

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**



**FACULTAD DE CIENCIAS NATURALES Y  
MATEMÁTICAS**

**DEPARTAMENTO DE ESTADÍSTICA**

**PROYECTO DE TITULACIÓN**

**PREVIO A LA OBTENCIÓN DEL TÍTULO DE:**

**“MAGÍSTER EN ESTADÍSTICA APLICADA”**

**TEMA:**

**ESTIMACIÓN DE LA MATRIZ DE CARGAS A TRAVÉS DE  
LA DESCOMPOSICIÓN SVD MEDIANTE UN MODELO  
FACTORIAL DE UNIDADES HOMOCEDÁSTICAS**

**AUTOR:**

**GUILLERMO ANDRÉS QUIJIJE ANDRADE**

Guayaquil - Ecuador

2024

# RESUMEN

El análisis factorial es una técnica multivariante ampliamente utilizada en el campo de las ciencias sociales, pero que en ocasiones es considerada como subjetiva o arbitraria en la selección de parámetros o la interpretación de los resultados obtenidos a partir de la misma. En el presente estudio se explora el análisis factorial homocedástico *HFA*, una variante basada en unicidades homocedásticas y descomposición SVD que estima la matriz de cargas utilizando el método de máxima verosimilitud y un conjunto con técnicas de criterios de información, para automatizar la selección y extracción de parámetros, además de simplificar el número de variables que intervienen en cada factor aplicando metodologías similares a la selección de modelos. Finalmente se comparará el desempeño del modelo *HFA* para determinar el número de factores a extraer y la capacidad para recuperar estructuras de ceros con diversas técnicas y criterios de rotación frecuentemente utilizados por la comunidad científica.

**Palabras Clave:** Análisis factorial, Descomposición SVD, criterios de información, rotaciones en la matriz de cargas

# ABSTRACT

Factor analysis is a multivariate technique widely used in the field of social sciences, but it is sometimes considered subjective or arbitrary in the selection of parameters or the interpretation of the results obtained from it. In this study, we explore homoscedastic factor analysis (HFA), a variant based on homoscedastic uniqueness and SVD decomposition. This method estimates the loading matrix using the maximum likelihood method and a set of information criteria techniques to automate the selection and extraction of parameters. Additionally, it simplifies the number of variables involved in each factor by applying methodologies similar to model selection. The primary objective of this research is to enhance the objectivity and reliability of parameter selection in factor analysis. Finally, the performance of the HFA model will be compared to determine the number of factors to extract and its ability to recover zero structures using various techniques and rotation criteria frequently utilized by the scientific community. This comparison aims to validate the effectiveness of the HFA model in providing more accurate and interpretable results in multivariate analysis.

**Keywords:** Factor analysis, SVD decomposition, information criteria, rotations in the loading matrix

## DEDICATORIA

A Dios por brindarme sabiduría y curiosidad particular por las matemáticas.

A mis padres Juanita y Richard por guiarme y alentarme a perseguir mis sueños.

A la mujer que amo, Alisson por ser mi compañera de vida y apoyarme en este viaje.

*“ ... y no olvidéis nunca que hasta el día en que Dios se digne en descifrar el porvenir al hombre, toda la sabiduría humana estará contenida en dos palabras: ¡Confiar y esperar!”*

*-El conde de Montecristo, Alexandre Dumas (1844)*

## **AGRADECIMIENTOS**

A Dios por darme salud, perseverancia y la valentía para afrontar este nuevo reto.

A mis padres Juanita y Richard, a quienes les debo todos mis logros y éxitos

A mi tutor PhD. Francisco Vera por darme la oportunidad de formar parte de su equipo de trabajo y guiarme a través de este proceso.

Al Mat. Juan José Illingworth por su apoyo financiero para llevar a cabo este programa, y su apoyo durante el mismo

## **DECLARACIÓN EXPRESA**

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

---

Ing. Guillermo Andrés Quijije Andrade

# TRIBUNAL DE GRADUACIÓN

---

PhD. Sergio Baúz Olvera  
PRESIDENTE

---

PhD. Omar Ruiz Barzola  
TUTOR

---

PhD. Francisco Vera Alcívar  
DOCENTE EVALUADOR

## ABREVIATURAS O SIGLAS

HFA	Homocedastic Factorial Analysis
SVD	Singular Value Decomposition
AIC	Akaike's Information Criteria
BIC	Bayesian Information Criteria
GIC	Generalized Information Criteria
mBIC	Modified Bayesian Information Criteria
EKC	Empirical Kaiser Criterion
SMT	Sequential Model Test
KGC	Kaiser-Guttman Criterion

## TABLA DE CONTENIDO

RESUMEN .....	I
ABSTRACT .....	II
DEDICATORIA .....	III
AGRADECIMIENTOS .....	IV
DECLARACIÓN EXPRESA .....	V
TRIBUNAL DE GRADUACIÓN .....	VI
ABREVIATURAS O SIGLAS .....	VII
TABLA DE CONTENIDO .....	VIII
LISTADO DE FIGURAS.....	X
LISTADO DE TABLAS.....	XI
CAPÍTULO 1 .....	1
1. INTRODUCCIÓN.....	1
1.1. Antecedentes.....	1
1.2. Descripción del problema.....	3
1.3. Objetivos.....	5
1.4. Hipótesis .....	5
1.5. Alcance .....	5
CAPÍTULO 2 .....	6
2. MARCO TEÓRICO .....	6
2.1. Descomposición SVD.....	6
2.2. Método de máxima verosimilitud.....	6
2.3. Análisis factorial homocedástico.....	7
2.4. Criterios de información.....	9
2.5. Determinación del número de factores.....	10
2.6. Rotaciones en la matriz de cargas factoriales.....	12
CAPÍTULO 3 .....	18
3. Metodología.....	18
3.1. Condiciones del modelo HFA .....	18
3.2. Test de Schott .....	19
3.3. Simulación.....	19

CAPÍTULO 4 .....	21
4.    Resultados.....	21
4.1.    Test Schott .....	21
4.2.    Factores.....	25
4.3.    Ceros Estructurales .....	26
4.3.1.    Método y puntos de corte .....	26
CAPÍTULO 5 .....	28
5.    Conclusiones y Recomendaciones .....	28
5.1.    Conclusiones.....	28
5.2.    Recomendaciones .....	29
CAPÍTULO 6 .....	30
6.    Referencias .....	30
CAPÍTULO 7 .....	32
7.    Apéndices y anexos .....	32
APÉNDICE A .....	32
APÉNDICE B .....	32
APÉNDICE C .....	33
ANEXO 1 .....	34
ANEXO 2 .....	35

## LISTADO DE FIGURAS

Figura 1: Ejemplo de rotación en la matriz de cargas factoriales .....	12
Figura 2: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 1 .....	22
Figura 3: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 2 .....	23
Figura 4: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 3 .....	23
Figura 5: Cantidad de casos excedentes en los que se rechazó la hipótesis nula mediante una prueba de Schott.....	24
Figura 7: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 1 en 10,000 simulaciones .....	27
Figura 8: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 2 en 10,000 simulaciones .....	27
Figura 9: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 3 en 10,000 simulaciones .....	27

## LISTADO DE TABLAS

Tabla 1: Número de casos en que se rechazó la hipótesis nula de la prueba de Schott en las simulaciones .....	21
Tabla 2: Casos en los que se extrajo exactamente 3 factores en 10,000 simulaciones para cada matriz de cargas .....	25
Tabla 3: Casos en 10,000 simulaciones en que se recuperó totalmente la estructura de ceros para cada matriz de cargas .....	26

# CAPÍTULO 1

## 1. INTRODUCCIÓN

El análisis factorial es un modelo matemático ampliamente utilizado en campos como las ciencias sociales y la psicología, para crear estructuras a partir de la correlación de un conjunto de variables medibles, a través de constructos formados por variables no medibles denominadas como factores latentes.

### 1.1. Antecedentes

#### 1.1.1. El modelo uni-factorial

Spearman (1904) interrelacionó las calificaciones 36 estudiantes en 6 puntuaciones obtenidas en pruebas de: *clásicos, francés, inglés, matemática, discriminación de tonos y música*.

Spearman formuló la teoría bifactorial de la inteligencia a través dos términos: el factor (latente) general  $g$  y el factor específico  $e$ .

Matemáticamente representado por:

$$y_i = \lambda_i g + e_i$$

Donde:

$y_i$  representa la puntuación obtenida en la  $i$ -ésima prueba

$\lambda_i$  representa la influencia de  $g$  sobre  $y_i$ , usualmente  $|\lambda_i| \leq 1$

Bajo ciertos supuestos que serán descritos posteriormente, se tiene:

$$var(e_i) = var(y_i) - \lambda_i^2$$

En otras palabras  $e_i$  es la parte de la varianza de  $y_i$  que no logró ser explicada por el factor general  $g$ .

#### 1.1.2. El modelo de múltiples factores

Thurstone (1935) expandió el modelo bifactorial de Spearman ya que afirmó que la inteligencia de una persona debe ser explicada por más de un factor general, y nombro a este nuevo conjunto como *factores comunes*  $f$ .

Nota: se cambiará la notación del factor general  $g$  por la del factor común  $f$ .

De esta forma la puntuación obtenida en la primera prueba estaría dada por:

$$y_1 = \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1r}f_r + e_1$$

La puntuación obtenida en la segunda prueba estaría dada por:

$$y_2 = \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2r}f_r + e_2$$

Finalmente, la puntuación obtenida en la p-ésima prueba estaría dada por:

$$y_p = \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pr}f_r + e_p$$

Generalizando el concepto de Thurstone y empleando el algebra matricial se obtiene:

$$Y = [y_1 \quad y_2 \quad \dots \quad y_p]^t$$

$$\Lambda_{p \times r} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1r} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pr} \end{pmatrix}$$

$$F = [f_1 \quad f_2 \quad \dots \quad f_r]^t$$

$$\varepsilon = [e_1 \quad e_2 \quad \dots \quad e_p]^t$$

Formulando así el modelo matemático más comúnmente utilizado.

$$Y - \mu = \Lambda F + \varepsilon \quad (1)$$

Donde  $\mu = E[Y]$  que representa el vector de medias de  $Y$ ,  $\Lambda$  es conocido como la matriz de *cargas factoriales* y  $\varepsilon$  el *vector de especificidades*.

Bajo los supuestos:

- i.  $E[F] = \mathbf{0}_{r \times 1}$
- ii.  $Cov(F) = I_{r \times r}$
- iii.  $E[\varepsilon] = \mathbf{0}_{p \times 1}$
- iv.  $Cov(\varepsilon) = \Psi_{p \times p} = \text{diag}(\psi_p^2)$
- v.  $\varepsilon$  y  $F$  son independientes, por lo tanto  $Cov(\varepsilon, F) = \mathbf{0}_{p \times r}$

Puede calcularse  $\Sigma = Cov(Y)$

$$\Sigma = \Lambda \Lambda^t + \Psi \quad (2)$$

Y se desprende:  $Cov(Y, F) = \Lambda$  o equivalentemente  $cov(y_i, f_j) = \lambda_{ij}$  que mide el grado de asociación entre la variable medible  $y_i$  y el factor latente  $f_j$ ; para  $\begin{cases} i = 1, \dots, p \\ j = 1, \dots, r \end{cases}$

El enfoque clásico del análisis factorial recae sobre la descomposición (2) para lo cual se han propuesto varias metodologías que van desde componentes principales, métodos de máxima verosimilitud y estimación por mínimos cuadrados.

De manera general se denomina comunalidad  $h_i^2$  a la proporción de la varianza explicada por un determinado factor, y unicidad  $\psi_i^2$  (o varianza específica) a la proporción de la varianza asociada a un determinado factor específico.

Matemáticamente se expresa por:

$$\sigma_i^2 = h_i^2 + \psi_i^2 \quad (3)$$

Donde:  $h_i^2 = \sum_{k=1}^r \lambda_{ik}^2$

## 1.2. Descripción del problema

Existen problemas asociados a la factorización de la matriz  $\Sigma$ , su posterior representación a partir de  $r$  factores y al hecho de que la estimación de la matriz de cargas factoriales  $\Lambda$  no es única.

### 1.2.1. El caso de Heywood

Puede suceder que la reconstrucción de  $\Sigma$  a partir de las diferentes técnicas para estimar a las matrices de cargas  $\Lambda$  o de covarianzas de las especificidades  $\Psi$  den paso a interpretaciones incoherentes desde el punto de vista de la estadística (e.g. varianzas negativas) (Heywood, 1931).

Por ejemplo:

Se tiene a la matriz de covarianzas  $\Sigma = \begin{pmatrix} 1 & 0.4 & 0.9 \\ 0.4 & 1 & 0.7 \\ 0.9 & 0.7 & 1 \end{pmatrix}$  y se desea factorizarla utilizando solo un factor ( $r = 1$ ), empleando (2):

$$\begin{pmatrix} 1 & 0.4 & 0.9 \\ 0.4 & 1 & 0.7 \\ 0.9 & 0.7 & 1 \end{pmatrix} = (\lambda_{11} \quad \lambda_{21} \quad \lambda_{31}) \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{pmatrix} + \begin{pmatrix} \psi_1^2 & 0 & 0 \\ 0 & \psi_2^2 & 0 \\ 0 & 0 & \psi_3^2 \end{pmatrix}$$

Luego de resolver el sistema de ecuaciones, las cargas son:

$$\lambda_{11}^2 = \frac{18}{35}, \lambda_{21}^2 = \frac{14}{45}, \lambda_{31}^2 = \frac{63}{40}$$

Empleando (3) se obtiene:

$$\psi_1^2 = 1 - \frac{18}{35}, \psi_2^2 = 1 - \frac{14}{45}, \psi_3^2 = 1 - \frac{63}{40}$$

Nótese que es posible factorizar (de forma única)  $\Sigma$  a través de  $\Lambda$  y  $\Psi$ , no obstante dado que  $\psi_3^2 \approx -0.58$ , esta factorización es estadísticamente incoherente<sup>1</sup>.

---

<sup>1</sup> La solución para  $\psi_3^2$  no es válida ya que este término representa a la varianza de la 3era especificidad.

### 1.2.2. Rotaciones de la matriz de cargas factoriales

Si se añade una matriz ortogonal<sup>2</sup>  $R$  al modelo (1) se tiene:

$$Y - \mu = \Lambda^* F^* + \varepsilon$$

Donde  $\Lambda^* = \Lambda R$  y  $F^* = R^t F$ , la matriz de covarianzas  $\Sigma$  permanece invariante a la rotación ocasionada por  $R$ .

$$\Sigma = \Lambda^* (\Lambda^*)^t + \Psi$$

$$\Sigma = \Lambda R R^t \Lambda + \Psi = \Lambda \Lambda + \Psi$$

Las matrices de cargas rotadas producen la misma matriz  $\Sigma$ , por lo que la solución a partir de la estimación de  $\Lambda$  y  $\Psi$  no siempre será única.

Usualmente las rotaciones de la matriz de cargas son empleadas con la finalidad de optimizar criterios asociados a la variabilidad, y que modifiquen el grado de relación entre las variables medibles y el factor latente.

Y por otro lado, se busca que la rotación no agregue cargas a variables apenas relacionadas al factor, a esta última parte se le denominará ceros estructurales de la matriz de cargas.

En el siguiente capítulo se describirá un conjunto de técnicas y sus criterios de optimización para rotar la matriz de cargas.

### 1.2.3. Sobre parametrización

El análisis factorial se basa en representar a la matriz  $\Sigma$  mediante la factorización a través de  $\Lambda$  y  $\Psi$ .

Se denota por  $\#(x)$  a la función que representa el número de parámetros de  $x$ , por lo tanto:

El número de parámetros de  $\Sigma$  es:

$$\#(\Sigma) = \frac{p(p+1)}{2} \quad (4)$$

Mientras que el número de parámetros empleados por la factorización es:

$$\#(\Lambda \Lambda^t + \Psi) = pr + p \quad (5)$$

Nótese que el número de parámetros de (5) depende de  $r$  (el número de factores latentes escogidos por el usuario).

La representación factorial resultante estará sobre parametrizada si:  $r > \frac{p-1}{2}$

---

<sup>2</sup> Si  $R$  es ortogonal  $RR^t = R^t R = I_{r \times r}$

## 1.3. Objetivos

### 1.3.1. Objetivo General

Comparar el desempeño del modelo de análisis factorial con unicidades homocedásticas versus el análisis factorial clásico

### 1.3.2. Objetivos Específicos

- i. Estimar la matriz de cargas mediante el método de máxima verosimilitud
- ii. Descomponer la matriz de covarianzas en sus valores y vectores singulares asumiendo que existen unicidades homocedásticas.
- iii. Comparar el desempeño de los criterios de información (*AIC*, *BIC*, *mBIC*, *GIC*) versus las técnicas clásicas para determinar el número de factores mediante simulaciones.
- iv. Comparar el desempeño método del iterativo basado en criterios de información (*AIC*, *BIC*, *mBIC*, *GIC*) versus las técnicas clásicas de rotación de la matriz de cargas para recuperar los ceros estructurales, mediante simulaciones

## 1.4. Hipótesis

Vera (2023) propone una variante al análisis factorial clásico denominado *Análisis factorial homocedástico* (HFA) basado en una matriz de covarianzas **spike**, que:

1. Evita la sobre parametrización
2. No permite rotaciones de la matriz  $\Lambda$ , ofreciendo una factorización (2) única
3. Propone una alternativa de estimación por máxima verosimilitud que considera restricciones sobre ceros estructurales en posibles cargas, obteniendo una estimación consistente del verdadero número de factores latentes y que evita casos de Heywood.

Y demuestra que la condición para aplicar el modelo HFA es que un conjunto pequeño de valores propios  $p - r$  de  $\Sigma$  sean iguales a un único termino para la unicidad  $\psi^2$ :

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_p = \psi^2 \quad (6)$$

Esta condición puede ser comprobada mediante una prueba de Schott.

## 1.5. Alcance

El presenta trabajo describirá y comparará el desempeño de las técnicas propuestas por Vera, versus con un conjunto seleccionado de métodos para determinar el número de parámetros y la cantidad de ceros estructurales recuperados en datos simulados. El estudio tuvo una duración de 8 meses, a continuación se presentarán los métodos y el desarrollo de la investigación.

# CAPÍTULO 2

## 2. MARCO TEÓRICO

### 2.1. Descomposición SVD

La descomposición en valores singulares SVD (Eckart & Young, 1936) es una técnica muy utilizada para reducir la dimensión de una matriz rectangular cualquiera.

Sea  $X \in M_{m \times n}$ , es posible aproximarla mediante:

$$X = U\Lambda V^t \quad (7)$$

Donde:

$\Lambda$  es una matriz diagonal de dimensión  $k = \min\{m, n\}$ , que contiene los valores singulares de  $A^t A$  si  $k = m$  o  $AA^t$  si  $k = n$ :

$$D = \text{diag}(\delta_1, \dots, \delta_k)$$

$U_{m \times k}$  es la matriz ortogonal que contiene los  $k$  –vectores propios de  $A^t A$

$V_{n \times k}$  es la matriz ortogonal que contiene los  $k$  –vectores propios de  $AA^t$

### 2.2. Método de máxima verosimilitud

El método de máxima verosimilitud tomando como referencia (Johnson & Wichern, 2007) es una alternativa de estimación de parámetros  $\theta$  empleados por modelos matemáticos, cuya única restricción es la imposición de una cierta distribución sobre el conjunto de datos observados.

Sean  $\{X_1, X_2, \dots, X_m\} \in R^p$  vectores aleatorios con función de probabilidad  $f_p(X_i|\theta)$ , se denota  $\mathcal{L}$  a la función de verosimilitud:

$$\mathcal{L} = \prod_{i=1}^n f_p(X_i|\theta) \quad (8)$$

Asumiendo que cada vector aleatorio sigue una distribución Normal Multivariante con vector de medias  $\mu$  y matriz de covarianzas  $\Sigma$  dada por:

$$f_p = (2\pi)^{-p/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(X-\mu)^t \Sigma^{-1} (X-\mu)} \quad (9)$$

Reemplazando la distribución multivariante de los vectores aleatorios en (8) se obtiene:

$$\mathcal{L} = (2\pi)^{-p/2} |\Sigma|^{-m/2} e^{-\frac{1}{2} \sum_{j=1}^m \text{tr}\{\Sigma^{-1} (X_j - \mu)(X_j - \mu)^t\}}$$

No obstante, para métodos de optimización se trabaja<sup>3</sup> con  $l = \log(\mathcal{L})$

Tomando el logaritmo a la función de verosimilitud se tiene:

$$l = \alpha - \frac{m}{2} \log|\Sigma^{-1}| - \frac{1}{2} \text{tr} \left[ \sum_{j=1}^m \Sigma^{-1} (X_j - \mu)(X_j - \mu)^t \right] \quad (10)$$

Donde  $\alpha = -\frac{p}{2} \log(2\pi)$  es una constante que no depende de los parámetros  $\mu$  y  $\Sigma$ .

### 2.3. Análisis factorial homocedástico

Vera (2023) propone factorizar  $\Lambda$  mediante la descomposición SVD de la forma:

$$\Lambda = LDR^t \quad (11)$$

Si se asume que el término de la unicidad es único (e igual a  $\psi^2$ ) para todas las  $p$  –especificidades,  $\Psi$  quedaría expresado en términos de la matriz  $I_p$  en (2):

$$\Sigma = LD^2L^t + \psi^2 I_p \quad (12)$$

Nótese que  $\Sigma$  ahora no se ve afectada por la matriz de rotación  $R$  por lo tanto ya no podrá ser estimada, lo que asegura que  $\Sigma$  permanezca invariante a las rotaciones y de esta manera se asegura que la matriz de cargas  $\Lambda$  sea **única**.

El modelo (1) quedaría redefinido de la siguiente manera:

$$Y - \mu = LDF + \varepsilon \quad (13)$$

Donde la matriz de cargas ahora es denominada *matriz de cargas escaladas* y se calculará por  $\Lambda = LD$ , asumiendo que el producto  $\Lambda^t \Lambda = D^2$  es una matriz diagonal.

Si cada vector columna  $y_1, \dots, y_p$  de  $Y$  sigue una distribución normal multivariante con vector de medias  $\mu$  y matriz de covarianzas (12), al reemplazar en la función de verosimilitud (10) se tiene:

$$l(\psi^2, \delta_j^2, L_j) = -\frac{m}{2} \left[ p \log(2\pi) + (p - r) \log(\psi^2) + \sum_{j=1}^r \log(\delta_j^2 + \psi^2) \right. \\ \left. + \sum_{j=1}^r \left( \frac{1}{\delta_j^2 + \psi^2} - \frac{1}{\psi^2} \right) L_j^t S L_j + \frac{1}{\psi^2} \text{tr}(S) \right]$$

---

<sup>3</sup> El término *log* hace referencia al logaritmo en base  $e$

Donde  $L_j$  representa al  $j$ -ésimo vector columna de  $L$  y  $S$  representa a la matriz de covarianzas muestral.

Dado que  $l$  depende de  $\psi^2$ ,  $\delta_j^2$  y  $L_j$ , las respectivas estimaciones obtenidas al maximizar la función  $l$  son:

$$\begin{aligned}\hat{\psi}^2 &= \frac{1}{p-r} \sum_{j=r+1}^p \hat{\lambda}_j \\ \hat{\delta}_j^2 &= \hat{\lambda}_j - \hat{\psi}^2, \quad j = 1, \dots, r \\ \hat{L}_j &= v_j, \quad j = 1, \dots, r\end{aligned}$$

Donde  $v_j$  representa al  $j$ -ésimo vector propio correspondiente al valor propio  $\hat{\lambda}_j$  respectivo de la matriz  $S$ .

Si se reemplazan las estimaciones  $\psi^2$ ,  $\hat{\delta}_j^2$ ,  $\hat{L}_j$  en la función de verosimilitud se alcanza su máximo valor:

$$\begin{aligned}\hat{l}_{max} &= -\frac{m}{2} (p \log(2\pi) + p + (p-r) \log(\hat{\psi}^2) \\ &\quad + \sum_{j=1}^r \log(\hat{\delta}_j^2 + \hat{\psi}^2))\end{aligned}\tag{14}$$

Adicional, si existen ceros estructurales en la matriz de cargas se debería modificar la función de máxima verosimilitud con restricciones adicionales que los consideren.

Vera (2023) demostró que las estimaciones de los parámetros de  $\psi^2$ ,  $\delta_j^2$  de  $\Sigma$  que consideren ceros estructurales son:

$$\hat{\psi}^2 = \frac{1}{p-r} \left( tr(S) - \sum_{j=1}^r \theta_j \right), \quad j = 1, \dots, r\tag{15}$$

$$\hat{\delta}_j^2 = \theta_j - \hat{\psi}^2, \quad j = 1, \dots, r\tag{16}$$

Donde  $\theta_j$  representa al valor propio más alto de la matriz de covarianzas  $S_{(-k_j)}$  que resulta de remover las filas y columnas correspondientes a los ceros estructurales.

## 2.4. Criterios de información

Bajo el modelo de unicidades homocedásticas, el número de parámetros de (12) es:

$$\#(\Sigma) = pr - (r + 1) \left( \frac{r}{2} - 1 \right)$$

Si se considera al vector de medias  $\mu$ , se tienen  $p$  parámetros adicionales. En total el número de parámetros  $k$  a estimar por el modelo HFA es:

$$k = (r + 1) \left( p - \frac{r}{2} + 1 \right) \quad (17)$$

Aun cuando solo se tome en cuenta el número de parámetros de la matriz de covarianzas  $\Sigma$ , el modelo HFA no estará sobre parametrizado si:

$$r < p - 1$$

Nótese que  $k$  depende de  $r$  (cantidad de valores propios distintos de  $\Sigma$ ), por lo que será estimada empleando criterios de información.

- AIC (Akaike, 1974):

$$AIC = -2\hat{l}_{max} + 2k$$

- BIC (Schwarz, 1978):

$$BIC = -2\hat{l}_{max} + k \log(n)$$

- mBIC (Wang, Li, & Leng, 2009):

$$mBIC = -2\hat{l}_{max} + k \left( \log \log \frac{p(p+3)}{2} \right) \log(n)$$

- GIC (Kim, Kwon, & Choi, 2012):

$$GIC = -2\hat{l}_{max} + kn^c, \quad 0 < c < 1$$

Nuevamente, si existen ceros estructurales el número de parámetros del modelo HFA sería:

$$k = (r + 1) \left( p + 1 - \frac{r}{2} \right) - \sum_{j=1}^r c_j \quad (18)$$

Donde  $c_j$  representa la cantidad de ceros estructurales en  $L_j$  ( $j$  -ésimo) vector columna de cargas.

## 2.5. Determinación del número de factores

A continuación se presentarán algunas de las técnicas más utilizadas para determinar el número de factores a extraer y las diferentes rotaciones de la matriz de cargas factoriales.

- **Criterio Kaiser-Guttman** (Guttman, 1954) & (Kaiser H. , 1960)  
Es la técnica más parsimoniosa, consiste en extraer todos los factores cuyos valores propios sean mayores que 1, ya que se postula que un factor debe explicar al menos su propia varianza.  
No obstante, esta técnica tiende a sobre estimar el número de parámetros, debido a que por errores relacionados al muestreo varios valores propios podrían ser mayores que 1.  
Se hará referencia a este criterio por sus siglas en inglés **KGC**
- **Análisis Paralelo** (Horn, 1965)  
Técnica basada en simular de manera aleatoria  $N$  conjuntos de prueba a partir de los datos, y que además no guarden relación con los mismos.  
Posteriormente se comparan los valores propios del conjunto de datos versus el promedio de los valores propios obtenidos en las  $N$ -simulaciones con la finalidad de extraer tantos factores (de la matriz de estudio) como el número de valores propios que superó al promedio (en las  $N$ -simulaciones).  
Por defecto se generarán  $N = 1,000$  simulaciones.  
Se hará referencia a este criterio por **PARALLEL**
- **Método Hull** (Lorenzo-Seva, Timmerman, & Kiers, 2011)  
Esta técnica consiste en generar  $(J + 1)$  –modelos en donde cada modelo se construye a partir de  $j$  factores, de tal manera que el primer modelo tiene  $j = 0$  factores, el segundo modelo  $j = 1$  y así sucesivamente hasta llegar al  $(j + 1)$ -ésimo modelo con  $j = J = p$  factores, en cada paso se calcula un determinado índice de bondad de ajuste  $GOF_j$ .  
Iterativamente, se comparan todos los modelos generados hasta que uno **no pueda ser clasificado** como *solución inviable*.  
La solución  $s_k$  creada a partir un modelo de  $k$  factores, será inviable si existe una solución  $s_{k'}$  creada a partir de un modelo con  $k'$  factores (en donde  $k' < k$ ) tal que:  
$$GOF_k > GOF_{k'}$$
  
De esta forma el método de HULL sugerirá aquella solución con mayor índice de bondad de ajuste alcanzado con el menor número de factores posibles.  
Se hará referencia a este criterio por **HULL**

- **Criterio empírico de Kaiser** (Braeken & van Assen, 2017)  
Si los datos siguen una distribución normal, los valores propios están asintóticamente distribuidos bajo un modelo Marchenko-Pastur. Se crea un nuevo conjunto de valores propios denominados *referencias* del mismo tamaño que los valores propios de  $\Sigma$ .

El primer valor propio de referencia será:

$$\lambda_{1,ref} = \left(1 + \sqrt{\frac{p}{n}}\right)^2$$

Donde  $p$  representa el número de columnas y  $n$  el número de filas de la matriz de datos ( $X_{n \times p}$ )

Mientras que el  $j$ -ésimo valor propio de referencia será:

$$\lambda_{j,ref} = \max \left\{ \frac{p - \sum_{i=0}^{j-1} \lambda_j}{p - j + 1} \left[1 + \sqrt{\frac{p}{n}}\right]^2, 1 \right\}$$

Nótese que a partir del 2do valor propio de referencia, se añade un ponderador que toma en cuenta el porcentaje de varianza explicado hasta ese momento, de esta manera se disminuye el riesgo sobre estimación de parámetros.

Se extraerán tantos factores como el número de valores propios  $\lambda$  de  $\Sigma$  que superaron a sus respectivos valores propios de referencia.

Se hará referencia a este criterio por sus siglas en inglés **EKC**

- **Modelos  $\chi^2$  secuenciales** (Auerswald & Moshagen, 2019)  
En este método se probarán iterativamente  $J$ -modelos con  $0 \leq j \leq J$  factores, en cada paso se realizará una prueba de la razón de verosimilitud.

$$H_0: \Sigma = \Sigma_0 = \Lambda \Lambda^t + \Psi$$

$$H_a: \Sigma \neq \Sigma_0$$

El estadístico de prueba asociado a el contraste de hipótesis es:

$$X^2 = \left(n - 1 - \frac{2p + 4r + 5}{6}\right) \log \left( \frac{|\widehat{\Lambda} \widehat{\Lambda}^t + \widehat{\Psi}|}{|\widehat{S}|} \right)$$

Que sigue una distribución  $\chi^2$  con  $\frac{1}{2}[(p - r)^2 - p - r]$  grados de libertad, donde:

- $\widehat{S}$  es el estimador de máxima verosimilitud la matriz de covarianzas muestrales
- $n, p$  son las filas y columnas de la matriz de datos respectivamente
- $r$  es el número de factores a retener

El algoritmo termina cuando se encuentre una matriz  $\Sigma_0$  que difiera significativamente de  $\Sigma$ , finalmente se recomendará el número de factores de  $\Sigma_0$

Se hará referencia a este criterio por sus siglas en inglés **SMT**

## 2.6. Rotaciones en la matriz de cargas factoriales

### 2.6.1. Ejemplo de Aplicación

Con frecuencia se busca rotar la matriz de cargas factoriales para alcanzar estructuras más sencillas que faciliten la interpretación desde el punto de vista del investigador.

Esta estructura sencilla, se alcanza idealmente cuando ciertas cargas obtienen valores relativamente altos en un factor y al mismo tiempo, valores relativamente bajos en los demás factores, por ejemplo:

Se tiene una matriz de cargas con  $p = 6$  variables y  $r = 2$  factores

$$\Lambda^t = \begin{bmatrix} 0.553 & 0.568 & 0.392 & 0.740 & 0.724 & 0.595 \\ 0.429 & 0.288 & 0.450 & -0.273 & -0.211 & -0.132 \end{bmatrix}$$

Dado que la matriz de cargas emplea únicamente dos factores, la matriz de rotación  $R$  estaría dada por la expresión:

$$R = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

Considerando  $\phi = 20^\circ$  se tiene:

$$\Lambda^* = \Lambda R$$

$$(\Lambda^*)^t = \begin{bmatrix} 0.373 & 0.435 & 0.214 & 0.789 & 0.753 & 0.604 \\ 0.592 & 0.465 & 0.557 & -0.003 & 0.049 & 0.079 \end{bmatrix}$$

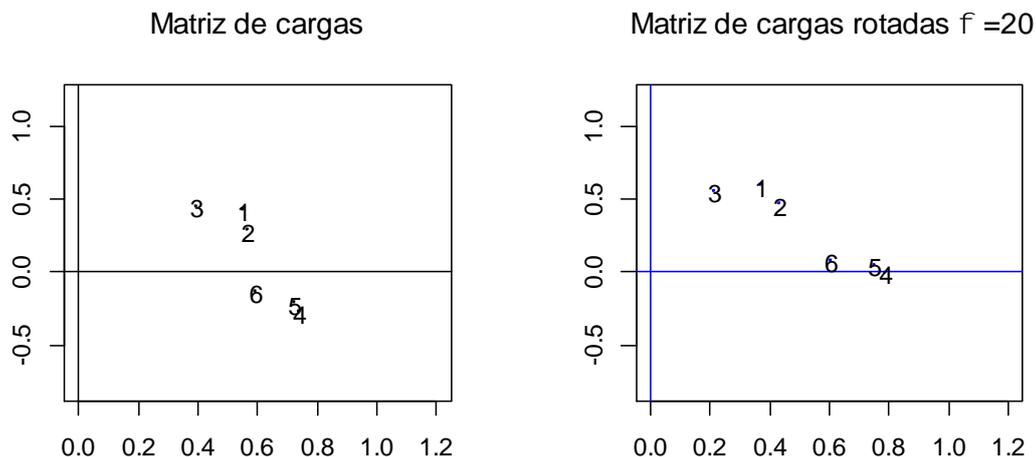


Figura 1: Ejemplo de rotación en la matriz de cargas factoriales

Fuente: Elaboración propia

En la figura 1 a la izquierda se observan graficadas las cargas de la matriz  $\Lambda$ , nótese que las variables **1 y 2** tienen valores moderadamente altos en el **1er factor** (eje x).

Por otro lado, las variables **4, 5 y 6** tienen valores relativamente altos en el **1er factor**, una posible interpretación sería considerar que las variables “todas” las variables están relacionadas con el **1er factor**.

No obstante al realizar una rotación considerando  $\phi = 20^\circ$ , se obtiene una agrupación diferente entre las variables y los factores.

A la derecha en la figura 1, se puede observar que las variables **1, 2 y 3** ahora solo tienen valores relativamente altos en el **2do factor**, mientras que el resto de las variables **4, 5 y 6** solo tienen valores relativamente altos en el **1er factor**, de esta manera se ha logrado aclarar los constructos formados por el análisis factorial.

Las rotaciones de la matriz de cargas serán estimadas mediante el paquete *GPARotation (GPA)* de Bernaards (2005).

### 2.6.2. Rotación Ortogonal

Se dice que  $\tilde{\Lambda}$  es una rotación ortogonal de la matriz  $\Lambda$  si:  $\tilde{\Lambda} = \Lambda_{r \times p} R_{p \times p}$  (para alguna matriz ortogonal  $R$ ), entonces  $\tilde{\Lambda}$  será la matriz obtenida luego de multiplicar la matriz de cargas  $\Lambda$  por la matriz de **rotación**  $R$ .

A continuación se detallarán algunas herramientas matemáticas para aplicar el algoritmo *GPA* en R:

#### Producto de Frobenius

Sean dos matrices  $m \times n$ ,  $A$  y  $B$  se notará al producto de Frobenius de la siguiente manera:

$$\{A, B\} = tr(A^t B)$$

De esta definición se desprende al concepto de:

#### Norma de Frobenius

$$\|A\| = \{A, A\}^{1/2}$$

La norma de Frobenius es la raíz cuadrada del producto de Frobenius de una matriz consigo mismo.

#### Producto de Hadamard

$$(A \cdot B)_{ij} = (A)_{ij}(B)_{ij}$$

$$(A / B)_{ij} = (A)_{ij}/(B)_{ij}$$

Nótese que el resultado del producto de Hadamard, es una matriz  $m \times n$  cuyos elementos son el producto (o la división) de elementos respectivos de  $A$  y  $B$ .

Se denotará  $X^2$  al producto de Hadamard de  $X$  consigo mismo, es decir:

$$X^2 = (X \cdot X)_{ij} = (X)_{ij}(X)_{ij}$$

### Gradiente de una rotación

Sea  $Q$  algún criterio de rotación particular, el objetivo del algoritmo *GPA* es optimizar  $Q(\tilde{\Lambda})$  sobre todas las posibles rotaciones ortogonales  $\tilde{\Lambda}$  de  $\Lambda$ .

Se notará por  $G_q$  a la matriz de derivadas parciales de  $Q(\tilde{\Lambda})$  respecto a los componentes de  $\tilde{\Lambda}$ .

$$G_q = \frac{dQ}{d\tilde{\Lambda}}$$

De esta manera  $G_q$  sería el gradiente de  $Q$  respecto a  $\tilde{\Lambda}$ . Además se puede representar el diferencial de  $Q$  evaluado en  $\tilde{\Lambda}$  mediante el producto de Frobenius:

$$dQ_{\tilde{\Lambda}}(d\tilde{\Lambda}) = \{G_q, d\tilde{\Lambda}\}$$

### 2.6.3. Algoritmo GPA

**Valores de Entrada:**  $\Lambda_{p \times r}$ ,  $R = I_{r \times r}$ ,  $\alpha = 1$

1. Se computa la rotación  $\tilde{\Lambda} = \Lambda R$
2. Se evalúa  $Q(\tilde{\Lambda})$  y se calcula el gradiente  $G_q$   
 Algoritmo iterativo 1: Se ejecutará hasta que  $s < 1 \times 10^{-5}$
3. Se computan las matrices:  $G = \Lambda^t G_q$ ;  $M = R^t G$ ;  $S = \frac{1}{2}[M + M^t]$
4. Se calcula la matriz:  $G_p = G - RS$
5. Se calcula el criterio de parada:  $s = \|G_p\|$
6. Se almacena el valor del criterio de rotación evaluado en la matriz  $\tilde{\Lambda}$ :  
 $f = Q(\tilde{\Lambda})$

Algoritmo iterativo 2: Se ejecutará hasta que  $f^* < f - \frac{1}{2}s^2\alpha$

- 6.1. Se establece la distancia  $\alpha^* = 2\alpha$
- 6.2. Se calcula la matriz  $X = R - \alpha^* G_p$
- 6.3. Se realiza la descomposición SVD de  $X$ :  $X' = SVD(X)$
- 6.4. Se calcula la matriz de rotación  $R' = UV$  a partir de los vectores propios izquierdos y derechos de  $X$
- 6.5. Se calcula la nueva matriz rotada  $\Lambda^* = \Lambda R'$
- 6.6. Se calcula su respectivo criterio de rotación evaluado en la nueva matriz rotada  $f^* = Q(\Lambda^*)$  y su gradiente  $G_q^*$
- 6.7. Se modifica la distancia  $\alpha^* = \alpha^*/2$
7. Se retornan los siguientes resultados:

$$R_{res} = R'; f_{res} = f^*; G_{q_{res}} = \Lambda^t G_q^*$$

Donde:

- $R_{res}$  es la matriz ortogonal que genera la rotación  $\tilde{\Lambda} = \Lambda R_{res}$
- $f_{res}$  es el valor optimizado del criterio de rotación  $Q(\tilde{\Lambda})$
- $G_{q_{res}}$  es el gradiente optimizado.

El algoritmo *GPA* requiere principalmente de:

- i. Un criterio de rotación a optimizar  $Q(\tilde{\Lambda})$
- ii. El gradiente del criterio de rotación (evaluado en  $\tilde{\Lambda}$ )  $G_q$  que será utilizado como criterio de parada.

En la siguiente sección se describirán brevemente los criterios de rotación y sus gradientes.

#### 2.6.4. Criterios de Rotación

**Quartimax** (Carroll, 1953) & (Neuhaus & Wrigley, 1954)

La función objetivo de la rotación *quartimax* es:

$$Q(\tilde{\Lambda}) = -\frac{1}{4} \sum_{i=1}^p \sum_{j=1}^r \tilde{\lambda}_{ij}^4$$

Donde  $\tilde{\lambda}_{ij}$  representa a los elementos de la matriz  $\tilde{\Lambda}_{p \times r}$ .

En este criterio se busca optimizar la suma de los elementos de la matriz de cargas rotada  $\tilde{\lambda}_{ij}$  elevados a la cuarta potencia.

Empleando las notaciones previamente descritas puede reescribirse a la función objetivo utilizando el producto de Frobenius:

$$Q(\tilde{\Lambda}) = -\frac{1}{4} \{\tilde{\Lambda}^2, \tilde{\Lambda}^2\}$$

De esta forma, el gradiente puede expresarse mediante:

$$G_q = -\tilde{\Lambda}^3$$

La idea detrás de esta rotación es disminuir el número de factores que explican una variable, saturando cargas aquellas cargas altas y minimizando cargas bajas.

**Varimax** (Kaiser H. , 1958)

La función objetivo de la rotación *Varimax* está dada por la expresión:

$$Q(\tilde{\Lambda}) = -\frac{1}{p} \sum_{j=1}^r \left[ \sum_{i=1}^p \lambda_{ij}^{*4} - \frac{1}{p} \left( \sum_{i=1}^p \lambda_{ij}^{*2} \right)^2 \right]$$

Donde  $\lambda_{ij}^* = \tilde{\lambda}_{ij}/h_i$  se denomina carga estandarizada, y se obtiene al dividir cada carga y su respectiva comunalidad.

Nuevamente empleando la notación previamente descrita se obtiene:

$$Q(\tilde{\Lambda}) = \frac{1}{4} \{ \tilde{\Lambda}^2, [I - C] \tilde{\Lambda}^2 N \}$$

Donde:

$I$  es la matriz identidad de dimensión  $p \times p$

$C$  es una matriz cuadrada cuyos elementos son  $1/p$

$$N_{r \times r} \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases}$$

Finalmente el gradiente sería:

$$G_q = \tilde{\Lambda} \cdot [(I - C) \tilde{\Lambda}^2 N]$$

La idea detrás de este criterio es polarizar las cargas en cada eje factorial, de manera que cada factor tenga un número reducido de variables con cargas altas.

Usualmente esta técnica facilita la interpretación de los factores ya que dependen de un número reducido de variables.

**Familia Crawford-Ferguson** (Crawford & Ferguson, 1970)

Conjunto de técnicas cuyo criterio de rotación es:

$$Q(\tilde{\Lambda}) = \frac{1}{4} [(1 - \kappa) \{ \tilde{\Lambda}^2, \tilde{\Lambda}^2 N \} + \kappa \{ \tilde{\Lambda}^2, M \tilde{\Lambda}^2 \}]$$

Donde  $N$  y  $M$  son matrices cuadradas de dimensión  $r$  y  $p$  respectivamente con ceros en su diagonal principal y unos fuera de ella.

$\kappa$  es un número en el intervalo  $[0,1]$ , destacando:

- $\kappa = 0$  se tiene una rotación Quartimax
- $\kappa = 1/p$  se tiene una rotación Varimax

- $\kappa = k/2p$ , se tiene una rotación Equamax donde  $k$  es el número de columnas y  $p$  es el número de filas de  $\tilde{\Lambda}$
- $\kappa = \frac{k-1}{p+k-2}$ , se tiene una rotación Parsimax

Finalmente el gradiente sería:

$$G_q = (1 - \kappa)\tilde{\Lambda} \cdot \tilde{\Lambda}^2 N + \kappa\tilde{\Lambda} \cdot M\tilde{\Lambda}^2$$

La idea detrás de este criterio es combinar los criterios Varimax y Quartimax de forma parsimoniosa.

# CAPÍTULO 3

## 3. Metodología

### 3.1. Condiciones del modelo HFA

En el capítulo 1 se expuso brevemente la condición del modelo HFA, a continuación se presentará la definición formalmente.

**Teorema.** -  $\Sigma$  tiene representación mediante el modelo HFA, con  $r$ - factores si y solo si sus últimos  $p - r$  valores propios son iguales.

Vera (2023) propone la demostración de la siguiente forma:

Si se supone la estructura (6) en los valores propios de la matriz  $\Sigma_{p \times p}$ , puede dividirse a sus respectivos vectores propios en dos grupos:

$$V_1 = [v_1 \quad v_2 \quad \cdots \quad v_r], \text{ donde } \lambda_1 \geq \cdots \geq \lambda_r$$

$$V_2 = [v_{r+1} \quad v_{r+2} \quad \cdots \quad v_p], \text{ donde } \lambda_{r+1} = \cdots = \lambda_p = \psi^2$$

Dado que  $\Sigma$  es una matriz simétrica, los vectores propios pueden calcularse de forma que sean ortogonales, lo que implica:

- 1)  $V_1^t V_1 = I_r$
- 2)  $V_2^t V_2 = I_{p-r}$
- 3)  $V_1^t V_2 = 0_r 0_{p-r}^t$
- 4)  $V_2^t V_1 = 0_{p-r} 0_r^t$

Además:

$$V_1 V_1^t + V_2 V_2^t = I_p \quad (19)$$

La descomposición en valores y vectores propios sería:

$$\Sigma = V_1 \Lambda_1 V_1^t + V_2 (\psi^2 I_p) V_2^t \quad (20)$$

Despejando  $V_2 V_2^t$  en (19), reemplazándose en (20) y factorizando se tiene:

$$\Sigma = V_1 (\Lambda_1 - \psi^2 I_r) V_1^t + \psi^2 I_p$$

Nótese que  $\Sigma$  se representó mediante (12) con:  $L = V_1$  y  $D^2 = \Lambda_1 - \psi^2 I_r$ , que es una matriz diagonal con elementos<sup>4</sup>  $\delta_j^2 = \lambda_j - \psi^2$ ;  $j = 1, \dots, r$ .

---

<sup>4</sup> Este resultado coincide con la estimación por máxima verosimilitud

Finalmente para probar la implicación en el otro sentido, se asumirá que  $\Sigma$  efectivamente tiene la forma (12) con  $D^2 = \text{diag}(\delta_1^2, \dots, \delta_r^2)$  y  $L^t L = I_r$ :

$$\Sigma L = L(D^2 + \psi^2 I_r)$$

Entonces,  $L$  sería la matriz que contiene los vectores propios de  $\Sigma$  correspondientes a los valores propios:  $\delta_j^2 + \psi^2$  de manera que los restantes valores propios de  $\Sigma$  (que son iguales a  $\psi^2$ ) y sus respectivos vectores propios serán parte del complemento ortogonal de  $L$ .

### 3.2. Test de Schott

Para probar el supuesto de unicidades homocedásticas

$$H_0: \lambda_{r+1} = \dots = \lambda_p = \psi^2$$

Vera sugiere utilizar una modificación del estadístico de Schott (2006) empleando los estimadores de máxima verosimilitud:

$$U = \frac{n-r}{2(\hat{\psi}^2)^2} \sum_{j=r+1}^p (\hat{\lambda}_j - \hat{\psi}^2)^2 \quad (21)$$

$U$  converge en distribución a  $\chi^2[(p-r)(p-r+1)/2 - 1]$

Se realizará el contraste de hipótesis al 10%, 5% y 1% de significancia.

### 3.3. Simulación

Bajo el modelo HFA,  $\Sigma$  tiene la representación (12), a continuación se reconstruirá la matriz  $\Sigma$  de  $p = 24$  variables con  $r = 3$  factores a partir de:

$$\delta_1^2 = 20, \delta_2^2 = 16, \delta_3^2 = 12, \psi^2 = 4$$

De tal manera que:

$$\Sigma_{24 \times 24} = [l_{i1} \quad l_{i2} \quad l_{i3}]_{24 \times 3} \begin{pmatrix} 20 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 12 \end{pmatrix}_{3 \times 3} \begin{bmatrix} l_{i1} \\ l_{i2} \\ l_{i3} \end{bmatrix}_{3 \times 24} + (4)I_{24 \times 24}$$

Ahora, la matriz  $\Sigma$  solo depende de los vectores columna  $l_{ij} \begin{cases} i = 1, \dots, 24 \\ j = 1, 2, 3 \end{cases}$ .

Se propondrán 3 distintas matrices de cargas, con la finalidad de medir la eficiencia del método de estimación por máxima verosimilitud y sus respectivos criterios de información, en aras de recuperar correctamente la estructura de ceros a distintos niveles de complejidad:

Matriz de cargas 1, contiene 48 ceros:

$$l_{ij} = \begin{cases} \sqrt{8} & i = 1, \dots, 24; \quad j = 1 \\ \frac{\sqrt{8}}{8} & i = 9, \dots, 16; \quad j = 2 \\ \frac{\sqrt{8}}{8} & i = 17, \dots, 24; \quad j = 3 \\ 0, & \text{resto} \end{cases}$$

En este primer caso, los espacios ocupados por los ceros estructurales tienen un patrón de distribución “sencillo” en cada vector de cargas, de manera que es el caso menos complejo por identificar.

Matriz de cargas 2, contiene 24 ceros:

$$l_{ij} = \begin{cases} \frac{\sqrt{24}}{24}, & i = 1, \dots, 8; j = 1 \\ \frac{\sqrt{12}}{12}, & i = 1, \dots, 6; j = 2 \\ \frac{\sqrt{12}}{12}, & i = 13, \dots, 18; j = 3 \\ -\frac{\sqrt{12}}{12}, & i = 7, \dots, 12; j = 2 \\ -\frac{\sqrt{12}}{12}, & i = 19, \dots, 24; j = 3 \\ 0, & \text{resto} \end{cases}$$

Matriz de cargas 3, contiene 40 ceros:

$$l_{ij} = \begin{cases} \frac{\sqrt{11}}{11}, & i = 1, \dots, 11; j = 1 \\ \frac{\sqrt{11}}{11}, & i = 14, \dots, 24; j = 2 \\ \frac{\sqrt{10}}{10}, & i = 10, \dots, 15; j = 3 \\ -\frac{\sqrt{10}}{10}, & i = 8, 9, 16, 17; j = 3 \\ 0, & \text{resto} \end{cases}$$

Se utilizará la función *rWishart* del paquete *stats* para realizar 10,000 simulaciones de la matriz  $\hat{S} = \frac{\Sigma}{n-1}$  considerando 7 tamaños de muestra ( $t$ ):

$$t = (50; 100; 200; 500; 1,000; 5,000; 10,000)$$

En la siguiente sección se reportará el desempeño del modelo HFA para:

- Determinar el número de factores a retener  $r$
- Detección de los ceros estructurales
- Detectar el número de casos en que se rechazó la hipótesis de igualdad de los  $p - r$  valores propios mediante la prueba de *Schott*

Adicional, se emplearán rotaciones de la matriz de cargas y técnicas para determinar el número de factores (secciones 2.5 y 2.6 respectivamente) para compararlos con los resultados obtenidos por el modelo HFA.

# CAPÍTULO 4

## 4. Resultados

En esta sección se presentarán los resultados:

1. Test de Schott
2. Aplicando criterios de información:
  - 2.1. Número de casos en los que se extrajo correctamente 3 factores
  - 2.2. Cantidad de ceros estructurales detectados
3. Según técnicas descritas en el capítulo 2:
  - 3.1. Número de casos en los que se extrajo correctamente 3 factores
  - 3.2. Cantidad de ceros estructurales detectados por rotaciones de la matriz de cargas

Para cada una de las simulaciones descritas en el capítulo 3.

### 4.1. Test Schott

$$\text{Hipótesis nula: } \lambda_4 = \dots = \lambda_{24} = \psi^2 = 4$$

Matriz	$\alpha$	50	100	200	500	1,000	5,000	10,000
Matriz de cargas 1	10%	1252 (12.52%)	1095 (10.95%)	1066 (10.66%)	1073 (10.73%)	976 (9.76%)	1022 (10.22%)	955 (9.55%)
	5%	670 (6.70%)	585 (5.85%)	552 (5.52%)	543 (5.43%)	515 (5.15%)	501 (5.01%)	461 (4.61%)
	1%	170 (1.70%)	126 (1.26%)	113 (1.13%)	103 (1.03%)	110 (1.10%)	97 (0.97%)	95 (0.95%)
Matriz de cargas 2	10%	1279 (12.79%)	1173 (11.73%)	1007 (10.07%)	1041 (10.41%)	976 (9.76%)	1001 (10.01%)	965 (9.65%)
	5%	712 (7.12%)	616 (6.16%)	513 (5.13%)	516 (5.16%)	475 (4.75%)	484 (4.84%)	474 (4.74%)
	1%	152 (1.52%)	114 (1.14%)	94 (0.94%)	95 (0.95%)	92 (0.92%)	101 (1.01%)	105 (1.05%)
Matriz de cargas 3	10%	1293 (12.93%)	1106 (11.06%)	1114 (11.14%)	1033 (10.33%)	1056 (10.56%)	997 (9.97%)	1017 (10.17%)
	5%	668 (6.68%)	565 (5.65%)	592 (5.92%)	527 (5.27%)	561 (5.61%)	512 (5.12%)	518 (5.18%)
	1%	142 (1.42%)	136 (1.36%)	120 (1.20%)	92 (0.92%)	112 (1.12%)	124 (1.24%)	101 (1.01%)

Tabla 1: Número de casos en que se rechazó la hipótesis nula de la prueba de Schott en las simulaciones

*Nota:* Adaptado de "On the Spike Covariance Structure and Factor Analysis with Homoscedastic Errors", de Vera, 2023

Los valores críticos basados en la distribución chi cuadrado (21) al 10%, 5% y 1% de significancia para la prueba de Schott serían: 257.88, 266.38, 282.81 respectivamente.

En la tabla 1 se muestran: el número de casos en los que se rechaza la hipótesis nula y en paréntesis el nivel de significancia empírico<sup>5</sup>  $\alpha^*$

En las figuras que se muestran a continuación se han graficado 3 líneas horizontales que corresponden a los niveles de significancia teóricos: 1%, 5%, 10%.

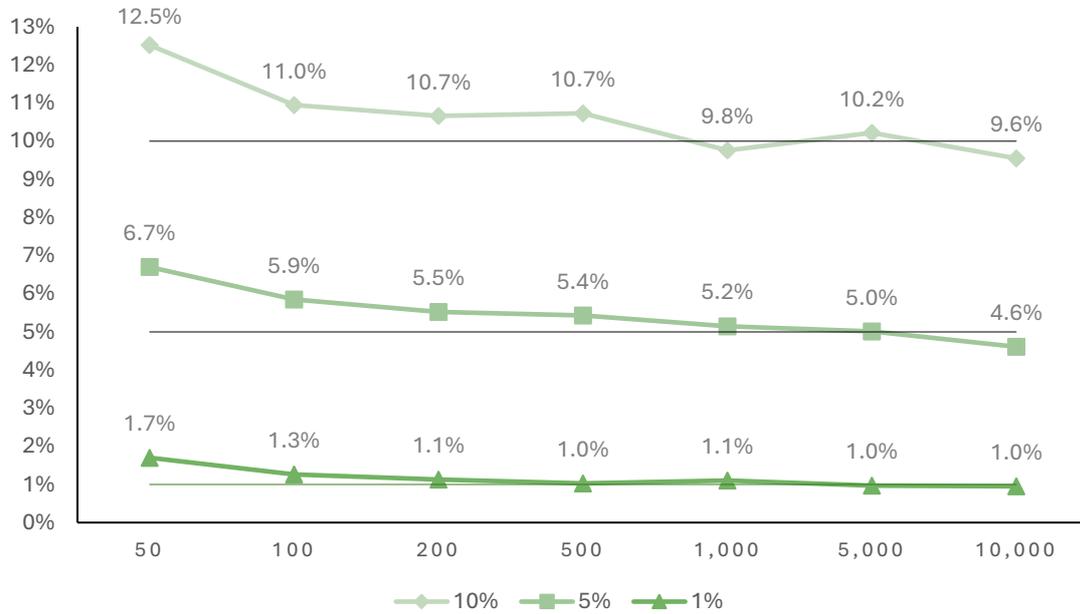


Figura 2: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 1

Fuente: Elaboración propia

Utilizando valores críticos con  $\alpha = 10\%$  y  $5\%$  las simulaciones a partir de la **matriz de cargas 1**, indicarían que se necesitarían construir matrices de covarianzas con más de 500 datos para obtener un  $\alpha^* \leq 10\%$  y  $\alpha^* \leq 5\%$  respectivamente.

Mientras que a partir de 200 datos se obtiene  $\alpha^* \leq 1\%$  utilizando un valor crítico con  $\alpha = 1\%$ .

<sup>5</sup> Representa el ratio entre la cantidad de veces en que se rechazó la hipótesis nula y el número de matrices de covarianzas distintas simuladas: 10,000

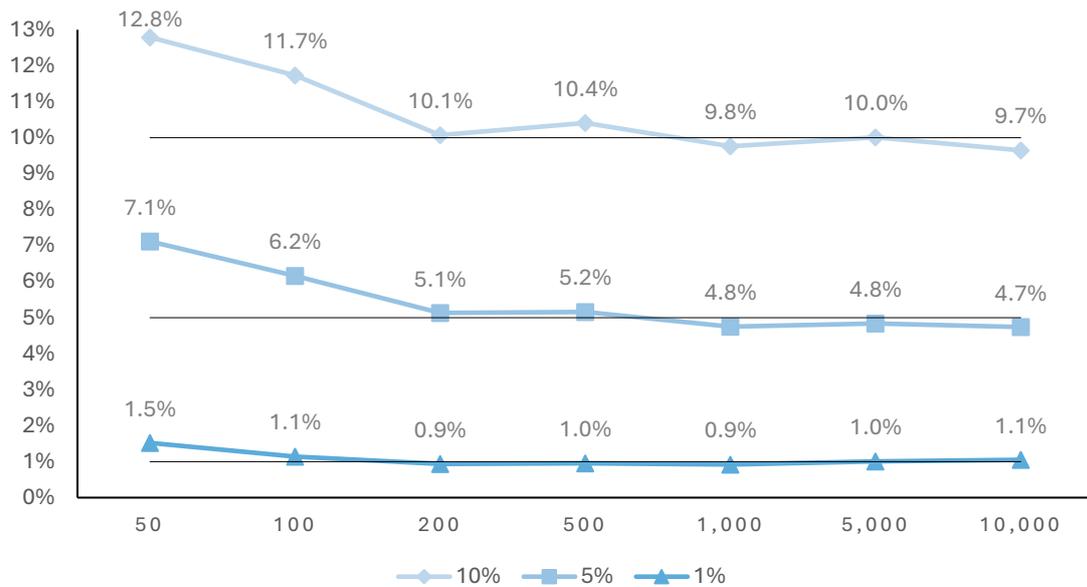


Figura 3: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 2

Fuente: Elaboración propia

Se obtuvieron resultados similares a los descritos en la matriz de cargas 1, se obtienen  $\alpha = 10\%$  y  $\alpha = 5\%$  a partir de 200 observaciones. Mientras que se alcanzó  $\alpha = 1\%$  a partir de las 100 observaciones

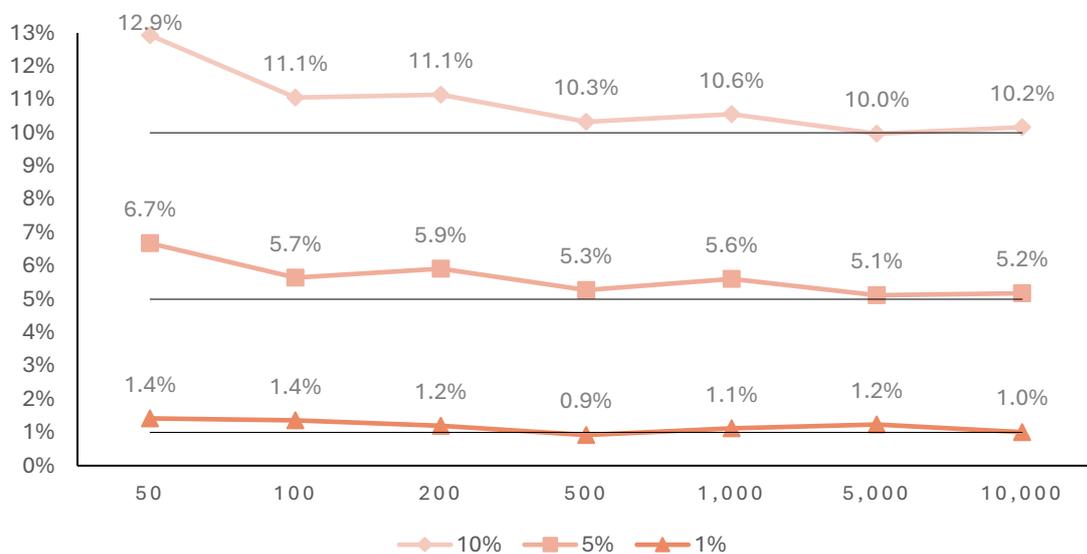


Figura 4: Cociente entre el número de casos en los que se rechaza la hipótesis nula mediante una prueba de Schott y 10,000 simulaciones en el conjunto 3

Fuente: Elaboración propia

No se alcanzó  $\alpha \leq 5\%$  estrictamente, no obstante se necesitó de más de 1,000 observaciones para obtener  $\alpha = 10\%$  y más de 500 observaciones para obtener  $\alpha = 1\%$ .

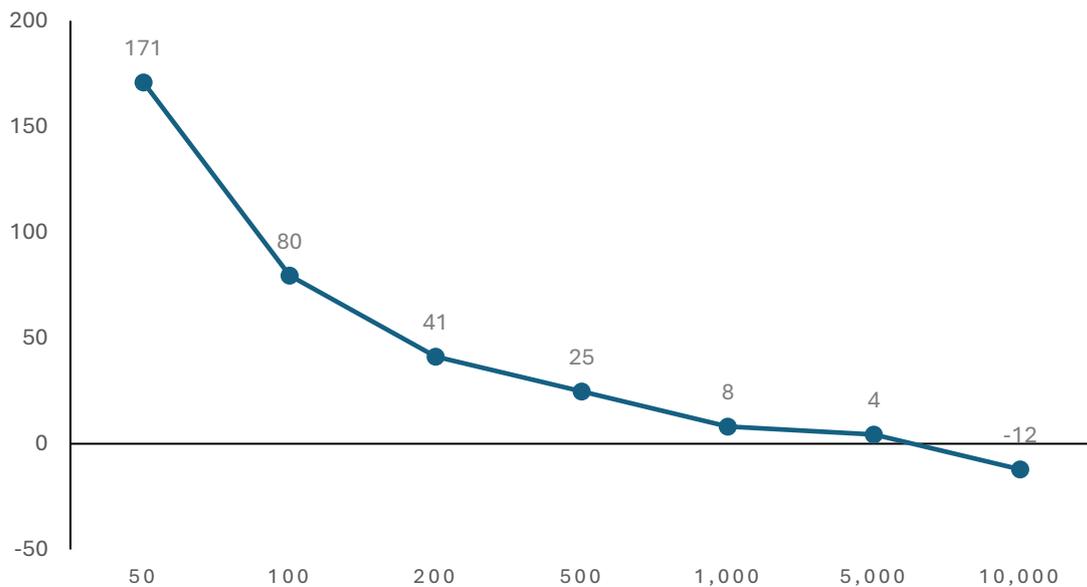


Figura 5: Cantidad de casos excedentes en los que se rechazó la hipótesis nula mediante una prueba de Schott

Fuente: Elaboración propia

Sin considerar el tipo de matriz de cargas que generó  $\Sigma$  ni el nivel de significancia del valor crítico, en la figura 5 se muestra el promedio de veces en los que se rechazó la hipótesis nula por encima de lo esperado, para cada tamaño de muestra analizado.

En términos generales se puede notar que a medida que crece el tamaño de muestra disminuye la cantidad de rechazos excedentes (ver Apéndice A).

No obstante cabe destacar que el error relativo más alto fue de apenas 1.71%.

## 4.2. Factores

	Técnica	50	100	200	500	1,000	5,000	10,000
Matriz de cargas 1	EKC	8,365	9,992	10,000	10,000	10,000	10,000	10,000
	HULL	7,058	7,845	8,430	8,686	8,820	8,799	8,837
	KGC	0	1	123	9,537	10,000	10,000	10,000
	PARALLEL	9,606	9,991	10,000	10,000	10,000	10,000	10,000
	SMT	1,555	6,765	8,606	9,189	9,381	9,462	9,510
	AIC	8,700	9,050	8,951	8,798	8,697	8,405	8,302
	BIC	3,425	9,436	10,000	10,000	10,000	10,000	10,000
	MBIC	8	1,424	9,732	10,000	10,000	10,000	10,000
	GIC	6	53	2,980	9,992	10,000	10,000	10,000
Matriz de cargas 2	EKC	7,878	9,962	10,000	10,000	10,000	10,000	10,000
	HULL	7,169	7,961	8,548	8,762	8,811	8,879	8,843
	KGC	0	0	271	9,856	10,000	10,000	10,000
	PARALLEL	9,560	9,992	10,000	10,000	10,000	10,000	10,000
	SMT	1,523	6,582	8,574	9,228	9,414	9,491	9,530
	AIC	8,701	8,964	8,965	8,811	8,681	8,394	8,347
	BIC	3,471	9,430	10,000	10,000	10,000	10,000	10,000
	MBIC	1	1,500	9,729	10,000	10,000	10,000	10,000
	GIC	0	80	2,901	9,995	10,000	10,000	10,000
Matriz de cargas 3	EKC	6,928	9,896	10,000	10,000	10,000	10,000	10,000
	HULL	6,594	7,780	8,302	8,580	8,685	8,725	8,667
	KGC	0	0	49	8,796	10,000	10,000	10,000
	PARALLEL	9,337	9,975	10,000	10,000	10,000	10,000	10,000
	SMT	1,498	6,614	8,539	9,241	9,344	9,479	9,464
	AIC	8,763	9,005	8,959	8,781	8,599	8,394	8,338
	BIC	3,452	9,437	10,000	10,000	10,000	10,000	10,000
	MBIC	4	1,462	9,731	10,000	10,000	10,000	10,000
	GIC	1	78	2,872	9,993	10,000	10,000	10,000

Tabla 2: Casos en los que se extrajo exactamente 3 factores en 10,000 simulaciones para cada matriz de cargas

Fuente: Elaboración propia

Independientemente del tipo de matriz de cargas utilizado en las simulaciones, el número de casos en los que se extrae exactamente 3 factores es muy similar por cada técnica empleada<sup>6</sup>. (ver Apéndice B)

El criterio de información AIC demostró ser el más eficiente con  $n = 50$  observaciones, no obstante **no es consistente al aumentar el número de observaciones**. BIC por otro lado obtuvo los mejores resultados con  $n \geq 100$ .

Sin embargo en los tres escenarios estudiados, el método de análisis paralelo (*PARALLEL*) supero a todos los criterios analizados, ya que escogió exactamente 3 factores con 95% de efectividad (promedio) cuando  $n = 50$ , 99.8% para  $n = 100$  y 100% con  $n \geq 200$ .

<sup>6</sup> La desviación estándar más alta fue 731 de 10,000 simulaciones y se obtuvo al aplicar el criterio de Kaiser (EKC)

## 4.3. Ceros Estructurales

### 4.3.1. Método y puntos de corte

Mediante la función *factanal* de la librería *stats* en R, se estimó la matriz de cargas utilizando el método de máxima verosimilitud considerando  $r = 3$  factores. Posteriormente, utilizando las funciones de la librería *GPARotation* se rotó la matriz de cargas estimada con los criterios descritos en la sección 2.6.

Se considera que una simulación recuperó **totalmente** la estructura de ceros si y solo si detectó todos los ceros presentes en la matriz de cargas.

Por defecto la función *factanal* utiliza como punto de corte 0.1, de forma que una carga  $\hat{l}_{ij}$  será considerada “en cero” si:

$$-0.1 \leq \hat{l}_{ij} \leq 0.1$$

A continuación se muestran los resultados en los diferentes esquemas de la matriz de cargas:

	Criterio	50	100	200	500	1,000	5,000	10,000
<b>Matriz de cargas 1</b>	No rotar	0	0	0	111	1,156	8,505	9,819
	Varimax	0	0	9	4,321	9,631	10,000	10,000
	Quartimax	0	0	9	4,255	9,615	10,000	10,000
	Equamax	0	0	8	4,338	9,651	10,000	10,000
	Parsimax	0	0	8	4,341	9,654	10,000	10,000
	AIC	3	1	1	3	2	1	0
	BIC	93	641	1,339	2,392	3,499	5,682	6,465
	MBIC	362	3,100	5,350	6,840	7,817	8,825	9,109
	GIC	353	3,568	7,030	8,616	9,213	9,718	9,791
<b>Matriz de cargas 2</b>	No rotar	0	0	6	371	1,972	9,061	9,946
	Varimax	0	0	0	0	0	0	0
	Quartimax	0	0	0	0	0	0	0
	Equamax	0	0	0	0	0	0	0
	Parsimax	0	0	0	0	0	0	0
	AIC	0	2	1	0	4	1	2
	BIC	0	65	283	466	670	1,068	1,244
	MBIC	0	31	949	1,716	2,061	2,829	3,175
	GIC	0	8	716	3,790	5,521	7,680	8,088
<b>Matriz de cargas 3</b>	No rotar	0	0	0	136	1,280	8,409	9,816
	Varimax	0	0	0	275	1,650	7,839	9,412
	Quartimax	0	0	1	505	2,881	9,269	9,924
	Equamax	0	0	0	176	1,107	6,492	8,504
	Parsimax	0	0	0	121	700	4,930	7,157
	AIC	0	2	3	4	5	4	9
	BIC	35	568	1,577	3,119	4,139	6,197	6,771
	MBIC	28	1,738	5,015	7,210	8,264	9,188	9,356
	GIC	22	1,576	6,266	8,727	9,517	9,858	9,875

Tabla 3: Casos en 10,000 simulaciones en que se recuperó totalmente la estructura de ceros para cada matriz de cargas

Fuente: Elaboración propia

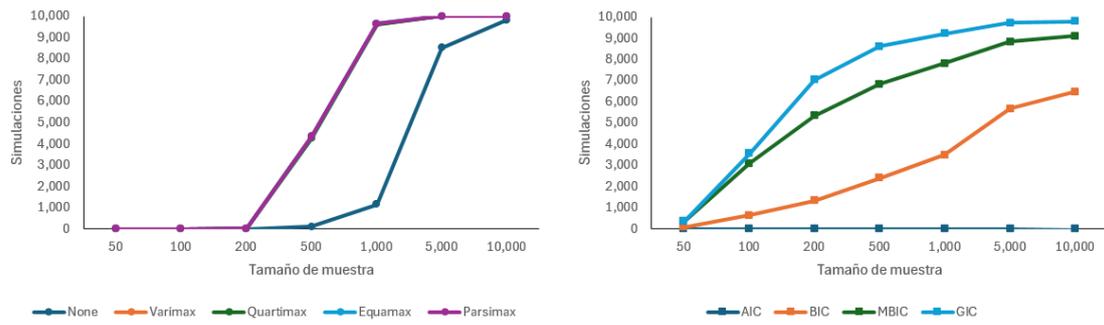


Figura 6: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 1 en 10,000 simulaciones

Fuente: Elaboración propia

Todas las rotaciones recuperan aproximadamente la misma cantidad de ceros, mientras que GIC es el criterio de información con mejores resultados.

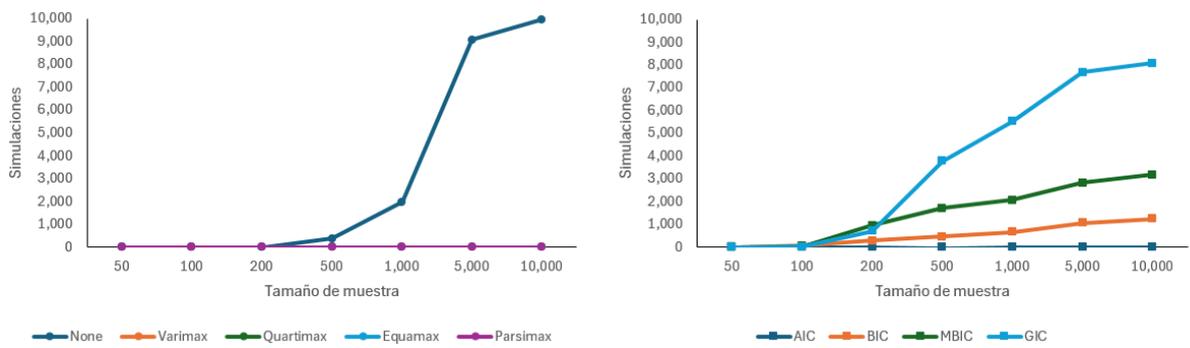


Figura 7: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 2 en 10,000 simulaciones

Fuente: Elaboración propia

No rotar la matriz recupera más ceros que rotarla, en  $n \leq 100$  no se recupera ningún cero. GIC obtiene los mejores resultados a partir de  $n > 100$ .

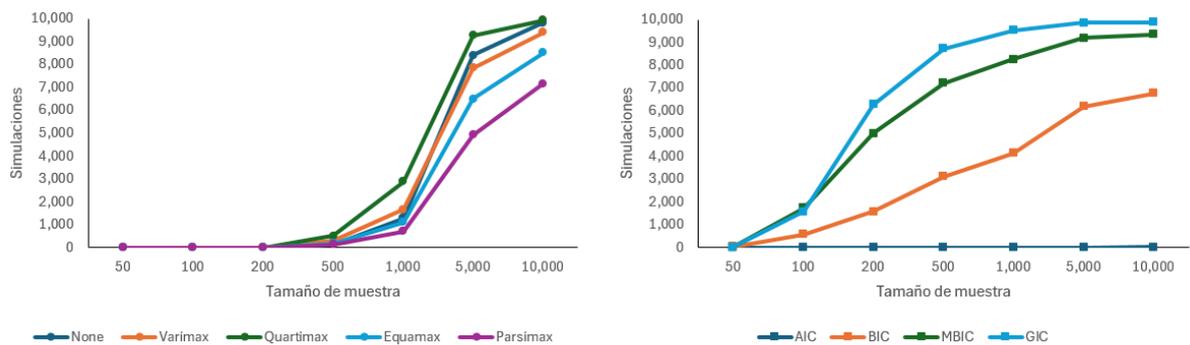


Figura 8: Casos en que se recuperó totalmente la estructura de ceros en la matriz de cargas 3 en 10,000 simulaciones

Fuente: Elaboración propia

Con  $n \leq 200$  no se recupera ningún cero, a partir de  $n > 200$  Quartimax produce los mejores resultados. Nuevamente GIC recupera la mayor cantidad de ceros.

# CAPÍTULO 5

## 5. Conclusiones y Recomendaciones

### 5.1. Conclusiones

#### Prueba de Schott:

- La modificación a la prueba de Schott propuesta por Vera rechaza ligeramente más veces (1.71% por encima) de lo esperado con  $n = 50$  observaciones
- No obstante  $n \geq 100$  rechaza entre 0.04% y 0.80% por encima de lo esperado, por lo que detecta eficientemente el número de vectores propios distintos  $\lambda_1, \lambda_2, \lambda_3$  del resto.
- El valor de significancia empírico  $\alpha^*$  converge hacia el valor de significancia teórico  $\alpha$  a medida que aumenta el tamaño de muestra, reduciendo así el número de casos en los que se rechaza (incorrectamente) la hipótesis de igualdad en los 21 valores propios restantes.

#### Determinación del número de factores:

- La estructura de la matriz de cargas no tiene efecto sobre el número de casos en los que se detecta extraer exactamente 3 factores, a diferencia de la técnica empleada o el tamaño de muestra  $n$ .
- AIC funciona mejor en muestras “pequeñas”  $n \leq 100$ , pero no aumenta su efectividad notablemente a medida que aumenta  $n$  y de hecho con  $n = 10,000$  no superó el 80% de efectividad.
- Los demás criterios de información alcanzan el 100% de efectividad a partir de  $n \geq 500$ .
- BIC obtiene más del 90% de efectividad con apenas  $n = 100$  observaciones y tiene el mejor desempeño entre los criterios de información
- No obstante la mejor técnica es el Análisis Paralelo obteniendo la efectividad más alta con  $n = 50$  y 100 observaciones (más del 90% y 99% respectivamente). Además de alcanzar el 100% de efectividad a partir de  $n = 200$ .

### **Recuperación de ceros en la estructura de la matriz de cargas:**

- Los criterios de rotación no mostraron resultados consistentes entre las diferentes matrices de cargas, a diferencia de los criterios de información.
- En la matriz de cargas 1, todas las rotaciones recuperaron aproximadamente la misma cantidad de ceros estructurales a partir de  $n \geq 200$  observaciones, este número creció con el tamaño de muestra  $n$ . Pero en muestras pequeñas ( $n < 200$ ) no se pudo recuperar ninguna estructuras de ceros.
- En la matriz de cargas 2 se recuperaron más ceros estructurales al no rotar la matriz, además ninguna rotación logro recuperar estructuras de ceros aún con  $n = 10,000$ .
- Finalmente en la matriz de cargas 3 solo se pudieron recuperar ceros estructurales a partir de  $n > 200$ , la rotación Quartimax obtuvo los mejores resultados alcanzando el 100% de efectividad con  $n = 10,000$  (sin considerar a los criterios de información)
- GIC obtuvo los mejores resultados en los tres escenarios analizados entre todas las técnicas estudiadas.

## **5.2. Recomendaciones**

- Antes de aplicar el modelo HFA al conjunto de datos, se debe verificar la redundancia de factores (o igualdad de los  $p - r$  últimos valores propios de  $\Sigma$ ).
- Para detectar la cantidad de factores a extraer, se puede utilizar el método de Análisis Paralelo o el criterio de información BIC siempre que se tengan  $n > 50$  observaciones.
- Rotar la matriz de cargas puede aumentar las saturaciones en ciertos factores o aportar a la interpretación de los constructos formados, pero si se desea conservar la estructura de ceros se deben considerar la estimación del método HFA y puntualmente del criterio de información GIC.

# CAPÍTULO 6

## 6. Referencias

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*.
- Auerswald, M., & Moshagen, M. (2019). How to Determine the Number of Factors to Retain in Exploratory Factor Analysis: A Comparison of Extraction Methods Under Realistic Conditions. *American Psychological Association*.
- Bernaards, C. (2005). Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. *University of California, Los Angeles*, 1-24.
- Braeken, J., & van Assen, M. (2017). An Empirical Kaiser Criterion. *Psychological Methods*, 450-466.
- Carroll, J. (1953). An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 23-38.
- Crawford, C., & Ferguson, G. (1970). A general rotation criterion and its use in orthogonal rotation. *Psychometrika*, 321-332.
- Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1-8.
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 149-161.
- Heywood, H. (1931). On Finite Sequences of Real Numbers. *Proceedings of the Royal Society of London*, 486-501.
- Horn, J. (1965). Rationale and test for the number of factors in factor analysis. *Psychometrika*, 179-185.
- Johnson, R., & Wichern, D. (2007). *Applied Multivariate Statistical Analysis*. New Jersey: Pearson Prentice Hall.
- Kaiser, H. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 187-200.
- Kaiser, H. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 141-151.

- Kim, Y., Kwon, S., & Choi, H. (2012). Consistent Model Selection Criteria on High Dimensions. *Journal of Machine Learning Research* 13, 1037-1057.
- Lorenzo-Seva, U., Timmerman, M., & Kiers, H. (2011). The Hull Method for Selecting the Number of Common Factors. *Multivariate Behav Res*, 340-364.
- Multivariate, A. (s.f.).
- Neuhauser, J., & Wrigley, C. (1954). The quartimax method: An analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology*, 81-91.
- Schott, J. (2006). A high-dimensional test for the equality of the smallest eigenvalues of a covariance matrix. *Journal of Multivariate Analysis*, 827-843.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 461-464.
- Spearman, C. (1904). 'General intelligence,' objectively determined and measured. *The American Journal of Psychology*, 201-293.
- Thurstone, L. (1935). *The Vectors of Mind*. University of Chicago Press: Chicago.
- Vera, F. (2023). On the Spike Covariance Structure and Factor Analysis with Homoscedastic Errors. *Unpublished paper*.
- Wang, H., Li, B., & Leng, C. (2009). Shrinkage Tuning Parameter Selection with a Diverging Number of Parameters. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 671-683.

# CAPÍTULO 7

## 7. Apéndices y anexos

### APÉNDICE A

Matriz	$\alpha$	50	100	200	500	1,000	5,000	10,000	Promedio
Matriz de cargas 1	10%	2.52%	0.95%	0.66%	0.73%	-0.24%	0.22%	-0.45%	<b>0.63%</b>
	5%	1.70%	0.85%	0.52%	0.43%	0.15%	0.01%	-0.39%	<b>0.47%</b>
	1%	0.70%	0.26%	0.13%	0.03%	0.10%	-0.03%	-0.05%	<b>0.16%</b>
Matriz de cargas 2	10%	2.79%	1.73%	0.07%	0.41%	-0.24%	0.01%	-0.35%	<b>0.63%</b>
	5%	2.12%	1.16%	0.13%	0.16%	-0.25%	-0.16%	-0.26%	<b>0.41%</b>
	1%	0.52%	0.14%	-0.06%	-0.05%	-0.08%	0.01%	0.05%	<b>0.08%</b>
Matriz de cargas 3	10%	2.93%	1.06%	1.14%	0.33%	0.56%	-0.03%	0.17%	<b>0.88%</b>
	5%	1.68%	0.65%	0.92%	0.27%	0.61%	0.12%	0.18%	<b>0.63%</b>
	1%	0.42%	0.36%	0.20%	-0.08%	0.12%	0.24%	0.01%	<b>0.18%</b>
<b>Promedio</b>		<b>1.71%</b>	<b>0.80%</b>	<b>0.41%</b>	<b>0.25%</b>	<b>0.08%</b>	<b>0.04%</b>	<b>-0.12%</b>	<b>0.45%</b>

Fuente: Elaboración propia

**Cantidad de veces en 10,000 simulaciones en que se rechazó la hipótesis nula diferente a lo esperado**

### APÉNDICE B

Técnica	50	100	200	500	1,000	5,000	10,000
EKC	7,724 (731)	9,950 (49)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)
HULL	6,940 (305)	7,862 (92)	8,427 (123)	8,676 (91)	8,772 (75)	8,801 (77)	8,782 (100)
KGC	0 (0)	0 (1)	148 (113)	9,396 (544)	10,000 (0)	10,000 (0)	10,000 (0)
PARALLEL	9,501 (144)	9,986 (10)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)
SMT	1,525 (29)	6,654 (98)	8,573 (34)	9,219 (27)	9,380 (35)	9,477 (15)	9,501 (34)
AIC	8,721 (36)	9,006 (43)	8,958 (7)	8,797 (15)	8,659 (53)	8,398 (6)	8,329 (24)
BIC	3,449 (23)	9,434 (4)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)
MBIC	4 (4)	1,462 (38)	9,731 (2)	10,000 (0)	10,000 (0)	10,000 (0)	10,000 (0)
GIC	2 (3)	70 (15)	2,918 (56)	9,993 (2)	10,000 (0)	10,000 (0)	10,000 (0)

Fuente: Elaboración propia

**Promedio y (desviación estándar) del número de casos en los que se extrajo exactamente 3 factores en las 10,000 simulaciones**

## APÉNDICE C

Matriz de cargas 1		
$\sqrt{8}/8$	0	0
0	$\sqrt{8}/8$	0
0	0	$\sqrt{8}/8$

Matriz de cargas 3		
$\sqrt{24}/24$	$\sqrt{12}/12$	0
$\sqrt{24}/24$	$-\sqrt{12}/12$	0
$\sqrt{24}/24$	0	$\sqrt{12}/12$
$\sqrt{24}/24$	0	$-\sqrt{12}/12$

Matriz de cargas 3		
$\sqrt{11}/11$	0	0
$\sqrt{11}/11$	0	$-\sqrt{10}/10$
$\sqrt{11}/11$	0	$-\sqrt{10}/10$
$\sqrt{11}/11$	0	$\sqrt{10}/10$
$\sqrt{11}/11$	0	$\sqrt{10}/10$
0	0	$\sqrt{10}/10$
0	0	$\sqrt{10}/10$
0	$\sqrt{11}/11$	$\sqrt{10}/10$
0	$\sqrt{11}/11$	$\sqrt{10}/10$
0	$\sqrt{11}/11$	$-\sqrt{10}/10$
0	$\sqrt{11}/11$	$-\sqrt{10}/10$
0	$\sqrt{11}/11$	0

*Fuente: Elaboración propia*

## ANEXO 1

	<b>Técnica</b>	<b>50</b>	<b>100</b>	<b>200</b>	<b>500</b>	<b>1,000</b>	<b>5,000</b>	<b>10,000</b>
<b>Matriz de cargas 1</b>	EKC	2.8343	2.9992	3	3	3	3	3
	HULL	2.7396	2.8102	2.8434	2.8628	2.8775	2.8783	2.8832
	KGC	7.4904	6.4263	4.8372	3.0465	3	3	3
	PARALLEL	3.0159	3.0009	3	3	3	3	3
	SMT	6.0694	3.5215	3.1802	3.0953	3.0715	3.0594	3.0546
	AIC	3.0967	3.102	3.1109	3.1272	3.139	3.1702	3.1829
	BIC	2.1279	2.9432	3	3	3	3	3
	MBIC	1.0209	1.6875	2.9732	3	3	3	3
	GIC	1.0149	1.1189	2.0665	2.9992	3	3	3
<b>Matriz de cargas 2</b>	EKC	2.7841	2.9962	3	3	3	3	3
	HULL	2.7502	2.8208	2.8595	2.8735	2.8793	2.8878	2.8841
	KGC	7.4333	6.3329	4.6705	3.0144	3	3	3
	PARALLEL	3.0035	3.0006	3	3	3	3	3
	SMT	6.0984	3.5459	3.1813	3.0914	3.0677	3.0568	3.0526
	AIC	3.0942	3.1109	3.1111	3.1275	3.1401	3.1721	3.1768
	BIC	2.1371	2.9428	3	3	3	3	3
	MBIC	1.0223	1.7021	2.9729	3	3	3	3
	GIC	1.0145	1.1266	2.0621	2.9995	3	3	3
<b>Matriz de cargas 3</b>	EKC	2.6895	2.9896	3	3	3	3	3
	HULL	2.6927	2.8188	2.8416	2.8589	2.8686	2.8724	2.8667
	KGC	7.5743	6.5596	5.0516	3.1209	3	3	3
	PARALLEL	3.0024	3.0021	3	3	3	3	3
	SMT	6.1053	3.5392	3.192	3.0884	3.0753	3.0583	3.059
	AIC	3.0926	3.108	3.1115	3.1304	3.152	3.1724	3.1786
	BIC	2.131	2.9432	3	3	3	3	3
	MBIC	1.0199	1.6907	2.9731	3	3	3	3
	GIC	1.013	1.1281	2.0536	2.9993	3	3	3

**Cantidad promedio de factores a extraer por cada técnica**

## ANEXO 2

	<b>Criterio</b>	<b>50</b>	<b>100</b>	<b>200</b>	<b>500</b>	<b>1,000</b>	<b>5,000</b>	<b>10,000</b>
<b>Matriz de cargas 1</b>	No rotar	25.98%	33.21%	44.27%	62.70%	78.01%	98.16%	99.84%
	Varimax	42.97%	61.32%	80.78%	96.88%	99.80%	100%	100%
	Quartimax	42.99%	61.28%	80.74%	96.86%	99.80%	100%	100%
	Equamax	42.91%	61.31%	80.78%	96.89%	99.81%	100%	100%
	Parsimax	42.88%	61.29%	80.78%	96.89%	99.81%	100%	100%
	AIC	56.77%	57.96%	58.86%	59.90%	60.85%	61.26%	61.46%
	BIC	71.82%	77.40%	80.83%	85.15%	88.27%	92.19%	93.45%
	MBIC	79.62%	85.09%	89.44%	92.60%	94.91%	96.87%	97.54%
	GIC	79.98%	86.48%	92.57%	96.94%	98.82%	99.68%	99.80%
<b>Matriz de cargas 2</b>	No rotar	25.15%	30.95%	40.10%	59.69%	76.32%	98.88%	99.95%
	Varimax	27.82%	28.89%	27.63%	23.44%	18.81%	7.51%	4.39%
	Quartimax	28.26%	29.38%	28.05%	23.14%	17.28%	3.64%	1.72%
	Equamax	27.60%	28.69%	27.58%	23.89%	19.91%	10.78%	7.27%
	Parsimax	27.39%	28.58%	27.62%	24.24%	20.86%	13.80%	10.51%
	AIC	47.19%	43.98%	44.14%	45.45%	46.17%	46.64%	46.92%
	BIC	61.21%	58.68%	61.19%	67.50%	70.56%	74.22%	75.49%
	MBIC	73.75%	70.77%	70.19%	76.97%	79.78%	83.07%	84.12%
	GIC	74.43%	75.54%	78.16%	86.26%	90.96%	95.65%	96.68%
<b>Matriz de cargas 3</b>	No rotar	25.16%	32.64%	42.68%	60.71%	76.38%	98.26%	99.85%
	Varimax	32.91%	44.94%	59.00%	75.54%	83.75%	96.40%	99.18%
	Quartimax	33.72%	46.54%	61.55%	80.01%	89.08%	99.17%	99.94%
	Equamax	32.47%	44.10%	57.53%	72.94%	80.10%	93.24%	97.36%
	Parsimax	32.14%	43.37%	56.36%	70.64%	76.69%	88.90%	93.91%
	AIC	52.48%	54.29%	56.58%	58.30%	59.43%	60.00%	60.39%
	BIC	66.99%	73.15%	79.94%	87.16%	90.55%	93.61%	94.24%
	MBIC	76.34%	80.32%	86.12%	93.35%	96.35%	97.97%	98.19%
	GIC	76.80%	82.13%	88.88%	96.24%	99.10%	99.88%	99.91%

### Porcentaje de verdaderos ceros promedio recuperados por cada técnica

$$\% = \frac{\text{Verdaderos ceros recuperados en promedio}}{\text{Verdaderos ceros}}$$