

B. NOTAS TEÓRICAS

B.1 ANÁLISIS FACTORIAL COMÚN

El análisis factorial es un nombre dado a una clase de métodos estadísticos multivariados cuyo propósito es reducir los datos en cantidad. Es también una técnica en la cual todas las variables son consideradas simultáneamente (Hair et.al., 1987).

El análisis de componentes principales es un método en el que los factores son basados en la varianza total. Con este análisis, las unidades (“1”) se usan en la diagonal de la matriz de correlación de las variables originales, lo cual implica que todas las varianzas son comunes o compartidas (Hair et.al., 1987). El método de componentes principales crea nuevas variables que son la combinación lineal de las variables originales y además son independientes entre sí (Terrádez, 2005).

Asimismo se puede considerar un análisis factorial clásico basado en el método de componentes principales, en el cual se habla de la “rotación”. Antes de seguir es importante notar que el análisis factorial común tiene las comunalidades en la diagonal de matriz de correlación de los datos, mientras que en componentes principales, la diagonal es la varianza máxima, es decir 1. El análisis factorial aquí presentado se realizó en base a la varianza máxima y no con las comunidades, por lo que el estudio factorial común sin rotación equivale al análisis de componentes principales.

Cuando la rotación no se lleva a cabo se alcanza el objetivo de reducir el número de datos pero la información obtenida probablemente no ofrezca una buena interpretación de las variables examinadas; por ello la rotación es recomendada porque simplifica la estructura de los factores (Hair et.al., 1987).

Hay dos métodos para realizar la rotación, el ortogonal y el oblicuo. Cuando el objetivo es utilizar los factores obtenidos en subsecuentes análisis estadísticos, se debería seleccionar una rotación ortogonal porque se elimina la colinealidad entre los factores. Sin embargo, si sólo se está interesado en obtener interpretaciones teóricas, el método oblicuo es útil debido a que es más realista empíricamente (Hair et.al., 1987).

En la práctica el objetivo de todos los métodos es simplificar las hileras o columnas de la matriz de factores para facilitar la interpretación. Por simplificar las hileras se entiende hacer tantos valores en cada hilera tan cercanos a cero como sea posible; en cambio al simplificar las columnas se hacen los valores de cada columna tan cercanos a cero como sea posible. Se han desarrollado tres métodos ortogonales: Quartimax, Varimax y Equimax: (Hair et.al., 1987).

Quartimax: Simplifica las hileras de una matriz de factores, pues se enfoca en rotar el factor inicial para que una variable tenga una carga alta en un factor y baja en otro.

Varimax: Simplifica las columnas. Éste es el método más usado porque parece dar una separación de factores más clara.

Equimax: No se concentra en simplificar hileras o columnas en específico, sino trata de lograr alguna de las dos.

Cabe mencionar que no hay reglas específicas para seleccionar una rotación ortogonal en particular, simplemente se utiliza la técnica que más se acople a las necesidades de un problema de investigación (Hair et.al., 1987).

*Interpretación de una salida de resultados en un análisis factorial de componentes principales:

Un análisis factorial común sin rotación corresponde a un análisis de componentes principales, y es de ahí donde podemos iniciar la interpretación de los resultados.

La cantidad de componentes principales extraídos son como máximo el número de variables originales analizadas; y para determinar el número de éstos a usar durante la investigación se emplean los *eigenvalues*, que son las varianzas de dichos componentes principales. De acuerdo con el criterio de Kaiser, se debe retener todo componente con un *eigenvalue* mayor a 1. También se puede definir el número de componentes principales según la cantidad de varianza de los datos explicada; por ejemplo, se retienen los componentes que juntos expliquen el 80% de la varianza total del fenómeno (esto se observa en la columna “*cummulative*” de la tabla de resultados). En general, no hay una regla que imponga un método determinado de selección de componentes principales, pues la decisión de cuántos usar depende del investigador. Por ejemplo, según el criterio de Kaiser, un primer componente principal explica aproximadamente el 80% de la variación total de los datos, sugiriendo que esta nueva variable explica adecuadamente la variabilidad total.

Después de saber cuántos componentes deben ser retenidos, éstos se rotan. Los *loadings* de la salida de resultados indican el peso relativo de una variable en cada componente. En otras palabras, el *loading*, o vector propio señala la correlación entre el factor y la variable original. A mayor valor absoluto de esta cifra, más importante es tal variable dentro de un cierto componente. De esta manera, los factores se nombran según las variables de mayor peso que éste agrupe.

Por último, habiendo ya analizado los resultados, definido el número de componentes a usar y rotado los factores, se emplean los *scores*¹ resultantes de la rotación como las nuevas variables encontradas para seguir con el estudio.

¹ Los *scores* son la evaluación de los factores en cada uno de los individuos (empleando los *loadings* estandarizados –coeficientes- y la variable original estandarizada).
Ejemplo, $\text{Factor1} = \text{coeficiente} * x1 + \text{coeficiente} * x2$

B.2 PANEL DE DATOS Y CORTE TRANSVERSAL

El corte transversal se entiende como un estudio estadístico que se realiza con la información de un conjunto de variables en un punto determinado en el tiempo (Glosario de Economía Mexicano, 2006); en cambio un panel de datos (también llamado longitudinal) es aquel que sigue una muestra de individuos a través del tiempo y provee múltiples observaciones en cada individuo de la muestra (Hsiao, 1989).

Para el caso de panel se puede considerar un conjunto de datos que contenga observaciones de n unidades de corte transversal (por ejemplo países) durante t periodos. Supongamos que cada observación contiene los valores de m variables de interés. Entonces, el conjunto de datos está formado por mnt valores (Cottrell, 2002).

Los datos deben ser ordenados por observación, de modo que la matriz de datos tendrá nt filas y m columnas. Existen dos posibilidades de ordenar las filas (Cottrell, 2002):

- Por unidad: La matriz de datos estaría compuesta de n bloques, cada uno con t filas. El primer bloque de t filas contiene las observaciones de la unidad 1 de la muestra para cada uno de los periodos; el siguiente bloque contiene las observaciones de la unidad 2 para todos los periodos; etc. Entonces, la matriz de datos parecería un conjunto de datos de series de tiempo construido verticalmente.
- Por periodo: La matriz de datos estaría compuesta por t bloques, cada uno con n filas. La primera de las n filas contiene las observaciones de cada unidad en el periodo 1. La matriz de datos sería un conjunto de datos de muestras de corte transversal acomodadas verticalmente.

También es importante mencionar que el panel de datos posee ventajas sobre el corte transversal porque aporta al investigador un gran número de datos, incrementando los grados de libertad y reduciendo la colinearidad entre las variables explicativas. Pero

principalmente, el panel permite analizar un número importante de cuestiones que no pueden ser estudiadas con corte transversal o series de tiempo (Hsiao, 1989).

EFFECTOS FIJOS

Se considera la siguiente ecuación:

$$Y_{it} = \beta X_{it} + a_i + e_{it} \quad (\text{B.1})$$

Donde a_i es un número fijo para cada individuo; es decir, que no varía en el tiempo, por ejemplo, el idioma oficial del país, el hecho de compartir o no frontera con un socio comercial, etc. El error, e_{it} , es generalmente llamado idiosincrásico o “*time-varying*” debido a que representa factores no observables que cambian con el tiempo y afectan a Y_{it} . (Wooldridge, 2002). Equivale a N modelos de regresión, uno para cada individuo, misma pendiente, y un intercepto específico para cada individuo (Sosa, 2004).

La razón principal de emplear panel de datos es permitir al efecto no observable, a_i , estar correlacionado con las variables explicativas; a_i es constante a través del tiempo y por ello se pueden diferenciar las variables (Wooldridge, 2002).

Estimación del modelo de Efectos fijos:

Si (B.1) satisface todos los supuestos clásicos, el estimador ELIO de β es el estimador MCO incorporando N-1 variables binarias (Sosa, 2004).

Lo anterior se puede representar en matrices de la siguiente manera:

$$\begin{pmatrix} \hat{\beta} \\ \hat{a} \end{pmatrix} = (Z'Z)^{-1} Z'Y$$

Donde Z es una matriz $NT * (K+(N-1))$ con todas las K variables explicativas (X) y las N-1 dummies. El problema es que Z es una matriz muy grande para invertir, por lo que recurrimos a una transformación que elimine a_i (Sosa, 2004):

$$Y_{it} = \beta X_{it} + a_i + e_{it} \quad (\text{B.2})$$

Tomando promedios por individuo:

$$\bar{Y}_i = \beta \bar{X}_i + a_i + \bar{e}_i \quad (\text{B.3})$$

Restando (1)-(2):

$$(Y_{it} - \bar{Y}_i) = \beta(X_{it} - \bar{X}_i) + e_{it} - \bar{e}_i \quad (\text{B.4})$$

Este modelo es llamado *within model*, y permite cierta correlación entre a_i y las variables explicativas en cualquier periodo. Por ende, cualquier variable explicativa que es constante a través del tiempo se elimina con la transformación de efectos fijos (Wooldridge, 2002).

Además se asume que los errores son homocedásticos y no están correlacionados (Wooldridge, 2002).

También se puede recurrir a la estimación de un modelo de efectos fijos a través del método de *primeras diferencias*; el cual se explica a continuación².

$$\text{Dado: } Y_{it} = X_{it}\beta + \delta_2 d2_t + \delta_3 d3_t + a_i + e_{it} \quad (\text{B.5})$$

Donde $d2_t$ adquiere el valor de 1 si se trata del periodo 2, y de 0 cuando se trata del periodo 1 o 3. Y $d3_t$ toma el valor de 1 si $t=3$, y de 0 cuando $t=1, 2$. El periodo base es $t=1$.

Se asume que los errores idiosincrásicos están no correlacionados con la variable explicativa en cada periodo (Wooldridge, 2002):

$$\text{Cov}(X_{itj}, u_{is}) = 0, \text{ para todo } t, s, j.$$

Si la a_i está correlacionada con X_{itj} , entonces X_{itj} está relacionada con el error compuesto $u_{it} = a_i + e_{it}$. Por lo tanto podemos eliminar a_i diferenciando periodos adyacentes.

En el caso de $y=3$ se sustrae el periodo uno del periodo dos, y el dos del periodo tres. Esto resulta en:

² Aquí se desarrolla el caso para $t=3$, pues es el que se presenta en este trabajo.

$$\Delta Y_{it} = \Delta X_{it}\beta + \delta_2 \Delta d2_t + \delta_3 \Delta d3_t + \Delta e_{it} \quad (\text{B.6})$$

Lo cual representa dos periodos de tiempo para cada país en la muestra. Si esta ecuación satisface los supuestos del modelo clásico lineal, entonces MCO arroja estimadores no sesgados (Wooldridge, 2002)³.

EFFECTOS ALEATORIOS

En este modelo se hacen supuestos importantes tales como: homocedasticidad, media cero del error y no correlación entre la variable aleatoria y el error (Sosa, 2004).

Corresponde a:

$$Y_{it} = \beta_{0i} + \beta_1 X_{it} + e_{it}$$

$$\beta_{0i} = \beta_i + v_i$$

$$\rightarrow Y_{it} = \beta_i + \beta_1 X_{it} + v_i + e_{it} \quad (\text{B.7})$$

Además se asume que el intercepto es una variable aleatoria, pues está en función de un valor promedio más un error aleatorio. El término v_i indica la desviación de la constante de la unidad de corte transversal (por ejemplo, el país) y no debe estar correlacionada con los errores. Por lo tanto, e_{it} es diferente para cada observación; y dado el efecto separado de la unidad de corte transversal indicada por v_i , el modelo de efectos aleatorios permite incluir variables que no cambian a través del tiempo como parte de los regresores (Yaffee, 2003).

TEST DE HAUSMAN

Existe una prueba para determinar si el método a usar para estimar un panel de datos es Efectos Aleatorios o Efectos Fijos. Hausman propone evaluar la hipótesis de ausencia de correlación entre a_i y X (Sosa, 2004). Esta prueba tiene distribución asintótica $\chi^2_{(K)}$ bajo:

³ En la presente investigación se compararon regresiones con Efectos Fijos y primeras diferencias. Los coeficientes fueron extremadamente similares, y por ello se decidió emplear el método de Efectos Fijos y reportar sólo esos resultados.

$H_0: E(X_{it} a_i) = 0$

$H_1: E(X_{it} a_i) \neq 0$

Lo que hace esta prueba en general es revisar la matriz de covarianza de los regresores del modelo con efectos fijos y aleatorios. Si no hay diferencia significativa entre las matrices, las correlaciones de los efectos aleatorios con los regresores son estadísticamente insignificantes. Entonces:

Si a_i está correlacionada con X:

- β de Efectos Aleatorios (EA) es inconsistente
- β de Efectos Fijos (EF) es consistente

Si a_i NO está correlacionado con X:

- β de EA es consistente y eficiente.
- β de EF es consistente.

Cuando las variables X están correlacionadas con a_i el estimador EF es consistente y generalmente válido (Sosa, 2004). Por lo tanto, rechazar H_0 sugiere que el estimador de EA es inconsistente. Por ende, el test de Hausman explora la consistencia de los estimadores (Sosa, 2004).

A partir de este punto, las notas se enfocan a la estimación por efectos fijos, pues es lo que se emplea en la presente investigación.

EFFECTOS FIJOS VS. PRIMERAS DIFERENCIAS

Cuando se tienen al menos tres periodos de estudio la estimación por medio de efectos fijos y primeras diferencias no son iguales. La decisión se basa en la eficiencia relativa de los estimadores, la cual se determina por la correlación serial de los errores. Cuando los errores no están correlacionados el método de efectos fijos es más eficiente que las primeras

diferencias y ya que generalmente se asume que los errores están no correlacionados, los efectos fijos son aplicados más comúnmente. Sin embargo, este supuesto puede ser erróneo.

Si los errores presentan correlación serial (en general positiva) entonces no hay correlación entre la diferencia de dichos errores y aplicar primera diferencia es mejor.

De acuerdo con Wooldridge probar si los errores presentan correlación serial después de aplicar efectos fijos es difícil. En cambio, se puede probar la correlación entre las diferencias de los errores, si dicha relación no se presenta se pueden usar primeras diferencias. Pero si existe correlación serial negativa en la diferencia de los errores los efectos fijos son probablemente mejores.

TEST DE EFECTOS FIJOS

Consideremos el modelo de efectos fijos en forma matricial:

$$Y_{it} = X_{it}\beta + Z a_i + e_{it} \quad (\mathbf{B.8})$$

En donde Z es una matriz de $N-1$ variables binarias por intercepto.

Bajo el supuesto de normalidad de e , la hipótesis:

$H_0: a_1 = a_2 = \dots = a_{N-1} = 0$ (ausencia de efectos fijos) puede evaluarse con un test F estándar de significancia conjunta de las $N-1$ variables binarias (Sosa, 2004).