminación sería entonces:

se clasifica en el grupo I si:

$$W'X' - [1/2]W'[\bar{x}_{.1} + \bar{x}_{.2}] > 0.$$

se clasifica en el grupo II si:

$$W'X' - [1/2]W'[\bar{x}_{.1} + \bar{x}_{.2}] < 0.$$

Conclusión: se obtuvieron puntuaciones discriminantes basados en la regla estadistica de optimizar la varianza entre los grupos con respecto a la varianza dentro de ellos. Si la primera es muy "grande" en relación a la varianza dentro de grupos, entonces consideramos que la función discriminante separa los grupos adecuadamente. Lo contrario también se cumple.

4.1.2 SOBRE LAS CONDICIONES DEL ANALISIS DISCRIMINANTE.

1. Para una aplicación adecuada del A.D. es necesario que se cumplan ciertas condiciones: la primera de ellas, que la distribución de las variables independientes debe ser normal multivariada; la segunda, que la estructura de covarianzas entre los grupos sea similar; es decir, se supone que la dispersión de los grupos es igual, aunque desconocida. Además, se suponen iguales costos de clasificación incorrecta -como se definió en el párrafo anterior-. Muchos autores, no obstante, afirman que el análisis discriminante no es muy sensitivo al relajamiento de estos supuestos, a menos que las violaciones sean extremas.

2. Aunque no se ha hecho explícito, las variables independientes deben estar medidas en una escala de intervalo o razón, en tanto que las variables dependientes se consideran en una escala categórica. Esto último quiere decir que los grupos en que se divida la información -dos o más- deben ser exhaustivos y mutuamente excluyentes.

3. Los métodos de cómputo de la función discriminante son dos: el simultáneo y el paso a paso. El primero contempla el conjunto completo de las variables independientes sin evaluar el comportamiento discriminante de cada una de ellas. Por el contrario, cuando esto último es lo deseado, entran una a una las variables en la función en consideración a su capacidad diferenciadora.

4. La obtención de funciones discriminante se hace a través de paquetes computacionales y la significancia estadística de la función se prueba con los estadísticos X^2 y D^2 -distancia de Mahalanobis-. No obstante son pruebas débiles y su significancia no es muy grande: si dos grupos son significativamente diferentes a un nivel del diez por ciento con tamaños de muestra suficientemente grandes, los centroides pueden ser virtualmente idénticos y no obstante podemos obtener una prueba significativa. Refiriéndose a estos estadísticos Hair comenta:

"...En consecuencia, el nivel de significancia de estos estadísticos es una indicación muy pobre de la habilidad de la función para discriminar entre los dos grupos. Para clarificar más ampliamente la utilidad del procedimiento de la matriz de clasificación, lo relacionamos con el concepto del R^2 en el análisis de regresión. Muchos de nosotros probablemente hemos leído artículos académicos en los cuales el autor ha encontrado relaciones estadísticamente significativas que explican solamente un diez por ciento (o menos) de la varianza, por ejemplo, cuando un R^2 es igual a 0.10. Usualmente este valor es significativamente diferente de cero simplemente por el gran tamaño de la muestra. Con análisis discriminante múltiple la razón del porcentaje correctamente clasificado es análogo al R^2 de la regresión. Esta razón revela qué tan bien clasificó a las unidades estadísticas la función discriminante; en tanto el R^2 nos dice qué tanto de la varianza ha explicado la ecuación de la regresión. La prueba F para la significancia del R^2 es, asimismo, análoga a las pruebas de significancia X^2 o D^2. En A.D. podríamos tener una distancia significativa entre dos o más grupos, aunque clasifiquemos correctamente sólo el 53 %."[2]

5. Con la advertencia anterior, después de que la prueba estadística indica que la función discrimina significativamente, se acostumbran desarrollar las llamadas matrices de clasificación, las cuales intentan proveer una mayor precisión del poder discriminatorio de la función. Antes de construir esta matriz se debe determinar el índice discriminante -el Z^* ya mencionado-. Este Z^* es el límite contra el cual se juzga en qué grupo debe caer una puntuación individual. Cuando los tamaños de los grupos son iguales, el puntaje óptimo límite será el promedio de los centroides. Desde luego, cuando el tamaño es diferente se utiliza el promedio ponderado de los centroides:

$$Z^* = \frac{M_{x1} + M_{x2}}{2} .$$

----

2. Hair et al., p. 96.

$$Z^{*} = \frac{n_1 M_{x1} + n_2 M_{x2}}{n_1 + n_2}$$

Donde:

M_{x1} y M_{x2}: representan los centroides de los grupos.

n_1 y n_2: indican los respectivos tamaños.

6. Se construyen matrices de clasificación para convalidar la función discriminante. La mecánica es la siguiente: se divide aleatoriamente la muestra total en dos grupos. Uno de ellos, llamado la muestra de análisis se utiliza para calcular la función discriminante; el otro grupo -llamado muestra de validación- se utiliza para construir la matriz de clasificación. En esta última las entradas de la diagonal representan el número de individuos correctamente asignados a su respectivo grupo y por fuera de la diagonal la cantidad de elementos clasificados erróneamente. Teniendo esta información es fácil establecer el porcentaje de elementos clasificado acertadamente.

7. Una vez que se ha establecido que la función discriminante es estadisticamente significativa y que se considera que la precisión en la clasificación es aceptable, se analizan los resultados a través de las ponderaciones discriminantes, la estructura de correlaciones y los valores parciales F. Con las ponderaciones discriminantes -las W de la ecuación 4.1- se examinan signo y magnitud de los coeficientes asociados a cada variable: se espera que ponderaciones más grandes acusen un mayor poder discriminatorio que aquéllas que experimentan coeficientes más chicos. No obstante estos juicios deben hacerse con mucha precaución. Veamos un comentario al respecto:

> "La interpretación de las ponderaciones discriminantes es análoga a la interpretación de las ponderaciones de los beta en análisis de regresión y por ende está sujeta a las mismas críticas. Por ejemplo, una poderación pequeña bien puede significar que la variable correspondiente es irrelevante en la determinación de una relación, o bien que ha sido parcialmente sacada de la relación porque hay un alto grado de multicolinealidad. Otro problema con el uso de las ponderaciones discriminantes es que están sujetas a considerable inestabilidad. Estos poblemas sugieren precaución al utilizar estas ponderaciones al interpretar los resultados del análisis discriminante."[3]

Como alternativa se utiliza la estructura de correlaciones, es decir, al coeficiente de asociación lineal simple entre cada variable independiente y la función discriminante. Esta correlación refleja la varianza de las variables independientes con la función discriminante y puede ser interpretada como la carga del factor que define la contribución relativa de cada variable independiente a la función discriminante. Cuando se utiliza el método de paso a paso, un medio adicional de interpretar la facultad discriminatoria relativa de las variables independientes es el uso de los valores parciales F. Examinando los puntajes significativos de cada F, se pueden ordenar descendentemente a las variables que tienen más capacidad diferenciadora. En fin, con estos tres criterios el investigador se puede formar una idea de la importancia de las respectivas variables para propósitos de discriminación.

----

3. Hair et al., p. 104.

4.2 ANALISIS DE CONGLOMERADOS.

En el análisis discriminante se contaba con información de n elementos con p variables clasificados previamente en subconjuntos; es decir, se conocía de antemano la pertenencia de cada uno de los elementos a sus respectivos grupos. Con el análisis de conglomerados o de cúmulos la meta es distinta: se trata de encontrar estos subconjuntos cuando se ignora la pertenencia de los individuos a un grupo predefinido, el propósito es conformar los cúmulos en los cuales se agrupan esos n objetos sobre la base de sus similaridades internas.

En principio, se puede pensar en un conglomerado como un subconjunto de objetos que están muy "cercanos" entre si. Desde luego, uno de los problemas que inmediatamente debemos enfrentar es lo que se pretende entender como "cercanía", la cual se puede concebir de muy diversas formas. No obstante podemos designar a un conglomerado, de una manera más genérica, en el sentido de una agrupación dispersa alrededor de un valor central. En una situación en la que tengamos tres variables y un buen número de observaciones sobre ellas, una gráfica en tres dismensiones puede ayudarnos a detectar visualmente la presencia de agrupaciones en el conjunto original de información.

Por este motivo, uno de los criterios más usuales en las técnicas de conglomerados es el que dos conjuntos pueden ser pensados como grupos separados dependiendo de las distancias entre sus valores centrales en relación a las distancias medias dentro de cada grupo, un criterio que guarda una gran similitud con la prueba de homogeneidad de clases usada en el análisis de varianza y muy parecido al criterio de discriminación que señalamos en el A.D. en la sección anterior.

Un primer acercamiento para hacer conglomerados es a través de la matriz de correlaciones de las variables. En este caso el agrupamiento se entiende como la "cercania" de vectores y una forma de medirla -ya lo veiamos en el primer capitulo- bien puede ser el valor absoluto del coeficiente de correlación entre los datos. Esto significa considerar una magnitud del coeficiente -por ejemplo un R = 0.7- por encima del cual suponemos que el agrupamiento de variables es estrecho y bastaria con una mirada a la matriz de correlaciones para escoger las variables que a nuestro juicio fomenten conglomerados.[4]

Debemos tener en cuenta que en el caso en mención el sentido del agrupamiento está asociado a conformar conglomerados entre las variables y no entre los elementos a los cuales se les miden esas p variables. La dificultad mayor estriba en el hecho de conformar una unidad de medición de las distancias entre los elementos que pueda estar libre de las escalas de medición que reflejan las variables, pero el cambio de escala altera los agrupamientos. Por esta razón es recomendable escalar las variables para que tengan varianza unitaria y por consiguiente se pueda disminuir drásticamente los efectos por la escala. Otra posibilidad es utilizar los componentes principales para reducir significativamente las dimensiones originales. En fin, como ya se vislumbra, hay diferentes técnicas y acercamientos para enfocar la construcción de conglomerados.

Podemos agrupar las diferentes técnicas de conglomerados en dos tipos: el primero, cuando el interés analítico es estrictamente descriptivo y no son necesarios supuestos distribucionales sobre la población de la que proviene la información; en el segundo tipo, por el contrario, se considera la presencia de un modelo subyacente en el cual existe un juego entre observaciones y un pequeño número de diferentes distribuciones asociadas.[5] En tanto nos interesa la utilización de los métodos descriptivos como una forma de organización de los resultados en una apreciación inicial de la información, nos concentraremos en el primer tipo, de los que forman parte los llamados métodos jerárquicos. Los modelos probabilísticos tienen limitaciones similares a las esbozadas en el M.A.F. en el sentido del intento

----

4. Véase al respecto un sencillo e interesante ejemplo en Kendall, p. 34-36.

5. Una formulación probabilística que responde a este segundo enfoque puede verse en Mardia, p. 361 y ss.

de reacomodar la realidad a las exigencias del modelo. Las técnicas descriptivas nos permiten aprehender la información con una organización y ordenamiento determinados que pretenden visualizar sus rasgos más característicos, basados en lo que los datos por sí mismos nos puedan ofrecer -y el investigador descubrir- antes que por su manipulación modelística.

4.2.1 LAS DISTANCIAS.

Antes de estudiar los métodos de jerarquización, conviene esbozar lo que se maneja en conglomerados como técnicas de detección de "cercanías". Veiamos que esta noción nos lleva a consideraciones sobre las "distancias" entre los individuos. En principio, podemos plantear como regla general que la distancia entre el elemento A y el elemento B sea la misma, cualquiera sea el sentido en que se la mida. Esto quiere decir que las medidas que estamos haciendo tienen una connotación de objetividad que no dependen de criterios como afinidad, afición o inclinación. Particularmente puede existir una mayor inclinación -de cualquier tipo- de B hacia A que de A hacia B y por lo tanto una distancia en la que sucediera algo como lo último perdería ese carácter de objetividad. Los estudios de preferencias individuales o de grupo sobre política, candidatos u opiniones pueden adolecer de infracciones a esa primera regla, de ahí que su afirmación no es tan trivial como aparece a primera vista. Para completar estas "reglas" sobre distancias, otras precisiones importante son que las distancias no pueden ser negativas y que la distancia entre un individuo con respecto a sí mismo tiene que ser cero. Cada una de las aseveraciones que hemos mencionado corresponde, respectivamente, a la simetría, la no-negatividad y la identidad como tres de las propiedades de las distancias. Precisemos: tenemos dos individuos, elementos u objetos -A y B- los cuales representan mediciones de sus características -variables-. Decimos entonces que una función real es una función de distancia si cumple con:

1. $d(A,B) = d(B,A)$	Simetría: la distancia de A hacia B es la misma que de B hacia A.
2. $d(A,B) \geq 0$	Positividad. La distancia entre A Y B no puede ser negativa.
3. $d(A,A) = 0$	La distancia de un individuo con sí mismo tiene que ser nula.

Si se añaden mayores propiedades a las distancias encontraremos lo que se denomina como métrica. Las propiedades adicionales son la difinitividad que nos dice que si la distancia entre **A** y **B** es cero es porque **A=B**. De esta forma siempre habrá un criterio distintivo en tanto los elementos sean diferentes. Hay otra propiedad llamada la desigualdad triangular que afirma que la distancia entre **A** y **B** no puede ser mayor que la suma de las distancias entre **A** y un tercer elemento **C** con la distancia entre **B** y **C**:

4. $d(A,B) = 0$ si y sólo si $A = B$: Definitividad.

5. $d(A,B) \leq d(A,C) + d(C,B)$: Desigualdad triangular.

A la distancia entre **A** y **B** -$d(A,B)$- se la concibe también como una medida de divergencia: se incrementa cuando la similaridad o cercanía entre **A** y **B** disminuye; por esta razón la $d(A,B)$ puede pensarse como un coeficiente de disimilaridad.

Dentro de las métricas más conocidas, la más sencilla de ellas es la llamada distancia euclidiana simple. Con p variables para los dos puntos de referencia el cuadrado de la distancia de A a B es esa distancia euclidiana:

$$d^2(A,B) = \Sigma_{i=1}[X_{iA} - X_{iB}]^2 \quad (4.2.1)$$

Donde los subíndices A y B se refieren al individuo respectivo.

La expresión (4.2.1) se puede interpretar geométricamente como la hipotenusa en p dimensiones y puede pensarse como una distancia en el sentido métrico común. Esta distancia euclidiana tiene algunas características que en determinados estudios la hacen especialmente útil. Una de éstas es su invariabilidad cuando se rotan los ejes, lo cual significa que si se usan puntajes obtenidos por componentes principales de las variables originales, las distancias no se alteran. Otra característica importante es el que sus cálculos se simplifican notoriamente: es muy común la utilización del promedio general de distancias euclidianas, lo que se traduce en que con n observaciones, habrían 1/2 n(n-1) distancias euclidianas promedio por computar. No obstante, por manipulación matemática la distancia euclidiana general promedio puede obtenerse fácilmente con una expresión como la siguiente:

$$[2/[n-1]]\Sigma_j\Sigma_i[X_{ij}-X_{i.}]^2 \quad (4.2.2)$$

Donde:
$X_{i.}$: es la media sobre las n observaciones en el iésimo grupo.

Cuando las escalas y unidades de medición de las variables son heterogéneas se acostumbra a utilizar la distancia de Pearson que no es más que la distancia euclidiana pero estandardizada:

$$K^2 = \frac{\Sigma_i[X_{iA}-X_{iB}]^2}{S_i}$$

Donde:
S_i : es la varianza de la iésima variable.

A diferencia de la euclidiana, la distancia de Pearson no se ve afectada por los cambios de escala en el proceso de uniformización de las variables debido a la influencia de la varianza de ellas en su formulación. En ocasiones, en lugar de escalar por la varianza se utiliza el rango, definido como la diferencia máxima entre las observaciones de la variable correspondiente:

$$R=\max_i[X_{iA}-X_{iB}] \quad .$$

Una métrica muy común es la llamada distancia generalizada o distancia de Mahalanobis, la cual se expresa como sigue:

$$[x_A - x_B]' S^{-1} [x_A - x_A] \quad (4.2.3)$$

Donde:

$x_{A \text{ ó } B}$: vector de observaciones respectivo.

S: matriz de varianza-covarianza de las variables.

La ventaja de la distancia de Mahalanobis es el que toma en cuenta la dispersión entre las variables y la correlación entre ellas. Es un valor independiente de la escala, pues permanece inalterada si las variables se multiplican por una constante o son reemplazadas, por ejemplo, por componentes principales.

Otras distancias conocidas y regularmente usadas son las distancias de Manhattan, de Minkowski, de Canberra, de Bhattacharyya, las cuales pueden verse como variaciones de las anteriores.[6]

----

6. Ver, por ejemplo, Mardia, p. 371.

4.2.2 METODOS JERARQUICOS.

A partir de cualquiera de las nociones de distancia vistas en el apartado anterior se puede construir una matriz de distancias entre los elementos, objetos o unidades de análisis. Los métodos jerárquicos encuentran los conglomerados teniendo como criterio un ordenamiento sucesivo de la magnitud de los diferentes niveles en que se establecen estas distancias.

En esta búsqueda de agrupaciones se puede tener en cuenta como punto de partida, por ejemplo, la distancia mínima entre todas las entradas de la matriz de distancias. Esto se traduce en que los elementos a los cuales corresponda la mínima distancia hallada forman el nivel más bajo de agrupación -es decir, la agrupación más "estrecha"- a esa específica altura de distancia. Con la subsiguiente distancia mínima entre todas las restantes se conforma otro nivel de agrupación de los elementos incluidos, del cual forma parte el primer grupo ya encontrado y así se continúa el procedimiento hasta obtener el último nivel de la distancia o, lo que significa lo mismo, la agrupación más grande. El nombre de método jerárquico obedece a que los subgrupos que se van formando están insertos en el siguiente del nivel superior, es decir, se van anidando paulatinamente.

De ordinario el proceso de encontrar estos conglomerados se acompaña de la construcción de un diagrama de árbol en el cual se van bosquejando los distintos niveles de agrupación consecutivamente encontrados en las diferentes distancias. El diagrama de árbol recibe el nombre de dendograma y en él se acostumbra a representar en el eje de las abscisas, los elementos o unidades y en el eje de las ordenadas, los niveles de las distancias.

La técnica anterior para encontrar cúmulos y dendograma se denomina de distancia minima y se destaca en este algoritmo: primero, que el encadenamiento entre los elementos a que da lugar puede hacer perder el proceso de aglomeración en torno a un núcleo central y, segundo, que cualquier transformación lineal de las variables iniciales no afecta la forma general de presentación de los conglomerados. Mientras esto último es una propiedad deseable, la primera puede desvirtuar la composición de las agrupaciones.

Es posible obtener los conglomerados y el dendograma respectivo comenzando con la distancia máxima presente en la matriz de distancias. El algoritmo es como el anterior, pero en sentido inverso: se van encontrando sucesivamente las agrupaciones de acuerdo a la mayor distancia restante. Con esta técnica no hay propiamente un efecto de encadenamiento, como en el caso precedente y el dendograma, por consiguiente, presenta una forma distinta.

La gráfica No. 4.3 muestra los dendogramas obtenidos por los dos métodos para una misma información.[7] En el dendograma superior se nota de inmediato la presencia del efecto de encadenamiento y la distinta estructuración de los grupos a los niveles "micro" con respecto al dendograma inferior. Desde un punto de vista "macro" hay coincidencia entre los dos conjuntos fundamentales. Es conveniente mencionar que el método de distancia mínima puede no proveer soluciones útiles por su sensibilidad al ruido presente en relación a los diferentes conglomerados y al subsecuente efecto de encadenamiento. Por otra parte, el método de la distancia máxima ofrece agrupaciones compactas pero no necesariamente garantiza encontrar todos los grupos allí donde las distancias entre ellos son menores que algún valor dado. Puede intentarse como alternativa de búsqueda de cúmulos una distancia promedio entre las dos anteriores.

----

7. Tomado de Mardia, pgs. 373 y 375.

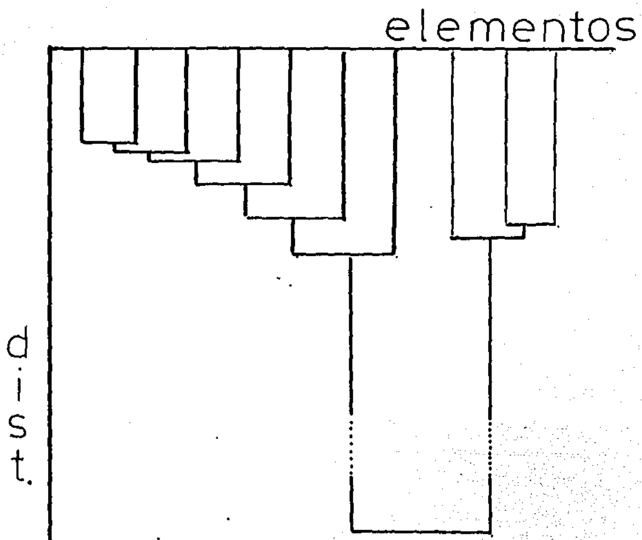
Sobre los métodos de optimización a través de modelos probabilísticos hay que destacar, adicionalmente, que por los fuertes supuestos implícitos pueden obtenerse conclusiones espurias cuando éstos no se cumplen. Estos procedimientos de optimización se sesgan hacia la búsqueda de conglomerados elípticos y esféricos. Si la información contiene conglomerados de distinta forma, éstos no pueden ser encontrados y la solución resultante confundirá de hecho la presencia de las agrupaciones originales.

Con los métodos jerárquicos el gran problema es la imposición de una subordinación que no tiene por qué estar presente en la información inicial. La técnica de dendogramas, por último, puede dar lugar a una considerable pérdida de información. La recomendación, por todas estas limitaciones, es una combinación de técnicas y de gráficas que vaya insinuando los conglomerados que se presentan en los datos originales. Recordemos que el propósito es esencialemente descriptivo.

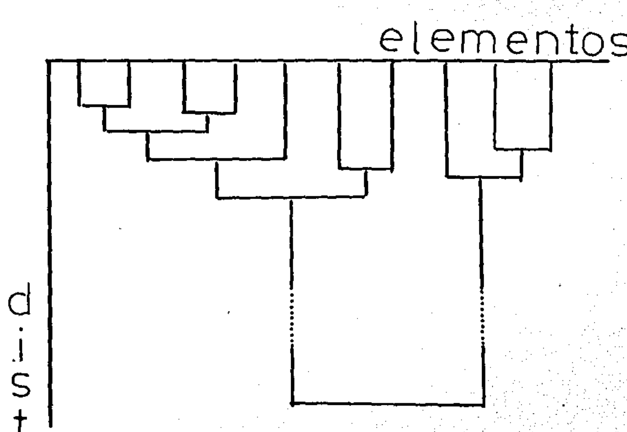
Desde luego, la presencia de cúmulos o conglomerados debe detectarse desde la informacióon objetiva. Esta aclaración no sobra porque es posible inducir agrupamientos cuando éstos no existen en la realidad, o despreciar cúmulos que forman parte integrante de los datos iniciales. Los problemas de interpretaciún son, pues, muy difíciles y la definición de los grupos que se encuentran es esencialmente un asunto donde la interpretación teórica y los aspectos analíticos deben conjugarse muy finamente con el instrumental estadístico de la presente técnica. La conformación de los conglomerados combina dos aspectos: uno de naturaleza objetiva y otro de carácter subjetivo. Este último muy importante por cuanto es el elemento guía del análisis. Recordemos que si los datos experimentan objetivamente agrupamientos distintos a los circulares o elípticos es muy difícil extraerlos a través de la técnica. Es en este punto en el que la tarea de reorganizar la información necesita de un cierto "olfato" y habilidad para que sin tergiversar el objeto de investigación pueda ser capaz de coadyuvar en el descubrimiento de los conglomerados.

GRAFICA No. 4.3

Dos tipos de dendogramas.



Dendograma de distancia mínima. Obsérvese el efecto de cadena y el anidamiento consecutivo que se va generando



Dendograma de distancia máxima. Los grupos más grandes se mantienen, pero los pequeños van a ser algo diferentes.

4.3 CORRELACION CANONICA.

El objetivo de la correlación canónica -C.C.- es encontrar las combinaciones lineales entre dos grupos de variables, destacando la estructura de correlación entre el juego de combinaciones lineales realizadas. Teniendo dos grupos de variables -**X** y **Y**- se busca asociarlas mediante combinaciones lineales detectando a cuál par de ellas corresponde las más alta correlación. A esta última se le llama primera correlación canónica y al par de combinaciones lineales asociadas de las **X** y las **Y** se les denomina las primeras variables canónicas. De manera similar se definen las restantes correlaciones y variables canónicas, cada una de ellas no correlacionada con las anteriores y posteriores y con la característica de tener la correlación más alta entre las variables canónicas remanentes.

En otras palabras, el proceso de análisis canónico es equivalente al desarrollo de un nuevo sistema de coordenadas en el espacio de cada conjunto de variables de tal forma que ese nuevo sistema presente ordenadamente la estructura de correlaciones. Se trata entonces, de encontrar las combinaciones lineales de las variables en cada conjunto con el objeto de que la primera tenga la facultad de expresar la correlación máxima de entre todas las combinaciones lineales posibles. Una segunda combinación de los dos conjuntos de variables iniciales reflejaría la segunda mayor correlación entre las restantes con la propiedad de no estar correlacionada con la primera ni con las últimas. Y así se sigue hasta culminar el juego completo de combinaciones lineales.

La primera combinación lineal representa las nuevas coordenadas en el cambio del sistema y la continuación del proceso implica la especificación completa del nuevo sistema de coordenadas. Se siente, en esta presentación, la cercana familiaridad de este procedimiento con la obtención de los componentes principales; la diferencia sustancial es que en los C.P. se consideran las interrelaciones dentro de un conjunto de variables, mientras que en C.C. se subrayan las interrelaciones entre dos grupos de variables.

Es notoria también una similaridad entre la C.C. y el modelo de regresión múltiple en el sentido de que en la segunda tenemos una variable Y expresada en función de un conjunto de variables X. En C.C. no hay una sola Y, sino un grupo de variables Y y de ahí que se le considere a la C.C. como la extensión multivariada de la regresión múltiple. No obstante, en C.C. no hay -o no debiera haber- una diferenciación entre variables explicadas y explicatorias o entre dependientes o dependientes, los dos grupos de variables se tratan simétricamente.

Veamos en cuál es la lógica de la C.C.: considerando dos conjuntos de variables aleatorias, $X_1, X_2, .., X_p$ y $Y_1, Y_2, .., Y_q$ con $q < p$ y medias y varianzas dadas, podemos expresar la varianza de los vectores $\mathbf{x}$ y $\mathbf{y}$ como:

$VAR[\mathbf{x}] = V_{11}$, además

$VAR[\mathbf{y}] = V_{22}$ y

$COV[\mathbf{x}, \mathbf{y}] = V_{12} = V'_{21}$ -->

Si contamos con dos vectores de coeficientes $\mathbf{a}$ y $\mathbf{b}$, podemos expresar las dos combinaciones lineales siguientes:

$U = \mathbf{a}'\mathbf{x}$ y

$W = \mathbf{b}'\mathbf{y}$ con

$\mathbf{a}' = [a_1 \ldots a_p]$ y

$\mathbf{b}' = [b_1 \ldots b_p]$.

y la correlación entre estas dos nuevas combinaciones lineales:

$$R_{U,V} = \frac{a' \ V_{12} \ b}{[a'V_{11}a \ \ b'V_{22}b]^{1/2}} \qquad (4.3.1)$$

Donde $a'V_{11}a$, $b'V_{22}b$ y $a'V_{12}b$ son respectivamente las matrices de covarianzas de $a'x$, $b'y$ y de $a'x$ con $b'y$. Como la correlación es independiente de la escala, se puede asumir que $a'V_{11}a$ y $b'V_{22}b$ son iguales a la unidad, lo que es muy útil para propósitos de simplificación.

El problema es entonces maximizar la correlación entre **U** y **W** lo cual significa encontrar los vectores **a** y **b** que hagan máxima la mencionada correlación. Usando los multiplicadores de Lagrange, se obtiene un máximo cuando las derivadas parciales de la función anterior con respecto a **a** y a **b** se igualen a cero:

$$[a'V_{12}b]^2 - \gamma[a'V_{11}a - 1] - t[b'V_{22}b - 1] = 0 \quad (4.3.2)$$

lo cual arroja que $\gamma = t = [a'V_{12}b]^2$.

Esto significa que el multiplicador de Lagrange es el máximo valor del cuadrado de la correlación entre $a'x$ y $b'y$. El proceso de solución del sistema definido por (4.3.2) permite establecer que γ es la raiz característica de la matriz :

$$V_{12}[V_{22}]^{-1}V_{21}[V_{11}]^{-1} \quad (4.3.3)$$

El segundo par de variables canónicas $-V_2$ y W_2- se escoge de manera tal que maximice $a'_2 V_{12} b_2$ con las restricciones de varianza unitaria y de no correlación con las primeras variables canónicas. De hecho se maximiza respetando las condiciones de ortogonalidad:

$$COV[U_1, U_2] = COV[W_1, W_2] = COV[W_1, V_2] = 0.$$

o lo que es lo mismo:

$$a'_1 V_{11} a_2 = b'_2 V_{22} b_2 = b'_1 V_{12} a_2 = 0.$$

Donde, después de cierto manejo matemático, se encuentra a γ_2 -el cuadrado de la segunda correlación canónica- como la segunda raíz característica de (4.3.3) con b_2 y a_2 como sus vectores propios asociados. El proceso se sigue hasta encontrar las (q-2) variables y correlaciones canónicas faltantes.

Conclusión: las relaciones entre **X** y **Y** pueden ser presentadas por las correlaciones entre las q pares de variables canónicas. Esto significa que se han reducido las p*q covarianzas presentadas en V_{12} a sólo q correlaciones canónicas.

El procedimiento también puede realizarse a través de las matrices de correlación y sus resultados son los mismos para las correlaciones canónicas, no asi para los coeficientes de las variables canónicas, pues éstos varian en consonancia a los cambios en la matrices de las que se extrajeron los valores y vectores propios. Se pueden realizar pruebas estadísticas a condición de que las variables en **X** y en **Y** tengan una distribución normal multivariada. No se entrará en mayores detalles respecto a las pruebas de hipótesis por la dureza misma de los supuestos que para esta tarea se exigirian. Adicionalmente, el propósito que estamos enfatizando en la C.C. es básicamente descriptivo.

CONCLUSIONES.

Decíamos en la introducción que la inquietud que dio origen a este ensayo surgió de preguntarse acerca de la posibilidad de combinar los M.E.M. con la investigación económica. Esta primera pregunta se partía en varios interrogantes: cuál es la naturaleza de los métodos multivariados ? Qué problemas especiales se presentan para su aplicación en el campo de la economía ? Cuáles aplicaciones han tenido mayor eco en los estudios económicos en México ?

El primer interrogante se resolvió palabras más, palabras menos, en los siguientes términos: dentro de los M.E.M. podemos ubicar dos grandes dimensiones, las cuales responden a propósitos y objetivos que se deben delimitar para evitar confusiones. Una primera es la dimensión descriptiva, entendida como una exposición alternativa, organizada y sintética de la información original, en donde se destaca bien una perspectiva distinta -caso de componentes-, bien características orgánicas básicas -análisis de cúmulos- , bien las asociaciones de los fenómenos en estudio -correlación canónica-. Concebimos entonces esta dimensión descriptiva como el resumen y síntesis de la información bruta de tal forma que se destaquen las principales características del objeto en estudio.

La segunda dimensión es la intención modelizadora y corresponde a la pretensión de ajustar la realidad a un predeterminado esquema funcional -en el sentido matemático- de comportamiento. No se trata entonces de descubrir en la realidad las características relevantes de la información -tarea a la que responde el espíritu descriptivo- sino que la meta es otorgarle a la realidad la particular explicación del modelo, se trata de concebir lo real desde un singular enfoque modelístico.

En la discusión sobre el método se cuestionó la posición axiomática y empirista de ciertas escuelas de pensamiento económico. Más precisamente las primeras versiones neoclásicas y las corrientes monetaristas modernas se encuentran impregnadas de esta forma de generación de conceptos y categorías. En ellas se incrusta, como en ninguna otra escuela, esa concepción de la matemática como el ápice de la pureza científica.

Entre los mejores ejemplos que ilustran el axiomatismo en la teoría económica tenemos todos los capítulos que en los textos de microeconomía se han heredado de los neoclásicos y que reciben el nombre de "Conducta del Consumidor": tanto las concepciones marginalistas pioneras, los planteamientos basados en las curvas de indiferencia y las restricciones de ingresos al igual que las teorías de las preferencias reveladas, comparten esa subordinación de lo real a las frías exigencias de los axiomas.

Entre lo más representativo de la confusión entre correlación y causalidad encontramos las teorías monetaristas actuales. Ese camino que va de la correlación a la causalidad se inaugura dentro del campo mismo de la estadística en las afirmaciones sobre el carácter espurio o no de la correlación. Recordemos que se definiía una correlación espuria cuando aspectos importantes de un análisis estadístico eran dejados de lado. De esta manera, al involucrarlos, el elemento espurio dejaba de serlo y el proceso en estudio no necesitaba de aspectos adicionales para su comprensión. Este es el motivo por el cual no es un lugar común advertir sobre el error de descubir causalidad cuando se corrobora una buena correlación.

Pero los atisbos de causalidad también se invocan ante la certeza de una predicción. En el terreno estadístico predicción y correlación van casados: una buena correlación posibilita que la magnitud de algunas variables pueda definirse de antemano. De igual manera una predicción aceptable se traduce hacia atrás en una asociación lineal más o menos fuerte entre las variables involucradas. Esta simetría en la predicción-correlación tiende a conjugarse, a concretarse como "explanans" teórico. A pesar de la simplicidad del enfoque, es de los más usados y buena parte del achacar a las relaciones lineales responsabilidades causales se debe al punto de vista metodológico sobre la formación de conceptos y categorías. Creemos que éste es precisamente el meollo: consiste en prestar y arrebatar a las consideraciones y resultados estadísticos conclusiones teóricas.

Es en este ambiente que se juzgan las aplicaciones que se trataron en el análisis factorial. En los estudios criticados un rasgo común es esa pretensión de apropiarse de la teoría por obra y gracia del "descubrimiento" de los factores subyacentes de la realidad. Esas atribuciones explicativas de encontrar los substratos de la realidad de las que presume el modelo de análisis de factores fueron aceptadas abierta o soterradamente en cada una de las investigaciones mencionadas. El cuestionamiento fundamental en cada uno de ellos tuvo a ésta como la referencia más inmediata.

Se constata que ese censurable ánimo modelizador es una intentona de generar teoría a partir de un modelo matemático explicativo: no se trata entonces de darle sustento empírico a un planteamiento, se trata de descubrir y elaborar ese planteamiento. Y este "logro", además, se escapa de las buenas intenciones del investigador, pues depende de lo que él considere como la convicción teórica del estudio, del modelo estadístico, de los problemas a que se enfrenta y con lo que espera de la investigación.

Una segunda inquietud llama a entender y respetar las condiciones, supuestos y lógica interna de las técnicas estudiadas. La irreverencia a estas reglas conduce a equivocados descubrimientos y desbroza el camino para trastocar el desarrollo de categorías con los resultados de la puesta en práctica de una técnica estadística. Si bien es cierto que es también lugar común el llamado a respetar condiciones y supuestos, la omisión a este precepto persiste porque se desconoce la técnica particular, se desvirtúa su popósito o se carece de convicciones teóricas propias que orienten el proceso de investigación. Estos tres últimos aspectos permiten que se vayan colando con rapidez las violaciones tanto a los supuestos como a la interpretación, forzando la tarea extracción de conlusiones y son responsables de los "descuidos" aludidos, por lo menos en lo que atañe a las aplicaciones que se han estudiado en este material.

Creemos que en las investigaciones con un enfoque teórico marxista se ha descuidado el análisis cuantitativo de la información. En lo referente a métodos multivariados como los componentes principales, análisis de cúmulos y correlación canónica hay un abundante arsenal de corte descriptivo, cuya utilidad depende del conocimiento y difusión de estas técnicas. Se insiste, de nuevo, en profundiar en los objetivos, las condiciones, los supuestos y los resultados que definen a cada uno de los procedimientos multivariados. En cambio, por los problemas ya criticados, el análisis de factores y, un poco menos, el análisis discriminante, no tendrían una buena cabida dentro de una concepción marxista. Por las características de que está dotada, asumimos que investigaciones con esta concepción, pueden enriquecerse en gran medida por el tratamiento multivariado de la información empírica.

En dos ámbitos consideramos que los procedimientos multivariados son de especial provecho. Uno, cauando se apela a una cierta técnica con el ánimo de apoyar empíricamente un análisis teórico. En este contexto la selección de la técnica responde a los requerimientos del problema planteado, y no al contrario, como tanto se ha insistido. Otro, por la particular apreciación que pueda ofrecer una visión alternativa de organización y presentación de la información. Esta forma distinta de plasmar los mismos hechos puede sugerir a la investigación nuevas modalidades, diferentes comportamientos, diversos nexos o, incluso, otras preguntas. Pero estos dos ámbitos no chocan entre sí, pueden abordarse simultáneamente, ser sustento uno del otro, privilegiar alguno etc. No hay una regla fija.

En nuestro medio los M.E.M. no son muy conocidos. La mayor prueba de esta afirmación es la aguda y casi absoluta escasez de estudios en los que se apliquen las relativamente nuevas técnicas estadísticas. Desafortunadamente el procedimiento más usado es el más cuestionable de todos -el análisis factorial-, lo que obedece a lo ambicioso y seductor de sus propósitos. De otro lado, buena parte del problema para las aplicaciones de los M.E.M. se debe a los complicados, tediosos y largos procedimientos de cálculo, los cuales hacen imposible efectuarlos con calculadoras de bolsillo. Se requiere un mínimo equipo computacional para poder llevarlos a cabo. Es muy posible que la difusión de computadoras extienda el uso de las técnicas nombradas.

Apéndice B.

APENDICE MATEMATICO

B.1 TRATAMIENTO ESTADISTICO.

Si denotamos a x_i como la i-ésima observación de la variable **x**, la media de las **n** observaciones de esa misma variable la representamos por **x.**:

$$\mathbf{x.} = [1/n]\ \Sigma_j x_j, \text{ con } j = 1..n \quad (5.1.1)$$

Esta última expresión la podemos representar matricialmente por:

$$\mathbf{x.} = [1/n]\ \mathbf{A}'\mathbf{X} \quad (5.1.2)$$

Donde:
A':vector de (1xn) donde cada uno de sus elementos es la unidad.
X :vector (nx1) de observaciones de la variable x.
Desglosando (5.1.2): $\mathbf{x.} = 1/n\ [1\ 1...1]\ [x_1\ x_2\ ...\ x_n]'$

Ahora bien, si tenemos **p** variables con **n** observaciones de cada una de ellas y deseamos obtener la media de cada variable, obtenemos un vector de medias para cada una de las **p** variables:

$$\mathbf{X}.' = [1/n]\ \mathbf{A}'\mathbf{X} \quad (5.1.3)$$

Donde :
A': sigue siendo el vector definido en (5.1.1)
X: matriz (nxp) de observaciones
X.': vector (1xp) en el cual sus **p** elementos son cada uno las medias de las variables.

Si llamamos a $x_{1.}, x_{2.}, \ldots, x_{p.}$ como las medias de las **n** observaciones para cada una de las **p** variables, la expresión (5.1.3) se puede desglosar de la manera siguiente:

$$\mathbf{X.'} = 1/n\ [1\ 1\ 1 \ldots 1] \begin{bmatrix} x_{11} & x_{21} & \ldots & x_{p1} \\ x_{12} & \ldots & \ldots & x_{p2} \\ \vdots & & & \vdots \\ x_{1n} & \ldots & \ldots & x_{pn} \end{bmatrix} = [x_{1.}, \ldots, x_{p.}]$$

Ahora bien, si a cada uno de los elementos de la matriz **X** le restamos su respectiva media, obtenemos una nueva matriz **X** pero en forma de desviaciones. Veamos paso a paso qué es lo que queremos decir:

$$\begin{bmatrix} x_{11} & x_{21} & \ldots & x_{p1} \\ x_{12} & \ldots & \ldots & x_{p2} \\ \vdots & & & \vdots \\ x_{1n} & \ldots & \ldots & x_{pn} \end{bmatrix} - \begin{bmatrix} x_{1.} & x_{2.} & \ldots & x_{p.} \\ x_{1.} & x_{2.} & \ldots & x_{p.} \\ \vdots & & & \vdots \\ x_{1.} & x_{2.} & \ldots & x_{p.} \end{bmatrix}$$

Esta diferencia de matrices la podemos representar por:

$$\begin{bmatrix} x_{11}-x_{1.} & x_{21}-x_{2.} & \ldots & x_{p1}-x_{p.} \\ x_{12}-x_{1.} & \ldots & \ldots & x_{p2}-x_{p.} \\ \vdots & & & \vdots \\ x_{1n}-x_{1.} & \ldots & \ldots & x_{pn}-x_{p.} \end{bmatrix} = X$$

Es decir, los elementos en la matriz **X** en forma de desviaciones son:

$$\begin{bmatrix} X_{11} & X_{21} & \ldots & X_{p1} \\ X_{12} & \ldots & \ldots & X_{p2} \\ \vdots & & & \vdots \\ X_{1n} & \ldots & \ldots & X_{pn} \end{bmatrix}$$

Al premultiplicar esta matriz **X** por su transpuesta, encontramos la matriz $\mathbf{X'X}_{p\times p}$ de sumas de cuadrados y productos cruzados:

$$\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & \dots & \dots & X_{2n} \\ \vdots & & & \vdots \\ X_{p1} & \dots & \dots & X_{pn} \end{bmatrix} \begin{bmatrix} X_{11} & X_{21} & \dots & X_{p1} \\ X_{12} & \dots & \dots & X_{p2} \\ \vdots & & & \vdots \\ X_{1n} & \dots & \dots & X_{pn} \end{bmatrix}$$

$$= \mathbf{X}'\mathbf{X} \quad (5.1.4)$$

Esto último significa que los elementos de la matriz **X'X** son los que siguen:

$$\begin{bmatrix} \Sigma[X_1]^2 & \Sigma X_1X_2 & \dots & \Sigma X_1X_p \\ \Sigma X_2X_1 & \Sigma X_{22} & \dots & \Sigma X_2X_p \\ \vdots & \vdots & & \vdots \\ \Sigma X_pX_1 & \Sigma X_pX_2 & \dots & \Sigma[X_p]^2 \end{bmatrix}$$

Es decir, por ejemplo, los elementos
$\Sigma[X_1]^2=\Sigma[x_{1j}-x_{1.}]^2$ y $\Sigma X_3X_2=\Sigma[x_{3j}-x_{3.}][x_{2j}-x_{2.}]$

Concluimos que en la diagonal principal de la matriz **X'X** se acomodan las sumas de cuadrados de las desviaciones de la respectiva variable con respecto a su propia media y fuera de la diagonal encontramos la sumatoria de los productos cruzados de las desviaciones de las variables. Como se observa, la matriz **X'X** es simétrica.

Todas estas transformaciones tienen razón de ser para encontrar la matriz de varianza-covarianza de las variables x:

$$\mathbf{COV}[\mathbf{X}] = \mathbf{S} = 1/[n-1]\ \mathbf{X}'\mathbf{X}. \quad (5.1.5)$$

En otras palabras, al multiplicar la matriz **X'X** por el escalar **1/[n-1]** obtenemos en la diagonal principal la varianza de las variables y fuera de la diagonal las covarianzas entre las variables.

Ahora bien, si denotamos por S_{ij} a la covarianza de la i-ésima variable con la j-ésima variable, entonces S_{ii} seria la covarianza de la variable con respecto a sí misma, es decir la varianza de la i-ésima variable. De ahi que el elemento **ij** de la matriz de correlaciones lo podemos representar por:

$$R_{ij}=\frac{S_{ij}}{[S_i S_j]^{1/2}}$$

Los elementos de la diagonal principal de la matriz de correlaciones serán igual a la unidad. La matriz **R** quedaría así:

$$\begin{bmatrix} R_{11} & R_{12} & . \; . \; . & R_{1p} \\ R_{21} & R_{22} & . \; . \; . & R_{2p} \\ \vdots & \vdots & & \vdots \\ R_{p1} & R_{p2} & . \; . \; . & R_{pp} \end{bmatrix}$$

Donde $R_{11}=R_{22}=. \; . \; .= R_{pp}=1$.

B.2 DIFERENCIACION EN MATRICES.

Teniendo los vectores $A'_{1\times n}$ y $X_{n\times 1}$, entonces de la expresión $y=A'X$ podemos obtener el vector de derivadas parciales con respecto a **X**:

$$\left[\frac{Dy}{DX}\right]=\left[\frac{dy}{dx_1},\frac{dy}{dx_2},\ldots,\frac{dy}{dx_n}\right]' \qquad (5.2.1)$$

Lo cual quiere decir que la derivada parcial de **y** con respecto al vector **X** es igual al vector **A**

Por ejemplo, la función: $y=2x_1+3x_2+5x_3$ se puede escribir:

$y=A'X=[2 \; 3 \; 5] \; [x_1 \; x_2 \; x_3]'$ entonces

$$\frac{Dy}{DX} = [2\ 3\ 5]'$$

De otra parte, el vector de derivadas parciales de la forma:

$z=X'AX$ es

$2AX$ (5.2.2

En nuestro ejemplo: $z=2[x_1]^2+2x_1x_2+3[x_2]^2$.

$z=[x_1\ x_2]\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$ Entonces su derivación es:

$$\frac{Dz}{DX} = 2\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}4x_1+2x_2\\ 2x_1+6x_2\end{bmatrix}=\begin{bmatrix}4 & 2\\ 2 & 6\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}.$$

Y estos resultados son de utilidad cuando se trata de encontrar el máximo de una función expresada matricialmente.

B.3 MULTIPLICADORES DE LAGRANGE.

En los procesos de maximización y minimización de funciones matemáticas, a menudo se incluyen restricciones para obtener el máximo o el mínimo de una función dada. El procedimiento más acostumbrado para obtener máximos y mínimos de funciones sometidas a restricciones de igualdad es el método de multiplicadores de Lagrange. Consiste en que la maximización o minimización de una función particular $G(x,y)$ sometida a la restricción dada por la función $H(x,y)=0$, pueda plantearse de la siguiente manera:

$$F(x,y,\gamma)=G(x,y)-\gamma H(x,y) \quad (5.3.1)$$

En la cual γ es el multiplicador de la función de restricción. La diferenciación de la función F respecto a las tres incógnitas da lugar a tres ecuaciones que igualadas a cero pueden resolverse para **x**, **y** y γ. La extensión de los multiplicadores de Lagrange para una función de **p** variables $F[x_1 \ldots x_p]$ sometida a **k** restricciones $H_j[x_1 \ldots x_k]=0$ donde $j=1..k$ y $r < p$ la representamos:

$$\bar{F}[x_1 \ldots x_p; \gamma_1 \ldots \gamma_k]=G[x_1 \ldots x_p]-\Sigma_j \gamma_j H_j[x_1 \ldots X_k]$$

Y la diferenciación parcial arroja **p+k** ecuaciones que se pueden resolver con igual número de incógnitas. En la matemática multivariada es usual maximizar funciones sujetas a restricciones como la de que la norma de un vector tiene que ser igual a la unidad. Si al vector $A=[a_1..a_2]$ se le exige norma igual a uno, esto significa que

$$\mathbf{A}'\mathbf{A}=[a_1]^2+[a_2]^2=1 .$$

Restricción que en la mayoría de los casos es útil para eliminar indeterminaciones. Un ejemplo típico de multiplicadores de Lagrange se encuentra en el proceso de obtención de los componentes principales. En este mismo apéndice se ilustra en la siguiente sección el proceso de obtención de valores característicos a partir de los multiplicadores de Lagrange.

B.4 VALORES Y VECTORES CARACTERISTICOS.

El problema de encontrar los valores -y vectores- característicos o latentes o primitivos o propios o eigenvalores de una matriz simétrica **B** se define como el proceso de encontrar un escalar γ y un vector asociado $\mathbf{a} = \mathbf{0}$ tal que cumpla con:

$$\mathbf{B}\mathbf{a}=\gamma\mathbf{a} \quad (5.4.1)$$

Al escalar γ le damos el nombre de valor característico y al vector **a** lo llamamos el vector característico asociado al γ encontrado. A la solución de la ecuación matricial $[\mathbf{B}-\gamma\mathbf{I}]\mathbf{a}=0$ en la cual se involucra un escalar γ desconocido, cuyo valor debe ser tal que $[\mathbf{B}-\gamma\mathbf{I}]$ sea una matriz singular, la llamamos la solución de la ecuación característica de la matriz **B**. Encontrar los valores γ que satisfacen la ecuación es encontrar los valores propios o eigenvalores o valores característicos de la matriz **B**. Veamos un ejemplo de obtención de valores y vectores propios de una matriz simétrica:

Si la matriz **B** es: $\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$

y queremos buscar un vector **a** que cumpla con (5.4.1), definimos una función Φ_1 de tal manera que podamos escribirla así:

$$\Phi_1=4[a_1]^2+4a_1a_2+[a_2]^2-\gamma[[a_1]^2+[a_2]^2-1] \quad (5.4.2)$$

Donde γ es el multiplicador de Lagrange y se ha hecho uso la restricción $[a_1]^2+[a_2]^2=1$, restricción que significa que la norma del vector es igual a la unidad. Matricialmente, la función anterior la podemos expresar como aparece en (5.4.3):

$$\Phi=\mathbf{a}'\mathbf{B}\mathbf{a}-\gamma[\mathbf{a}'\mathbf{a}-1] \quad (5.4.3)$$

La derivada parcial de (5.4.3) con respecto al vector **a** e igualada a cero es:

$$2\mathbf{B}\mathbf{a}-2\gamma\mathbf{a}=\mathbf{0}.$$

O sea que $\mathbf{B}\mathbf{a}-\gamma\mathbf{a}=\mathbf{0}$.

por lo tanto: $[B-\gamma I]a=0$ (5.4.4)

Donde I es la matriz identidad.

Para encontrar los valores γ diferentes de cero, es necesario que la matriz

$[B-\gamma I]$ sea singular.

De esta manera, tenemos entonces que:

$$\begin{vmatrix} 4-\gamma & 2+0 \\ 2+0 & 1-\gamma \end{vmatrix} = (4-\gamma)(1-\gamma)-4 = \gamma^2-5\gamma = 0$$

Resolviendo este polinomio de segundo grado arroja que $\gamma_1=5$ y $\gamma_2=0$

Asociado con cada valor propio tenemos un vector característico a de la ecuación (5.4.4). Encontremos entonces ese vector asociado, primero con γ_1:

$$[B-\gamma I]a= \begin{bmatrix} 4-\gamma_1 & 2 \\ 2 & 1-\gamma_1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

lo cual significa que:

$$-a_1 + 2a_2 = 0$$
$$2a_1 - 4a_2 = 0$$

Usando la restricción de que $a'a=1$, encontramos que aproximadamente $a_1=.8944$ y $a_2=.4472$. De manera similar se procede con γ_2 y se encuentra que $a_1=.4472$ y $a_2=-.8944$.

Podemos entonces concretar a (5.5.1) :

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} .8944 \\ .4472 \end{bmatrix} = 5 \begin{bmatrix} .8944 \\ .4472 \end{bmatrix}$$

Y encontramos que el vector $A_1=[.8944 \quad .4472]'$ es el vector característico asociado al primer valor propio $\gamma_1=5$.

Adicionalmente los vectores A_1 y A_2 asociados con dos distintos eigenvalores γ_1 y γ_2 de una matriz simétrica son mutuamente ortogonales, es decir, que $A_1'A_2=0$. La ortogonalidad la podemos interpretar como que los vectores son linealmente independientes y perpendiculares entre sí.

Si en lugar de trabajar con la matriz B, lo hacemos con la matriz de covarianzas de las variables x, S, como se define en (5.1.5) que cumple con los requisitos exigidos para la matriz B, e incluimos la matriz A integrada en sus columnas por los vectores propios asociados a las raíces características de la matriz S, entonces la matriz A es una matriz de transformación ortogonal que representa una rotación rígida de las variables x.

De esta manera podemos establecer que

$$SA_1=\Gamma_1 A_1 \quad . \, .$$

y si consideramos todos los vectores A_1 en la matriz A, tendremos:

$$SA=A\Gamma \quad (5.4.5)$$

Donde Γ es una matriz cuya diagonal principal está compuesta por los sucesivos valores característicos y fuera de la diagonal los elementos son iguales a cero. Si premultiplicamos ambos miembros de (5.4.5) por la matriz **A'**:

$$\mathbf{A'SA}=\mathbf{A'A}\Gamma=\mathbf{I}\Gamma=\Gamma \quad (5.4.6)$$

Recordemos que si **A'A** es ortogonal y además los vectores que lo componen son de longitud unitaria, entonces **A'A** es ortonormal. Del resultado (5.4.6) se obtiene el desarrollo de los componentes principales presentado en el capítulo segundo.

Algunas propiedades de los valores característicos:

1.- Las raíces características de una matriz simétrica definida positiva son todas positivas.

2.- El producto de las raíces características de una matriz simétrica se igual al determinante de la matriz.

3.- La suma de las raíces características de una matriz simétrica es igual a la traza de la matriz.

4.- Para cualquier matriz simétrica definida positiva, las raíces características de $\mathbf{B}^{-1}$ son el recíproco de las raíces características de **B**. Los vectores propios de **B** y $\mathbf{B}^{-1}$ son idénticos.

Apéndice C.

BIBLIOGRAFIA.

1. ANDERSON, T.W.(1958). AN INTRODUCTION TO MULTIVARIATE STATISTICAL ANALYSIS. John Wiley and Sons, Inc. New York.

2. APPENDINI et AL.(1972). "Desarrollo Desigual en México, 1900 y 1962." en DEMOGRAFIA Y ECONOMIA, El Colegio de México. Vol VI, Número 1, 1972. P. 1 - 40. México, D.F.

3. _______________.(1977). "Construcción de zonas para una análisis de la estructura agraria en México: una aplicacion del análisis factorial." DOCUMENTOS DE TRABAJO DEL CEED No. 3. El Colegio de México. México, D.F.

4. APPENDINI, Kirsten.(1983)."La Polarización de la Agricultura Mexicana: un análisis a nivel de zonas agrícolas en 1970" en ECONOMiA MEXICANA. Serie Temática No. 1. Sr. Agropecuario: CIDE. Departamento de Economía. México, D.F.

5. ASTORI,Danilo.(1978). ENFOQUE CRITICO DE LOS MODELOS DE CONTABILIDAD SOCIAL. Siglo XXI. México.

6. BALIBAR, Etienne.(1979)."De Bachelard a Althusser: el concepto de 'corte epistemológico'." en LA FILOSOFIA Y LAS REVOLUCIONES CIENTIFICAS. Teoría y Praxis. Grijalbo. México.

7. BARNETT, Vic.(1981). INTERPRETING MULTIVARIATE DATA. John Wiley and Sons. Barnett, Ed. New York.

8. BLAUG, Mark.(1980). THE METHODOLOGY OF ECONOMICS. Cambridge University Press. Cambridge.

9. ____________(1976)."Kuhn versus Lakatos or paradigms versus research programmes in the history of economics." en METHOD AND APPRAISAL IN ECONOMICS. Edited by Spiro Latsis. Cambridge University Press. Cambridge.

10. BOLCH B. AND HUANG.(1974). MULTIVARIATE STATISTICAL METHODS FOR BUSINESS AND ECONOMICS. Prentice Hall. New Jersey.

11. CAÑEDO et AL.(1977)."La Mortalidad y su relación con factores sociales, económicos y culturales." en SALUD PUBLICA DE MEXICO. Epoca V, Vol. XIX, Número 6. Nov-dic. México D.F.

12. CUE MANCERA, Agustín.(1983)."Matemáticas y Economía." en ECONOMIA INFORMA No.103. Facultad de Economía. UNAM.

13. DESAI, Meghnad.(1981). TESTING MONETARISM. Frances Pinter (Publishers) Ltda. London.

14. DOW, Sheila.(1985).MACROECONOMIC THOUGHT. A Methodological Approach. Basil Blackwell Ltd. Great Britain.

15. FEYERABEND, Larry.(1985). "Cómo Defender a la Sociedad Contra la Ciencia." en REVOLUCIONES CIENTIFICAS. Ian Hacking, compilador. Breviarios del F.C.E. México.

16. FEYNMAN, R.P.(1965). THE CHARACTER OF PHYSICAL LAW. MIT Press. Cambridge.

17. FRISCH, Ragnar.(1970). "De la Teoría Utópica a las Aplicaciones Prácticas: El Caso de la Econometría." en LECTURAS No. 25. FCE. México.(1978).

18. GIRI, Narayan.(1977). MULTIVARIATE STATISTICAL INFERENCE. Academic Press. London.

19. GOULD, S.J.(1986). THE MISMEASURE OF MAN. W.W. Norton and Company. London.

20. GREEN, P. and CARROL, D.(1976). MATHEMATICAL TOOLS FOR APPLIED MULTIVARIATE ANALYSIS. Academic Press. New York.

21. HACKING, Ian.(1985)."La Filosofía de la Ciencia Según Lakatos." en REVOLUCIONES CIENTIFICAS. Ian Hacking, compilador. Breviarios del F.C.E. México.

22. HAIR et Al.(1979). MULTIVARIATE DATA ANALYSIS (WITH READINGS). Petroleum Publishing Company. Oklahoma.

23. KANTOROVICH, Leonid.(1975)."Las Matemáticas en la Economía: Logros, Dificultades, Perspectivas." en LECTURAS No.25. FCE. México. (1978).

24. KENDALL, M.G.(1968). A COURSE IN MULTIVARIATE ANALYSIS. Hafnfer Publishing company, Inc. New York.

25. KRAMER, Clyde.(1972). A FIRST COURSE IN METHODS OF MULTIVARIATE ANALYSIS. Virginia Politechnic Institute. Virginia.

26. KUZNETS, Simon.(1971)."El Crecimiento Económico Moderno: Hallazgos y Reflexiones." en LECTURAS No. 25. FCE. México. (1978).

27. _______________(1979). INVESTIGACION CUANTITATIVA DEL CRECIMIENTO ECONOMICO. Ariel. Barcelona.

28. LAKATOS, Imre.(1985)."La Historia de la Ciencia y sus Reconstrucciones Racionales." en REVOLUCIONES CIENTIFICAS. Ian Hacking, compilador. Breviarios del F.C.E. México.

29. LATSIS, et al.(1976). METHOD AND APPRAISAL IN ECONOMICS. Edited by Spiro Latsis. Cambridge University Press. Cambridge.

30. LEDERMAN, Walter.(1984). STATISTICS. HANDBOOK OF APPLICABLE MATHEMATICS. Vol. VI. Part B. John Wiley and Sons. Bristol.

31. LEFF Z., Enrique.(1979). "Hacia una biosociología del conocimiento científico." en LA FILOSOFIA Y LAS REVOLUCIONES CIENTIFICAS. Teoría y Praxis. Grijalbo. México.

32. LOYOLA, David.(1982). "Estadística Matemática y Economía: Reflexiones Teórico-Metodológicas." en DOCUMENTOS DE TRABAJO No.2. Serie Matemática Aplicada. CIDE. Noviembre. México.

33. MACFIE, A.L.(1955)."The Scottish traditioin in economic thought." en SCOTTISH JOURNAL OF POLITICAL ECONOMY No.21.

34. MARDIA, Kent.(1979). MULTIVARIATE ANALYSIS. Academic Press. London.

35. MARRIOT, F.H.C.(1974). THE INTERPRTATION OF MULTIPLE OBSERVATIONS. Academic Press. London.

36. MARTIN MORENO, Sergio.(1985)."Notas Sobre la Aplicación de la Matemática a la Investigación Económica." en INVESTIGACION ECONOMICA No. 174. UNAM. México.

37. ____________________(1987). "La hipótesis de la estructura dual de la industria: el caso de la econom ía mexicana" en ESTUDIOS ECONOMICOS. Vol. 2. No. 1. El Colegio de México. México, D.F.

38. MORRISON, D. (1967). MULTIVARIATE STATISTICAL METHOD. Mc Graw-Hill. New York.

39. NAFINSA.(1984). LA ECONOMIA MEXICANA EN CIFRAS. Nacional Financiera, S.A. Edición 1984. México, D.F.

40. OVERAL and KLETT.(1972). APPLIED MULTIVARIATE ANALYSIS. Mc Graw-Hill. New York.

41. PIATIER, Andre.(1967). ESTADISTICA Y OBSERVACION ECONOMICA. 2 T. Ariel. Barcelona.

42. POPPER, Karl.(1985)."La Racionalidad de las Revoluciones Cientificas." en REVOLUCIONES CIENTIFICAS . Ian Hacking, compilador. Breviarios del F.C.E. México.

43. SALAS, C. y VALLE, A.(1983)."El Uso de las Matemáticas en Piero Sraffa." Cuadernos de la División de Postgrado de la Facultad de Economía. U.N.A.M. México, D.F.

44. SANCHEZ V., Adolfo.(1980). FILOSOFIA DE LA PRAXIS. Grijalbo, tercera edición. México.

45. SHAIKH, Anwar.(1984)."Cuentas de ingreso nacional y categorias marxistas." en ECONOMIA: TEORIA Y PRACTICA. No. 4. U.A.M. México, D.F.

46. SRIVASTAVA, M.S. and KHATRI, C.G.(1979). AN INTRODUCTION TO MULTIVARIATE STATISTICS. North Holland. New York.

47. STOHS. M.(1983). "Uncertainty' in Keynes's General Theory: A rejoinder. HISTORY OF POLITICAL ECONOMY, 15 (Spring).

48. TAKEUCHI, K.(1982). THE FOUNDATIONS OF MULTIVARIATE ANALYSIS. Wiley Eastern Limited. New Delhi.

49. TINBERGEN, Jan.(1969). "El Uso de Modelos: Experiencia y Perspectivas." en LECTURAS No. 25. FCE. México.

50. TINTNER, Gerhard.(1967). ECONOMETRICS. John Wiley and Sons. Inc. New York.

51. VALLE BAEZA, Alejandro.(1984)."Una Nota sobre la Matematización de la Teoría Económica y la Docencia." en ENSAYOS No. 3. División de Estudios de Posgrado. Facultad de Economía. UNAM. México.

52. WIMSATT, W.C.(1981). "Robustness, reliability and overdetermination." en SCIENTIFIC INQUIRY AND THE SOCIAL SCIENCES. Ed. by Brewer and Collins. San Francisco.

CONTENIDO.