Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ciencias Matemáticas
Unidad de Posgrado

Método Tucker3 en el análisis factorial dinámico para evaluar la dinámica macroeconómica del Perú 2007 – 2021

TESIS

Para optar el Grado Académico de Magíster en Estadística Matemática

AUTOR

José Luis VÁSQUEZ PÉREZ

ASESOR

Mg. Emma Norma CAMBILLO MOYANO

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Vásquez, J. (2025). *Método Tucker3 en el análisis factorial dinámico para evaluar la dinámica macroeconómica del Perú 2007 – 2021*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	José Luis Vásquez Pérez
Tipo de documento de identidad	DNI
Número de documento de identidad	43829088
URL de ORCID	https://orcid.org/0000-0002-2815-2044
Datos de asesor	
Nombres y apellidos	Emma Norma Cambillo Moyano
Tipo de documento de identidad	DNI
Número de documento de identidad	15377390
URL de ORCID	https://orcid.org/0000-0003-3173-9425
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Oscar Antonio Robles Villanueva
Tipo de documento de identidad	DNI
Número de documento de identidad	32762171
Miembro del jurado 1	
Nombres y apellidos	Maria Estela Ponce Aruneri
Tipo de documento de identidad	DNI
Número de documento de identidad	10010612
Miembro del jurado 2	
Nombres y apellidos	Ysela Dominga Agüero Palacios
Tipo de documento de identidad	DNI
Número de documento de identidad	09059647
Miembro del jurado 3	
Nombres y apellidos	Fernando Cesar Camones Gonzales

Tipo de documento de identidad	DNI
Número de documento de identidad	10459353
Datos de investigación	
Línea de investigación	A.3.2.1. Análisis Multivariante
Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: Lima Provincia: Lima Distrito: Lima Latitud: -12.056445 Longitud: -77.085994
Año o rango de años en que se realizó la investigación	2022 - 2024
URL de disciplinas OCDE	Estadísticas, Probabilidad **https://purl.org/pe-repo/ocde/ford#1.01.03**

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS MATEMÁTICAS
VICEDECANATO DE INVESTIGACIÓN Y POSGRADO
UNIDAD DE POSGRADO

ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGÍSTER

Siendo las 10:05 horas del día viernes veinticuatro de enero de dos mil veinticinco, en el Laboratorio de Posgrado de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos, el Jurado Examinador de tesis, presidido por el Dr. Oscar Antonio Robles Villanueva e integrado por los siguientes miembros: Mg. María Estela Ponce Aruneri (Jurado Evaluador), Mg. Ysela Dominga Agüero Palacios (Jurado Informante), Mg. Fernando Cesar Camones Gonzales (Jurado Informante) y la Mg. Emma Norma Cambillo Moyano como Miembro Asesor; se reunieron para la sustentación de la tesis titulada: «MÉTODO TUCKER3 EN EL ANÁLISIS FACTORIAL DINÁMICO PARA EVALUAR LA DINÁMICA MACROECONÓMICA DEL PERÚ 2007 - 2021» presentado por el Br. José Luis Vásquez Pérez, para optar el Grado Académico de Magister en Estadística Matemática.

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, según tabla adjunta, habiendo obtenido el Bachiller **JOSÉ LUIS VÁSQUEZ PÉREZ**, el calificativo de 17 Muy Bueno

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Examinador recomienda que el Consejo de Facultad apruebe el otorgamiento del **Grado Académico de Magíster en Estadística Matemática** al Bachiller **JOSÉ LUIS VÁSQUEZ PÉREZ.**

Siendo las 11:15 horas, se levantó la sesión, firmando para constancia la presente Acta:

Dr. Oscar Antonio Robles Villanueva
Presidente

Mg. María Estela Ponce Aruneri
Miembro

Mg. Ysela Dominga Agüero Palacios
Miembro

Mg. Fernando Cesar Camones Gonzales
Miembro

Mg. Emma Norma Cambillo Moyano
Miembro Asesor

CERTIFICADO DE SIMILITUD

Yo, Emma Norma Cambillo Moyano en mi condición de asesora acreditada con Resolución Decanal N° 000076-2022-D-FCM/UNMSM de la Tesis, cuyo título es "**MÉTODO TUCKER3 EN EL ANÁLISIS FACTORIAL DINÁMICO PARA EVALUAR LA DINÁMICA MACROECONÓMICA DEL PERÚ 2007 – 2021**, presentado por el bachiller

JOSÉ LUIS VÁSQUEZ PÉREZ

para optar el grado de Magister en Estadística Matemática.

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante el software de similitud textual TURNITIN, el documento evaluado cuenta con el porcentaje de17% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado correspondiente.

Mg. Emma N. Cambillo Moyano

DNI 15377390

Dedicado a

Antuan, Alanis y Daniel

"Transire suum pectus mundoque potiri"

Arquímedes

Agradecimientos

Deseo expresar mi profundo agradecimiento a mi alma mater, la Universidad Nacional Mayor de San Marcos. Esta institución, a pesar de los desafíos y altibajos, se mantiene como la principal casa de estudios de nuestro país. Debo reconocer que sin su influencia y educación, yo no sería la persona que soy hoy ni habría emprendido este viaje académico que, a pesar de sus dificultades, me ha permitido crecer y aprender.
Es importante destacar que, como economista, San Marcos jugó un papel fundamental al impulsarme a explorar mi pasión por la estadística. Fue en esta universidad donde descubrí mi amor por la investigación y el estudio, lo cual me ha llevado a profundizar en campos como la estadística. Por lo tanto, mi gratitud hacia San Marcos es doble, ya que me brindó las bases para mi desarrollo académico y profesional.
Sin embargo, quiero resaltar especialmente a una figura en esta travesía académica: Ledyard Tucker. A través de su trabajo y legado en estadística, he encontrado una fuente inagotable de inspiración. Las historias, teorías y métodos con las que contribuyó a la disciplina me han guiado en la realización de esta tesis. En la actualidad, más que nunca, disfrutamos de los frutos de su trabajo, sembrados a principios del siglo pasado. Esta tesis no sería posible sin su influencia y guía. Por todo ello, dedico este pequeño agradecimiento a la memoria de Ledyard Tucker, cuyo legado sigue vivo en cada página de esta investigación. Sus contribuciones a la estadística y a mi formación académica merecen un reconocimiento especial en este camino que he recorrido.
De la misma manera, deseo expresar mi agradecimiento a todos mis profesores de la Maestría en Estadística Matemática de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos por su dedicación y tiempo. En particular, quiero destacar a la profesora Emma Cambillo Moyano, cuya guía y apoyo continuo fueron fundamentales para llevar a cabo esta investigación; en el desánimo, encontré en sus palabras una mano para levantarme.

Índice general

Agradecimientos **3**

Abstract **8**

1 Introducción **10**

1.1 Situación problemática 13

1.2 Formulación del problema 16

1.3 Justificación teórica 18

1.4 Justificación práctica 20

1.5 Beneficios del uso del modelo Tucker3 21

1.6 Objetivos 22

1.6.1 Objetivo general 22

1.6.2 Objetivos específicos 22

2 Marco teórico **23**

2.1 Antecedentes 24

2.2 Conceptos previos 41

2.2.1 Análisis de componentes principales 41

2.2.2 Análisis factorial 42

2.2.3 Descomposición en valores singulares 43

2.2.4 Mínimos Cuadrados Alternados 43

2.3 El análisis de tres vías 45

2.4 Estructuración de la Matriz de Datos: matriciación o matricización . . 50

2.5 Modelo Tucker3 60

2.5.1 Propiedades del modelo Tucker3 64

2.5.2 Ilustrando el modelo Tucker3 65

2.5.3 Interpretación del arreglo central 68

2.5.4 Puntuaciones (*scores*) y cargas (*loadings*) en el modelo Tucker3 71
2.5.5 Estandarización de los datos 72
2.5.6 Estimación de parámetros del modelo Tucker3 78
2.5.7 Función de pérdida 79
2.5.8 Estabilidad de la estimación parámetros 80
2.5.9 Inercia 82
2.5.10 Casos particulares del modelo Tucker 3 84
2.5.11 Algoritmo Tucker3 89

3 Metodología **91**
3.1 Tipo de investigación 91
3.2 Muestra y variables 92
3.3 Obtención de datos 93
3.4 Procesamiento de los datos 94
3.5 Técnicas de recolección de datos 95

4 Dinámica macroeconómica en el Perú, 2007-2021 **96**
4.1 Análisis de las actividades económicas 98
4.2 Aplicación del modelo 104
4.2.1 Dimensionalidad de la solución 104
4.2.2 Estimación y representación de los componentes 107
4.2.3 Trayectorias 114

5 Resultados y discusión **118**

6 Conclusiones **121**
6.1 Sugerencias para investigaciones futuras 122

Bibliografía **124**

A Apéndice. Análisis de componentes principales **133**

B Apéndice. Análisis factorial **142**
B.0.1 Unicidad del modelo 146
B.0.2 Número máximo de factores comunes 147

B.0.3 Estimación de los parámetros 148
B.0.4 Estimación de los factores 149
B.0.5 Uso de la matriz de correlación en la estimación 151
B.0.6 Rotación de los factores 152

C Apéndice. Relación entre el análisis factorial y el análisis de componentes principales 154

D Apéndice. Descomposición en valores singulares 156

E Apéndice. La Descomposición en Valores singulares (SVD), el Análisis de Componentes principales (ACP) y el Análisis factorial 166

F Programación del modelo de Tucker 3 en R 169

Word count

Índice de tablas

Tabla 2.1: Notación para un arreglo central de tres vías 69

Tabla 4.1: Contribución de componentes a la inercia total de los dos modelos106

Tabla 4.2: Contribución de triadas de componentes a la inercia total . . . 107

Índice de figuras

Figura 2.1: Número de publicaciones sobre Tucker3, 1933-2021 26

Figura 2.2: Citas de artículos de tres y múltiples vías de investigadores de diferentes disciplinas. 27

Figura 2.3: Perfil de datos de tres vías 46

Figura 2.4: Arreglo de datos de tres vías $x_{ijk};\ i = 1, ..., I; j = 1, ..., J; k = 1, ..., K$ 46

Figura 2.5: Disposición de las vías individuos, variables y condiciones, vistas como *slices* de arreglos 47

Figura 2.6: Matriz de tres vías, cortada en fibras horizontales, verticales y profundas 48

Figura 2.7: Despliegues separados en los lados frontales de un arreglo de tres vías 51

Figura 2.8: Representación del modelo Tucker3 63

Figura 2.9: Centrado a través de la primera vía 75

Figura 2.10: Visualización de centrado y la normalización 77

Figura 2.11: Scree plot multiway 85

Figura 2.12: Modelo Tucker1 con $\mathbf{A_k} = \mathbf{A}$, pero diferentes $\mathbf{B_k}$ 87

Figura 2.13: Modelo Tucker1 con $\mathbf{A_k} = \mathbf{A}$ y $\mathbf{B_k} = \mathbf{B}$ 88

Figura 4.1: Producto Bruto Interno a precios constantes de 2007 99

Figura 4.2: Sectores con mayor Valor Agregado Bruto en la estructura sectorial, 2007-2021 (Porcentaje) 100

Figura 4.3: Sectores con mayor Valor Agregado Bruto en la estructura sectorial, 2007-2021 (Porcentaje) 101

Figura 4.4: Sectores con mayor menor Agregado Bruto en la estructura sectorial, 2007-2020 (Porcentaje) 102

Figura 4.5: Estructura sectorial, 2007-2021 (Porcentaje) 103
Figura 4.6: Scree plot multiway considerando Lima y Callao 105
Figura 4.7: Comparación de los resultados del criterio *Diff-Diff* 106
Figura 4.8: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Considerando Lima y Callao) 108
Figura 4.9: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Excluyendo Lima y Callao) 108
Figura 4.10: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Excluyendo Lima y Callao) 109
Figura 4.11: Distancias de los Departamentos del Perú en plano factorial (1,2) 110
Figura 4.12: Variables sobre el plano (1,2) en Tucker 3 (Excluyendo Lima y Callao) 112
Figura 4.13: Variables sobre el plano (1,2) en Tucker 3 (Aislando la actividad económica de Petróleo, Minería y gas) 112
Figura 4.14: Distancias de las variables en el plano (1,2) 113
Figura 4.15: Distancias de las variables en el plano (1,2) (sin considerar Petróleo, Minería y Gas 113
Figura 4.16: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2) 114
Figura 4.17: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2) 115
Figura 4.18: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2) 116
Figura 4.19: Trayectorias de los departamentos en el plano (X,Y) en 2007 . 117
Figura 4.20: Trayectorias de los departamentos en el plano (X,Y) en 2021 . 117

Abstract

This thesis focuses on the application of the Tucker 3 model in a three-way context to analyze macroeconomic dynamics at the departmental level in the period 2007-2021. A comprehensive review of the literature on three-way analysis was conducted, emphasizing the importance of understanding its theoretical aspects and practical applications, especially when time is a key dimension.

Although the Tucker 3 model has been used in various fields, its use in economics has been limited, prompting this research to explore its characteristics and applicability in this specific context. The model was successfully applied to a three-dimensional dataset representing departments, economic variables, and time periods, explaining 86,6 % of the total variability.

An essential aspect was the exclusion of Lima and Callao due to their disproportionate influence. Despite this exclusion, the component structure remained robust, supporting the validity of the model.

The results of the Tucker 3 model provided a deeper understanding of the macroeconomic dynamics of Peru's departments, identifying notable trends in the relationships between departments, variables, and time, including geographic affinities and the influence of location on regional economic relationships.

Resumen

Esta tesis se enfoca en la aplicación del modelo Tucker 3 en un contexto de tres vías para analizar la dinámica macroeconómica a nivel departamental en el período 2007-2021. Se realizó una revisión exhaustiva de la literatura sobre el análisis de tres vías, subrayando la relevancia de comprender sus aspectos teóricos y aplicaciones prácticas, especialmente cuando el tiempo es una dimensión clave.

A pesar de que el modelo Tucker 3 se ha utilizado en diversas áreas, su uso en economía ha sido limitado, lo que motivó esta investigación para explorar sus características y aplicabilidad en este contexto. El modelo se aplicó exitosamente a un conjunto de datos tridimensional que representaba departamentos, variables económicas y periodos de tiempo, logrando explicar el 86,6 % de la variabilidad total.

Un aspecto esencial fue la exclusión de Lima y Callao debido a su influencia desproporcionada. A pesar de esto, la estructura de componentes se mantuvo sólida, respaldando la validez del modelo.

Los resultados del modelo Tucker 3 proporcionaron una comprensión más profunda de la dinámica macroeconómica de los departamentos del Perú, identificando tendencias notables en las relaciones entre departamentos, variables y tiempo, incluyendo afinidades geográficas y la influencia de la ubicación en las relaciones económicas regionales.

Capítulo 1

Introducción

En la actual era digital, los datos se han convertido en un artículo de consumo debido a la cantidad masiva de información generada y recopilada por las organizaciones y las personas a diario. De hecho, los datos son un recurso valioso que permite mejorar la toma de decisiones, personalizar productos, optimizar procesos y avanzar en la investigación. En cuanto a este último punto, es importante destacar que el uso de datos masivos permite llegar a conclusiones más acertadas, siendo sus mayores beneficios la capacidad de identificar patrones y tendencias, lo cual es sumamente útil en campos complejos, como la economía.

De acuerdo con lo anterior, el uso de datos masivos puede aumentar la eficiencia de la investigación, lo que se alinea con la utilización de diversas herramientas y técnicas de análisis multivariante para examinar grandes cantidades de datos. Esto permite llevar a cabo procedimientos más eficientes en un tiempo más corto, lo cual constituye un elemento valioso en investigaciones que requieren el análisis de grandes cantidades de datos y puede ahorrar tiempo y recursos valiosos.

En virtud de lo anterior, la presente investigación es una confluencia de modelos estadísticos multivariantes, que se presentan generalizados en un único modelo denominado Tucker3. Este modelo se utiliza para entender y evaluar la dinámica macroeconómica en Perú entre 2007 y 2021.

En ese sentido, la dinámica macroeconómica es una rama de la economía que se centra en el estudio de los comportamientos y fluctuaciones de las grandes variables económicas a nivel agregado de la economía, como el producto, el empleo, los precios, la inversión, el gasto público, el comercio internacional, entre otros. Se enfoca en entender

cómo estas variables interactúan entre sí, cómo se afectan mutuamente (interdependencia) y cómo influyen en el crecimiento económico. Para ello, se consideran datos en un arreglo de tres vías: los departamentos del país, las variables macroeconómicas y el tiempo. Estos datos serán útiles para aplicar el modelo Tucker3 con el objetivo de analizar y explicar dicha dinámica macroeconómica.

Metodológicamente, todo parte del análisis de tres vías, también conocido como análisis de tres modos o análisis de tres factores, que es una técnica de análisis multivariante utilizada para analizar datos en los que se observan relaciones entre tres tipos diferentes de variables. De acuerdo con ello, los datos se representan en una matriz tridimensional en la que cada dimensión representa un tipo de variable, con el objetivo de identificar las estructuras subyacentes en los datos y representarlas de manera más simple y fácil de entender.

Considerando lo anterior, el modelo Tucker3 es una generalización del análisis factorial a múltiples vías, lo que permite representar la variabilidad de las diversas vías en términos de un número más pequeño de factores o dimensiones latentes.

El análisis factorial es una técnica multivariante de análisis de datos que se utiliza para representar un conjunto de variables correlacionadas en términos de un número más pequeño de factores o dimensiones latentes. El análisis factorial es una técnica bidimensional que se aplica a matrices. En ese sentido, el modelo Tucker3 es una extensión del análisis factorial a múltiples vías, lo que permite representar la variabilidad de las diversas vías en términos de un número más pequeño de factores o dimensiones latentes.

Dado que el modelo Tucker3 es una técnica de análisis factorial multidimensional que se aplica a arreglos de múltiples vías, puede considerarse un análisis factorial dinámico cuando una de las vías es el tiempo, lo que combina el análisis factorial con el análisis del sistema dinámico, es decir, aquel que transcurre o evoluciona en el tiempo.

El modelo Tucker3 es una generalización de la descomposición en valores singulares (SVD, por sus siglas en inglés). La SVD es una técnica de factorización matemática que se utiliza para representar una matriz en términos de tres matrices simples. El modelo Tucker3, en cambio, se puede utilizar para representar arreglos de cualquier dimensionalidad y, por lo tanto, es una herramienta más flexible para el análisis de datos multidimensionales.

En función de lo mencionado, el objetivo de la presente investigación es evaluar la dinámica macroeconómica utilizando el modelo Tucker3 en un análisis factorial dinámico. Se analizarán los datos macroeconómicos comprendidos entre 2007 y 2021 de los departamentos del país. Esto permitirá identificar las tendencias y patrones en el crecimiento económico regional. El alcance de la investigación se limita a los departamentos del país y a los indicadores económicos de crecimiento. No se abordarán otros aspectos macroeconómicos, como los precios, el consumo o la inversión.

La presente investigación se divide en siete capítulos. El primero considera los aspectos fundamentales en los cuales se basa la tesis, como la situación problemática, la formulación del problema, su justificación y los objetivos que busca alcanzar. El segundo capítulo es el marco teórico, en el cual se explicarán cada uno de los temas básicos y fundamentales para el desarrollo de la tesis, destacando la teoría del análisis factorial y los componentes principales, la descomposición en valores singulares, el análisis de tres vías y finalmente el modelo Tucker3 y su desarrollo.

El tercer capítulo es la descripción metodológica, que involucra el tipo de investigación, la muestra recogida, la obtención de los datos y su procesamiento, la técnica de recolección y el análisis de la información. El cuarto capítulo desarrolla el análisis de la dinámica macroeconómica del Perú entre 2007 y 2021, considerando algunos hechos estilizados. Con ello, se aplica el modelo Tucker3 en su aplicación del análisis factorial dinámico para evaluar la dinámica macroeconómica. El quinto capítulo presenta los resultados y la discusión, es decir, el análisis, interpretación y discusión de la aplicación. El sexto capítulo presenta los resultados, finalmente, el séptimo capítulo las conclusiones.

1.1. Situación problemática

El análisis multivariado tiene como objetivo el estudio de los datos obtenidos de la medición de más de una variable sobre un conjunto de individuos u objetos, los cuales pueden ser representados en una matriz de individuos por variables, denominado como un arreglo de dos vías. Para analizar conjuntos de datos que contienen un número relativamente grande de variables se aplican técnicas de reducción de dimensiones como el análisis de componentes principales (ACP), análisis factorial (AF) , análisis de correspondencia (AC) o escalamiento multidimensional (MDS); a éstas técnicas multivariadas (ACP, AF, AC y EM) en la escuela francesa se les denominada métodos factoriales y están relacionadas con modelos matemáticos y estadísticos, las cuales se han desarrollado para extraer la máxima información contenida en los datos, reducir variables, analizar las puntuaciones observadas de muchos individuos con miras a la aplicación posterior de modelos estadísticos.

Estos métodos datan desde los inicios del siglo XX con los intentos de Karl Pearson, Charles Spearman, y otros para definir y medir la inteligencia. Debido a su temprana asociación con constructos, tales como la inteligencia, el análisis factorial fue desarrollado principalmente por científicos interesados en la psicometría; los argumentos alrededor de las interpretaciones psicométricas de varios estudios iniciales y la falta de recursos computacionales impidieron su inicial desarrollo como un método estadístico. El advenimiento de las computadoras de gran velocidad ha generado un renovado interés en los aspectos teóricos y prácticos del análisis factorial. La mayoría de las técnicas originales fueron abandonadas y controversias iniciales han venido resolviéndose a raíz de recientes desarrollos. Sigue siendo cierto, sin embargo, que cada aplicación de la técnica debe ser examinada por sus propios méritos para determinar su éxito.

La propuesta esencial del análisis factorial es describir, si es posible, las relaciones de un número grande de variables, contenidas en una matriz de covarianza, en términos de un menor número de variables latentes, subyacentes, no observables, llamadas factores.

Básicamente, el modelo factorial es motivado por el hecho que las variables pueden ser agrupadas a través de sus correlaciones. Es decir, suponer que todas las variables dentro de un grupo están altamente correlacionadas, pero tienen correlaciones bajas entre

variables en grupos diferentes. Entonces, es concebible que cada grupo de variables represente un constructo subyacente, o factor, que es responsable de las correlaciones observadas. El análisis factorial dinámico no es una técnica nueva, por el contrario, se ha utilizado en campos como la econometría y la psicología desde mediados de los años ochenta; sin embargo, sus inicios datan de los años cincuenta.

Cattell (1952) sugiere el análisis de una trayectoria observada de una serie de tiempo como funciones vectoriales, es decir, mediciones repetidas en un solo sujeto en muchas ocasiones, por medio del modelo factorial habitual. Así, Cattell propuso la aplicación especial del análisis factorial tradicional con una estructura de covarianzas de series de tiempo, donde las variables observadas pueden concebirse como causadas por una serie de factores o variables latentes de menor dimensión (Molenaar, 1985; Zuur et al., 2003)

Un problema que subyace en lo mencionado por Cattell era la forma adecuada en la que se debe manejar una estructura de covarianza que involucra retardos. Inicialmente, se tenía en cuenta las relaciones simultaneas entre los componentes de una serie de tiempo multivariante, esto generaba a la perdida de las relaciones temporales; no obstante, en 1963 el mismo Cattell propone un método iterativo de desplazamiento de factor variable, lo que implicaba la búsqueda iterativa del retardo en el que cada componente de una serie de tiempo multivariante se correlaciona al máximo con un factor dado.

Así, Molenaar (1987) define al análisis factorial dinámico como una rama particular de análisis de series de tiempo que se ocupa de la función de covarianza desfasada en el tiempo que refleja las relaciones entre las series de componentes observadas que pueden evolucionar a lo largo de varias unidades de tiempo. Como lo describe Anderson (1963), la presencia de las dependencias en los rezagos impide el uso del análisis factorial estándar o tradicional. En su defecto, sugiere un enfoque generalizado enmarcado en el modelado dinámico de procesos. Por lo tanto, se llega a un modelo de factor dinámico en la que la parte común de una serie de tiempo multivariada observada se concibe como la salida de un filtro causal, cuya entrada es una serie de factores latentes de menor dimensión.

Este filtro causal al cual se refiere Anderson (1963), puede interpretarse como una

secuencia de cargas factoriales rezagadas que describen las relaciones entre las observaciones dependientes del tiempo y las fuentes latentes. Con esto, se puede considerar que el factor dinámico es un modelo causal de la función de covarianza completa, en el marco de las técnicas de reducción de dimensiones, especialmente diseñada para datos de series de tiempo (Molenaar, 1987, 1985; Zuur et al., 2003).

El análisis factorial dinámico puede proceder a lo largo de dos líneas diferentes, pero formalmente equivalentes. La primera puede llevarse a cabo en el dominio del tiempo, donde, condicional a los valores fijos de ciertos parámetros, el modelo factorial dinámico se reescribe como un sistema de ecuaciones estructurales que puede ajustarse mediante técnicas de máxima verosimilitud. Este enfoque puede utilizarse con una serie de tiempo multivariante observada de longitud relativamente pequeña, aunque requiere una especificación a priori (Molenaar, 1987, 1985).

En la segunda línea, el análisis factorial dinámico se puede llevar a cabo en el dominio de la frecuencia, lo que deriva en un conjunto de modelos estructurales de valor complejo indexados por frecuencia, este enfoque no solicita una estructura a priori, sólo requiere la disponibilidad de una serie de tiempo multivariante de longitud suficientemente larga.

En esencia, la fortaleza del análisis factorial dinámico recae sobre el hecho que este tipo de modelos pueden describir tanto la estructura como las relaciones de factores latentes con rezago temporal, lo cual los convierte en alternativas esenciales para estudiar el cambio intraindividual. Con ello, se cumple el propósito de combinar el análisis factorial y el análisis de las series de tiempo (Chow et al., 2011).

En el análisis factorial dinámico, la dimensión de los datos ya no se circunscribe a dos vías, tal como la versión clásica o tradicional, ya aplicaría una dimensión adicional, el número de casos o el tiempo. En esta nueva estructura, la necesidad de descomponer la variabilidad de cada una de las vías es fundamental para determinar los factores subyacentes a los datos, con el objetivo de representarlos de manera resumida.

La situación se puede complicar más si consideramos más de tres vías, es decir, considerar conjuntamente el número de casos y el tiempo, lo cual puede llevar a mayores complicaciones, para un mejor análisis de los datos. Evidentemente, la complejidad para determinar las fuentes de variabilidad estarán dadas por métodos iterativos que permitan separar y maximizar la variabilidad en cada una de las vías. En tal sentido,

es apropiado contar con un análisis riguroso del método que permita desarrollar este propósito.

En el caso de un arreglo de dos vías, cada elemento de la matriz de datos de I individuos en J variables, es caracterizado por un par de índices, i para el modo unidades y j para el modo variables. Cuando el dato elemental se refiere a tres o más índices (vías o criterios de clasificación), entonces se tiene datos multivías. Particularmente, si se considera 3 vías, los datos multivías serán de la forma siguiente:

$$\mathbf{X}^{(\mathbf{I}\times\mathbf{J}\times\mathbf{K})} = \mathbf{x_{ijk}}, \quad i = 1, ..., I; j = 1, .., J; k = 1, ...; K \tag{1.1}$$

donde: i denota la vía del tipo unidades, j denota la vía del tipo variables, k denota la vía del tipo ocasiones. El método de Tucker3 consiste en descomponer la matriz de tres vías $\mathbf{X}^{(\mathbf{I}\times\mathbf{J}\times\mathbf{K})}$ mediante el algoritmo Tucker3, procedimiento iterativo que se basa en la descomposición en valores singulares generalizados de diferentes arreglos matriciales hasta alcanzar la convergencia de las mismas.

1.2. Formulación del problema

La situación problemática anterior expresa una necesidad teórica y empírica que permita obtener resultados aplicables. El uso masivo de datos multivariantes que involucra diversas variables o características, y las relaciones que pueden derivarse de dichas dependencias, permite analizar las relaciones subyacentes de los datos o reducir la dimensionalidad sin pérdida de información. Sin embargo, esta característica multivariante que implica muchas dimensiones conduce a modelos complejos y difíciles de comprender.

En general, el modelo Tucker3 es una herramienta que constituye una generalización de varios métodos estadísticos. Por un lado, es una generalización del análisis factorial y el análisis de componentes principales para el análisis de tres vías, lo que permite determinar la estructura subyacente de los datos y, bajo ciertos supuestos, maximizar la variabilidad con la menor pérdida de información. Por otro lado, es una generalización de la descomposición en valores singulares, ya que la matriz de tres vías se describe mediante el producto de tres matrices (una matriz central o núcleo, una matriz de direcciones derecha y una matriz de direcciones izquierda) que contienen la información esencial de la matriz multidimensional.

El modelo Tucker3 se utiliza en una variedad de aplicaciones, como el análisis de datos de imágenes, el procesamiento de señales, la representación de datos y la compresión de datos. En ese sentido, el modelo de Tucker3 es especialmente útil para analizar y visualizar datos que tienen una estructura subyacente compleja, ya que permite una representación compacta de la información.

En la actualidad, el uso de los datos masivos ha experimentado un gran impulso debido al desarrollo de las redes sociales, la mensajería, las comunicaciones, el *Streaming* de música, vídeos y películas, así como el desarrollo de aplicaciones que han simplificado la vida humana. Esta evolución ha llevado a la generación de enormes cantidades de datos como consecuencia de la interacción humana en el mundo digital. En este contexto, la inteligencia artificial y la minería de datos han experimentado un auge complejo, lo que ha dado lugar a herramientas multivariantes que permiten extraer la máxima información de los datos multivariantes obtenidos en este mundo digital.

La presente investigación busca utilizar datos multivariantes de complejas dimensiones en una aplicación en la economía peruana, con el fin de determinar la dinámica de las variables macroeconómicas departamentales del Perú en el período comprendido entre 2007 y 2021. En esta complejidad de los datos, el modelo Tucker3 permitirá llevar a cabo un análisis complejo para determinar las relaciones subyacentes de los datos, con el objetivo de identificar características no observables y comprender la dinámica de las variables económicas de los departamentos del país.

La estructura de los datos permite que la aplicación del modelo Tucker3 sea un análisis factorial, ya que en este se determinan las relaciones de las variables y los individuos en el tiempo. Sin embargo, el tiempo tiene una importancia crucial, ya que constituye una dimensión dentro de la estructura de los datos a analizar. Por lo expuesto, en los siguientes apartados se presentarán la justificación y los objetivos que persigue la presente investigación.

1.3. Justificación teórica

El modelo Tucker3 es una poderosa herramienta que se utiliza para describir, analizar y visualizar datos multidimensionales a través de una matriz multidimensional. En ese sentido, ofrece amplios beneficios, ya que permite representar de manera compacta los datos de diferentes vías mediante una descomposición generalizada en valores singulares en una matriz central y matrices de direcciones. Esto, a su vez, facilita el análisis y la visualización de la estructura subyacente de los datos multidimensionales. Finalmente, es importante porque permite reducir significativamente la dimensionalidad de los datos con la mínima perdida información.

Dada la variedad de beneficios que ofrece el modelo Tucker3, se utiliza con frecuencia en el campo de la inteligencia artificial y la minería de datos para analizar y visualizar datos multidimensionales complejos, ya que posibilita una representación más compacta de la información.

De manera más específica, el modelo Tucker3 permite analizar y visualizar datos de imágenes multidimensionales, como imágenes médicas o satelitales. Además, se utiliza en el procesamiento de señales para describir y analizar señales multidimensionales, como señales de audio o vídeo, lo que permite extraer características importantes de la señal, como patrones de sonido o movimiento.

Otras aplicaciones se encuentran en la representación de datos, especialmente en aplicaciones que involucran grandes cantidades de datos, donde es necesario reducir su dimensionalidad para facilitar el análisis. Del mismo modo, se utiliza en la compresión de datos, ya que al representar la información mediante una matriz central y matrices de direcciones, es posible reducir significativamente la dimensionalidad de los datos sin perder información importante.

Como se mencionó anteriormente, también se utiliza en problemas de procesamiento de lenguaje natural, como el análisis semántico de texto, donde se busca capturar la relación entre las palabras y frases en un texto y su significado. Del mismo modo, en el campo de la recomendación, se utiliza para analizar las interacciones de los usuarios con un sistema de recomendación y generar recomendaciones personalizadas.

En el ámbito económico, el modelo Tucker3 es útil debido a su capacidad revelar los componentes tendenciales de lo arreglos de tres vías. En ese sentido, se pueden destacar

los siguientes usos del modelo de Tucker3:

- Análisis de mercados financieros: Se utiliza para analizar patrones en los datos de la bolsa de valores y generar predicciones sobre el comportamiento del mercado. Puede utilizarse para analizar la relación entre diferentes activos financieros y detectar patrones de correlación.

- Análisis de datos de precios: Se utiliza para analizar y describir la estructura subyacente de los precios en diferentes mercados. Esto puede incluir el análisis de precios de materias primas, productos manufacturados o servicios.

- Análisis de datos de producción: Se utiliza para analizar y describir la estructura subyacente de la producción en diferentes industrias. Esto puede incluir el análisis de la producción de bienes y servicios, y la relación entre la producción y el empleo.

- Análisis de datos de consumo: Se utiliza para analizar y describir la estructura subyacente del consumo en una economía. Esto puede incluir el análisis de las tendencias en el consumo, la relación entre el consumo y el ingreso, y el impacto de las políticas económicas sobre el consumo.

- Análisis de datos de inversión: Se utiliza para analizar y describir la estructura subyacente de la inversión en una economía. Esto puede incluir el análisis de las tendencias en la inversión, la relación entre la inversión y el crecimiento económico, y el impacto de las políticas económicas sobre la inversión

En resumen, el modelo de Tucker3 se utiliza en economía para analizar y describir la estructura subyacente de los datos económicos, como los precios, la producción, el consumo, e inversión, y generar predicciones y análisis sobre el comportamiento del mercado y los impactos de las políticas económicas.

1.4. Justificación práctica

Como se ha mostrado en la sección anterior, el modelo Tucker3 permite la identificación de estructuras de componentes latentes de mayor orden en los datos, lo que puede ser útil para la comprensión de las relaciones complejas entre las variables. En específico, cuando la estructura de componentes latentes está determinada por un conjunto de datos de series de tiempo, el modelo Tucker3 es capaz de manejar datos con una estructura temporal y puede ser utilizado para identificar patrones y tendencias en los datos a lo largo del tiempo.

De acuerdo con lo anterior, la investigación propuesta tiene una justificación práctica importante en diversos campos. En particular, en la economía, sus usos son varios, entre los cuales podemos distinguir la búsqueda de patrones para la predicción y el análisis subyacente de los datos económicos. Los resultados de esta investigación se utilizarán para desarrollar un modelo que permita entender la estructura que subyace a los datos productivos de la economía peruana entre 2007 y 2021, lo cual será sumamente útil para realizar predicciones y encontrar tendencias entre los diversos sectores económicos en los departamentos del país.

Si bien esta aplicación ha sido poco frecuente, la capacidad del modelo Tucker3 en el análisis multidimensional permitiría determinar los patrones de las tres vías propuestas: departamentos (individuos), nivel de producción y otras (variables), y el tiempo (ocasiones). En ese sentido, el modelo permitirá tener un conocimiento más amplio en relación a la dinámica económica de los departamentos del país y, por agregación, de toda la economía nacional.

Además, el presente trabajo también se utilizará para evaluar la efectividad del modelo Tucker3 en datos temporales, con miras a realizar predicciones sobre la dinámica productiva o, por lo menos, respecto a las relaciones subyacentes que determinan dichos patrones. Por lo tanto, en esta investigación se buscará ajustar el modelo Tucker3 a los datos económicos del Perú entre 2007 y 2021, con el fin de entender mejor las relaciones complejas entre las variables económicas departamentales durante ese transcurso de tiempo.

1.5. Beneficios del uso del modelo Tucker3

En economía, el modelo Tucker3 y sus variantes se utilizan en una variedad de aplicaciones, como la predicción de series de tiempo económicas, la detección de patrones en datos financieros y la estimación de modelos económicos multivariantes. También se utilizan en la evaluación de riesgos financieros, la identificación de factores económicos subyacentes y la segmentación de mercados financieros.

Además, el modelo Tucker3 y otras representaciones del mismo pueden ser útiles en la investigación de relaciones complejas entre variables económicas y en la estimación de modelos econométricos. Esto se debe a que estos modelos son capaces de capturar interacciones no lineales y patrones complejos en los datos económicos, lo que los hace valiosos para la investigación económica.

Aplicar el modelo Tucker3 para entender la dinámica económica de un país requiere un enfoque cuidadoso y detallado para recopilar y analizar los datos relevantes. En primer lugar, se deben identificar las variables económicas que se desea incluir en el análisis, como el PBI, la inflación, el desempleo, la tasa de interés, etc. Es importante asegurarse de tener una cantidad suficiente de datos históricos para cada variable para poder realizar un análisis riguroso.

Una vez que los datos están disponibles, se pueden aplicar técnicas estadísticas para normalizarlos y eliminar cualquier sesgo o tendencia espurea. Luego, se puede utilizar el modelo Tucker3 para analizar las interacciones entre las variables económicas. El modelo Tucker3 utiliza un enfoque matemático para descomponer la información en tres componentes: los modos de ajuste, las coordenadas y los valores singulares. Estos componentes se utilizan para identificar las interacciones más fuertes entre las variables económicas y para determinar la dirección y la magnitud de estas interacciones.

El modelo Tucker3 permite analizar la evolución del PBI de manera más detallada y profunda, de manera que es útil para identificar la contribución de factores externos e internos al crecimiento económico y para analizar cómo estos factores interactúan entre sí para influir en el PBI a lo largo del tiempo. De esta manera, el modelo Tucker3 ayuda a entender la dinámica económica de un país, ya que permite identificar las tendencias y las causas subyacentes que afectan el crecimiento económico.

Los resultados obtenido indican la existencia de relaciones estrechas entre determinadas variables económicas, lo que podría tener implicaciones importantes para la política económica. También pueden proporcionar una comprensión más profunda de los factores que impulsan o limitan el crecimiento económico en el país.

1.6. Objetivos

La tesis se enfoca en aplicar el modelo Tucker3 a un conjunto de datos económicos departamentales del Perú durante un período de tiempo entre 2007 y 2021; y con los resultados investigar las relaciones complejas entre los indicadores económicos que describen la dinámica económica en el país. El objetivo está en línea con comprender como estas relaciones pueden ayudar a mejorar las decisiones económicas y a ayudar a las empresas y los inversores a tomar decisiones informadas.

1.6.1. Objetivo general

El objetivo central de la presente investigación es utilizar el método Tucker3, en una estructura de tres vías mediante el análisis factorial, para evaluar la dinámica macroeconómica del Perú entre 2007 y 2021.

1.6.2. Objetivos específicos

Los objetivos específicos de la presente investigación son:

- Describir los principios teóricos y los soportes analíticos del método Tucker3 en el análisis factorial dinámico.
- Determinar la dimensionalidad óptima que permita el mejor ajuste del modelo Tucker3.
- Estimar los componentes de las tres vías para el análisis de la dinámica macroeconómica del Perú entre 2007 y 2021.
- Representar de la evolución de los departamentos y variables en espacio estimado mediante el método Tucker3.

Capítulo 2

Marco teórico

La presente tesis presenta un conjunto de técnicas estadísticas que se usan conjuntamente de manera que, la relación existente entre ellas es tan estrecha que no se puede dilucidar las fronteras entre ellas. El método *Tucker3* es una generalización del análisis de tres vías, que se basa en un conjunto de técnicas para la descomposición de las matrices, tales como el análisis factorial y la descomposición en valores singulares. Con ello, el arreglo de tres vías es posible ser descompuesto en un conjunto de matrices que permiten determinar subyacentes de las vías.

El análisis factorial dinámico, por su parte, es una generalización del análisis factorial clásico, donde el tiempo se añade como una vía de condición en un arreglo de dos vías. De esta manera, la filosofía del análisis factorial está presente, pues se expresa los valores reales en función a un conjunto de nuevas variables latentes o no observables. Todo ello en base a un modelo estadístico que considera supuestos de normalidad en los factores y los residuos.

Cambiar la perspectiva de un análisis de dos vías a uno de múltiple vías conlleva una ilación secuencial que, en la medida de lo posible, se pretende abordar. Para ello, en el presente capítulo se abordan los diversos modelos y métodos que permitan aclarar el análisis de tres vías que, en esencia, es el corazón sobre el cual gira todo el marco teórico. En ese sentido, se inicia con el análisis de componentes principales, el análisis factorial, la descomposición en valores singulares hasta llegar al análisis de tres vías, donde se especifican los supuestos, la estructuración y modelos existentes; siendo el modelo *Tucker3*, el de mayor preocupación.

2.1. Antecedentes

El modelo Tucker3 es una técnica estadística diseñada para el análisis de datos de múltiples dimensiones. Puede considerarse una generalización del Análisis de Componentes Principales (PCA) y la Descomposición en Valores Singulares (SVD), ampliando así el espectro del análisis multidireccional Andersson and Bro (1998). Tucker formuló dos modelos, uno de componentes principales y otro de factor común, junto con varios procedimientos computacionales, lo que abrió las puertas al análisis de datos tridimensionales. Otros autores, como Levin en 1965 Levin (1965) y Appellof en 1981 Appellof and Davidson (1981), avanzaron en las aplicaciones del modelo Tucker, independientemente de los trabajos iniciales de Tucker. El problema central del análisis factorial tridimensional o de tres vías radicaba en el análisis simultáneo de todas las dimensiones y las relaciones entre los factores.

En particular, el modelo Tucker3 puede considerar tres dimensiones: (i) los individuos, (ii) las variables y (iii) las condiciones. Cuando las condiciones se refieren al tiempo, se convierte en un análisis dinámico. En este sentido, el modelo factorial se presenta como un análisis factorial dinámico, que puede verse tanto como un modelo de regresión que relaciona variables observadas con variables latentes (Rella, 2020), como una adición independiente de una dimensión a la estructura de datos bidimensionales, creando así un arreglo tridimensional que representa tres conjuntos de entidades: individuos, variables y ocasiones (Amaya and Pacheco, 2002).

El modelo Tucker3 fue introducido por primera vez en las ciencias sociales por Tucker en 1963 en su artículo "Implications of factor analysis of three-way matrices for measurement of change". Posteriormente, en los años 1964 y 1966, se refinó con mejoras en la descripción matemática y propuestas de estimación de los parámetros (Tucker, 1963; Tucker et al., 1964; Tucker, 1966). Desde entonces, el modelo Tucker3 ha ganado aceptación.

En la década de los setenta, el análisis tridimensional se expandió hacia el escalamiento multidimensional, principalmente gracias a Carroll and Chang (1970), quien propuso una generalización del análisis de Eckart-Young para descomponer arreglos de tres dimensiones. Gandour y Diehl en 1978 Gandour and Harshman (1978) utilizaron el

modelo PARAFAC para analizar el escalamiento multidimensional en datos de percepción del sonido. Este análisis buscaba determinar las dimensiones subyacentes a la percepción del tono lingüístico y cómo los antecedentes lingüísticos de un individuo influyen en su percepción. Los autores concluyeron que la mayoría de los hablantes de lenguas tonales, como el tailandés y el yoruba, tienen patrones perceptivos diferentes a los hablantes de lenguas no tonales, como el inglés.

Además, sus conclusiones sugieren que los modelos de escala multidimensional de múltiples dimensiones permiten un método eficiente para explorar espacios perceptivos en los que el número y tipo de dimensiones no se conocen previamente, además de determinar cómo estas dimensiones pueden estar relacionadas con características especiales, como las lingüísticas, culturales, educativas, etc.

A finales de la década de los setenta, destacan los estudios realizados por Bentler and Lee (1979), quienes propusieron una versión vectorial aleatoria del análisis factorial del modelo de componentes de tres vías de Tucker. Los autores emplearon el método de mínimos cuadrados generalizados para estimar los parámetros del modelo y un algoritmo de Gauss-Newton para el ajuste empírico. Se calcularon los errores estándar de los estimadores y se presentó una estadística de bondad de ajuste. Se concluyó que, con el aumento de la disponibilidad de enfoques para el análisis de tres vías, el desarrollo de estudios aumentaría considerablemente.

En una línea similar, Harshamn y Lundy llevaron a cabo investigaciones sobre el análisis factorial de tres vías y el escalamiento multidimensional (Harshman, 1984; Harshman and Lundy, 1984). Los autores compararon diversos modelos para el análisis de tres vías, centrándose en el método PARAFAC (parallel factor analysis) como una particularidad del modelo Tucker3.

Posteriormente, en los años ochenta, comenzó un proceso de desarrollo de la literatura sobre el análisis de tres vías y, en particular, sobre el modelo Tucker3. Esto se evidenció en las versiones estocásticas de Bloxom (1968), Bentler and Lee (1978) y Lee and Fong (1983). Más adelante, en 1983, Kroonenberg (1983) presentó un método de estimación mejorada de mínimos cuadrados alternos para el modelo original de Tucker3, así como una programación para llevar a cabo el análisis. Esta contribución formó parte del libro "Three-mode principal component analysis: Theory and applications", que describía el estado del arte del modelo de componentes de Tucker. Kiers (2000)

estandarizó la notación de los modelos Tucker2 y Tucker3. Muchas mejoras posteriores fueron abordadas por Coppi and Bolasco (1989), Geladi (1989) y Kroonenberg.
En los años noventa y posteriores, se produjo una expansión tanto en el desarrollo como en la citación de los trabajos sobre análisis de múltiples vías, el modelo Tucker3 y sus particularidades. Fue precisamente en este período cuando estos métodos se utilizaron de manera gradual en otras disciplinas, en particular en la agricultura (como en el trabajo de Basford et al. (1991)), la ecología (Beffy (1992)), la espectroscopia (Leurgans and Ross (1992)) y la química (Smilde (1992)).

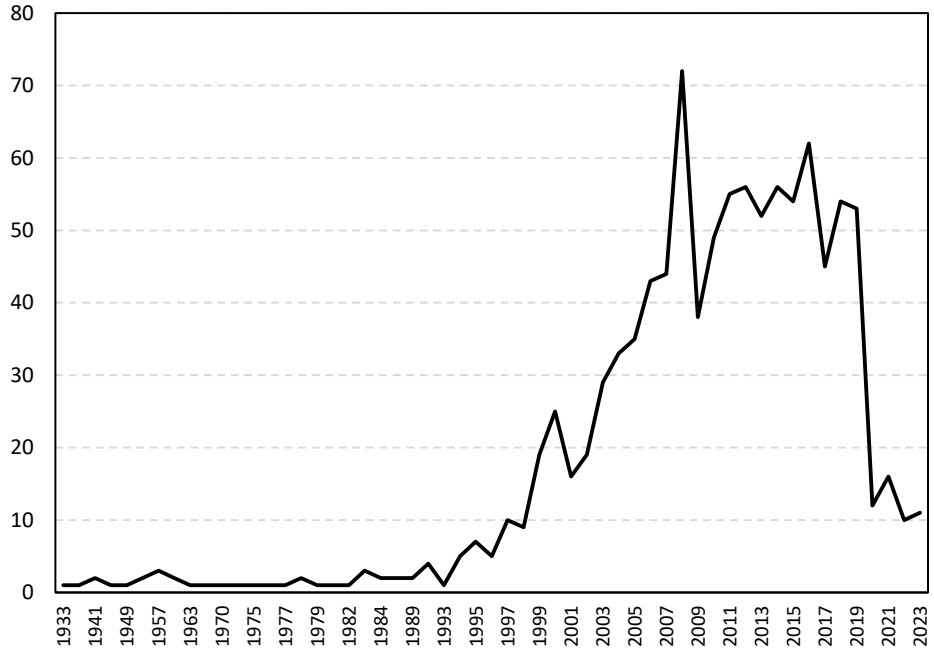


Figura 2.1: Número de publicaciones sobre Tucker3, 1933-2021

Nota. Elaborado en base a datos de Publish or Perish (PoP) 8.17.

La figura 2.1 representa el número de publicaciones relacionadas con el modelo Tucker3. En ella, se observa un marcado aumento a partir de 1992, alcanzando su punto más alto en 2008. Podemos considerar que el auge en el desarrollo de la literatura relacionada con el modelo Tucker3 se produjo principalmente en las décadas de los noventa y la primera del siglo XXI. Este fenómeno se explica en gran medida por el incremento en la recopilación de datos gracias a nuevos instrumentos de recolección, así como al avance de las tecnologías de la información. En este contexto, Smilde (1992) señaló que el crecimiento de los modelos de múltiples vías se debía a su capacidad

para variar todos los factores simultáneamente con el fin de optimizar una respuesta, dependiendo de los factores involucrados.

Dado que el modelo Tucker3 se encuentra en el ámbito de los modelos de tres o múltiples vías, Kroonenberg (2016) presenta la evolución de las citas en este campo en diversas disciplinas desde 1965 hasta 2016. Estas cifras se ilustran en la figura 2.2, donde se refleja una tendencia similar a la figura 2.1, con un aumento significativo en las citas a partir de la década de los noventa. Además, se proporciona información sobre las citas en áreas como la psicología, química, detección de señales y matemáticas.

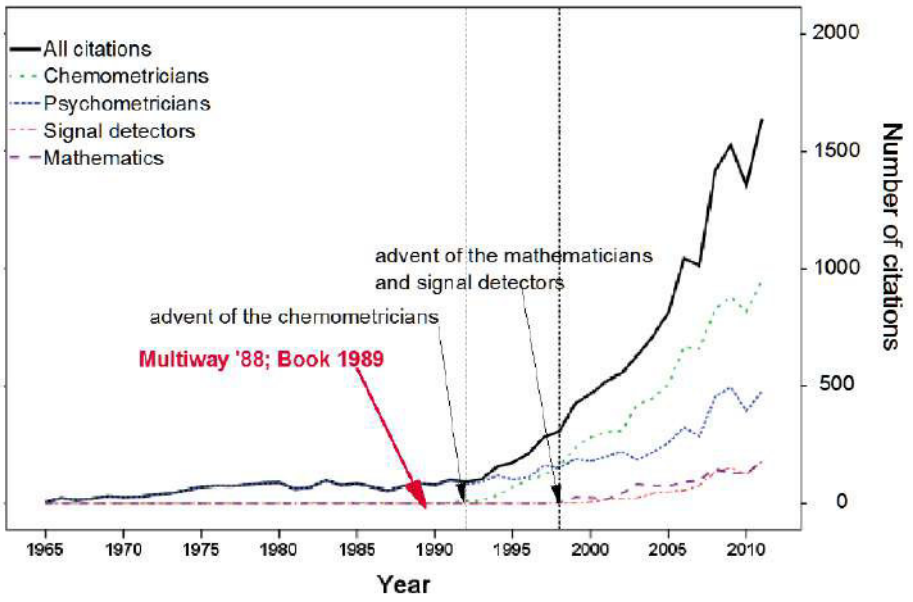


Figura 2.2: Citas de artículos de tres y múltiples vías de investigadores de diferentes disciplinas.

Nota. Tomado de (Kroonenberg, 2016, 6).

La figura 2.2 refleja un mayor desarrollo en los campos de la química y la psicología. En contraposición, las matemáticas no evidencian un crecimiento significativo. Sin embargo, como señaló Kroonenberg (2016), es interesante notar que matemáticos teóricos como Hitchcock ya habían concebido modelos de tres vías en 1927, aunque su uso no se popularizó hasta aproximadamente 1998, cuando los modelos de múltiples vías atrajeron su atención. De hecho, como respalda Kroonenberg (2016), los artículos de los pioneros en modelos de tres vías, como Tucker y Harshman, son citados principalmente en trabajos de disciplinas como las ciencias naturales.

Las ciencias sociales y del comportamiento parecen estar más inclinadas hacia la aplicación de nuevas teorías y pruebas estadística inferencial en lugar de enfoques descriptivos. También es posible que muchos científicos sociales se centren en identificar las causas de los fenómenos y predecir el comportamiento futuro, algo para lo que los métodos de tres vías no fueron inicialmente diseñados, aunque se han dado los primeros pasos en esa dirección.

Otro factor que podría influir es el deseo legítimo de mantener los análisis lo más simples posible para evitar discrepancias con los editores de revistas y sus revisores. Además, la falta de conocimiento sobre métodos estadísticos avanzados multivariantes, en particular el análisis de tres o múltiples vías destinado a modelar diferencias individuales y descubrir patrones, también podría ser relevante. Además, la ausencia de teorías sólidas que respalden los modelos estadísticos para el análisis de tres vías podría limitar su uso, incluyendo el modelo Tucker3.

Dentro de la bibliografía relacionada con los modelos Tucker3, uno de los documentos más citados es el artículo de Bro (1997), donde se presenta el método de descomposición multidireccional PARAFAC y su aplicación en quimiometría. PARAFAC es una extensión del Análisis de Componentes Principales (PCA) para matrices de mayor orden. A diferencia del PCA, PARAFAC no presenta problemas de rotación y puede recuperar espectros puros de datos espectrales de múltiples vías. No obstante, a diferencia del PCA, los componentes no pueden estimarse de manera sucesiva, ya que esto resultaría en un modelo con un ajuste deficiente. PARAFAC guarda una gran similitud con el modelo Tucker3; de hecho, puede considerarse como un modelo Tucker3 con restricciones (un PCA en dos vías). En este sentido, ambos constituyen modelos de descomposición multilíneal que descomponen la matriz en un conjunto de puntuaciones y cargas.

En la misma línea, Bro and Kiers (2003) proponen un nuevo enfoque de diagnóstico denominado diagnóstico de consistencia central o CORCONDIA para determinar la cantidad adecuada de componentes en modelos multidireccionales. Este enfoque se aplica principalmente al modelo de análisis factorial paralelo (PARAFAC) y también a otros modelos que pueden considerarse como variantes restringidas del modelo Tucker3. CORCONDIA se basa en evaluar la idoneidad del modelo estructural mediante

el análisis de datos y los parámetros estimados de modelos gradualmente aumentados. En el caso del modelo PARAFAC, que utiliza combinaciones dimensionales de componentes en todas las vías, se considera apropiado cuando la inclusión de otras combinaciones de los mismos componentes no mejora significativamente el ajuste del modelo. La recomendación es seleccionar el modelo más grande que aún sea adecuado en términos de ajuste.

De la misma manera, los principales artículos en relación a los modelos Tucker3 señalan que Andersson and Bro (1998) desarrollan un algoritmo con el propósito de mejorar la velocidad de convergencia de los modelos multidireccionales. En tal sentido, su estudio realiza 9 implementaciones diferentes del algoritmo Tucker3 con un conjunto de datos medidos a partir de mediciones espectrofluorimétricas en 65 muestras, y tiene dimensiones de $65 \times 40 \times 311$, que representan aproximadamente 6,5 MB de tamaño. Los autores encuentran una estimación de dimensión de $8 \times 8 \times 8$ con un nivel de convergencia de $3{,}720345874925 \times 10^{10}$.

El algoritmo realizado por Andersson and Bro (1998) contrasta con el realizado por Kroonenberg (1983), en el que se sostiene que no vale la pena resolver vectores propios altamente precisos, ya que el algoritmo resultante obtendrá una estructura de iteración dentro de iteración que requiere demasiado esfuerzo computacional. Después de todo, durante las iteraciones, los problemas propios planteados solo se formulan en términos de factores intermedios. Además, encuentran que se requiere precisión, quizás más que velocidad, en las implementaciones de modelos Tucker3 de grandes matrices de datos para producir resultados en el menor tiempo posible.

Posteriormente, Andersson and Henrion (1999) desarrollaron un algoritmo general para el modelo PCA múltiple y Tucker3. Su objetivo era maximizar una función objetivo diferenciable que dependía de un conjunto de matrices ortogonales, para luego abordar el problema de estimar matrices de transformación ortonormales que permitieran rotar las matrices centrales hacia una estructura más simple. Por su parte, Andersson and Bro (2000) crearon un "Toolbox" en MATLAB, un algoritmo que modela datos de múltiples vías y que abarca diversos modelos, incluyendo aquellos de la familia Tucker. Estos modelos han demostrado ser altamente aplicables en el campo de quimiometría.

En su artículo, Henrion (1998) se propuso simplificar la matriz central de cargas del modelo Tucker3, conocida como núcleo, la cual resume las cargas de las tres vías. Para

lograrlo, implementó una estrategia sencilla en la que se seleccionó un subconjunto de vías para la transformación, lo que generó una estructura más simple en las matrices. A este enfoque de optimización condicional de la estructura específica del núcleo se le denominó método SIMPLIMAX, que, de manera similar a los métodos de diagonalización, estableció una estructura ideal para el núcleo (con un número fijo de ceros que debía aproximarse lo máximo posible).

Uno de los resultados clave de esta aplicación fue la posibilidad de establecer una simplicidad en los componentes de todas las vías. Sin embargo, se observó que el núcleo dejaba de ser completamente simple. Esto se debía a que, aunque se lograba simplificar en componentes para variables, las matrices de componentes para objetos y condiciones seguían siendo complejas. De los tres núcleos optimizados condicionalmente, el que presentaba componentes simplificados para las variables lograba un núcleo con una estructura casi tan simple como en el caso óptimo. Sin embargo, la matriz de componentes para objetos y condiciones aún distaba de ser simple. En última instancia, se consideró que el núcleo con las condiciones simplificadas representaba un buen compromiso en este ejemplo.

Estas conclusiones reflejan una evaluación razonable teniendo en cuenta las dimensiones del arreglo. No obstante, el éxito de la primera aplicación motivó a Henrion (1998) a expandir el ejercicio a un conjunto de datos más amplio de tres vías, bajo la suposición de que la relativa simplicidad previamente observada podría mantenerse. En este sentido, se incluyeron datos de 102 laboratorios con cuatro muestras de sedimentos distintas, resultando en un arreglo con $I = 88$ individuos, $J = 15$ variables (Hg, Cd, Pb, Cu, Ni, Cr, Zn, As, Fe, Mn, Al , Ca, Mg, Tl, Sn) y $K = 4$ condiciones (muestras de sedimento). Para simplificar, se optó por una solución con dos componentes en cada vía.

Los resultados revelaron que la principal fuente de variación en el error del estudio era uniforme con respecto a los cuatro tipos de sedimentos, indicando que los errores eran estables y reproducibles en diferentes muestras. Es relevante destacar que, a diferencia de lo observado en algunos estudios similares, no se evidenció un error significativo atribuible a la inestabilidad del trabajo, es decir, no se encontraron sobreestimaciones significativas en una muestra seguidas de subestimaciones extremas en otra. Por el contrario, el error bruto específico se mantuvo estable en diferentes sedimentos. Esta

breve discusión ilustra cómo es posible obtener una visión general rápida y gráfica de las principales fuentes de error en un estudio entre laboratorios de tres vías.

En el estudio llevado a cabo por Estienne et al. (2001), se realizó un análisis con seis vías utilizando el modelo Tucker3 y una matriz de gran dimensionalidad. Los datos se obtuvieron a partir de un electroencefalograma durante la fase de prueba de un nuevo fármaco destinado a afectar la actividad cerebral. Sus conclusiones destacaron la eficacia del modelo Tucker3 en la extracción de información de conjuntos de datos complejos, especialmente cuando el PCA clásico no resultó en información útil.

En otro ámbito, Pravdova et al. (2001) llevaron a cabo mejoras en el algoritmo de mínimos cuadrados alternos (ALS), que tiende a ser altamente sensible a valores atípicos en los datos. Los valores atípicos de una muestra periférica pueden ejercer una influencia significativa en el modelo resultante. Para abordar este desafío, se exploraron diagnósticos estadísticos y se propuso la construcción de un algoritmo robusto. En el segundo enfoque, los procedimientos de modelado se centraron en la creación de modelos que representaran la mayoría de los datos sin ser afectados por los valores atípicos. Estos procedimientos robustos se caracterizaron por el concepto de "punto de ruptura", que se define como el porcentaje de valores atípicos en los datos que aún permite que el modelo produzca estimaciones adecuadas. Como resultado, un subconjunto de datos sin valores atípicos se denominó "subconjunto limpio".

Para evaluar estos métodos, Pravdova et al. (2001) utilizaron un conjunto de datos en forma de arreglo de dimensiones $50 \times 10 \times 10$, generado mediante una simulación de Monte Carlo con dos factores en cada vía. Se construyeron dos modelos Tucker3 (X1 y X2) para explicar el 60 % y el 90 % de la varianza de los datos, respectivamente. Además, se contaminaron los conjuntos de datos iniciales con diferentes porcentajes de valores atípicos, 20 % y 40 %, respectivamente.

A partir de estas condiciones, realizaron 100 iteraciones con diversos porcentajes de datos atípicos, lo que condujo a los siguientes resultados: el algoritmo robusto de Tucker3 siempre convergió hacia una solución adecuada, sin que los valores atípicos influyeran en el resultado final. Por otro lado, el algoritmo Tucker3 demostró ser muy susceptible a los valores atípicos y, dependiendo de la inicialización de las cargas, podía converger hacia diferentes soluciones. La versión robusta propuesta del modelo Tucker3 funcionó

correctamente en conjuntos de datos que contenían no más del 20 % de datos discordantes. En última instancia, los modelos robustos construidos para los conjuntos de datos X1 y X2 con un 20 % de valores atípicos, es decir, conjuntos de datos con un porcentaje diferente de varianza explicada, no se vieron afectados por los valores atípicos.

De manera similar, Henrion (1998) aplicó el modelo Tucker3 a una matriz de datos tridimensional de menor tamaño. El conjunto de datos comprendía mediciones de 16 muestras (individuos) en cinco concentraciones de metales pesados (Cd, Hg, Pb, Cu) (variables) en tres órganos (riñón, hígado, músculo) (condiciones), lo que resultó en un arreglo de dimensiones $16 \times 5 \times 3$. El modelo Tucker3 fue descrito como una generalización de componentes principales, análisis factorial y mínimos cuadrados.

Comon et al. (2009) examinan resultados teóricos y proponen algoritmos numéricos. Estos algoritmos son posteriormente comparados en situaciones previamente clasificadas por Richard Harshman. Entre las conclusiones que se extraen de su estudio, los autores destacan que la reducción de dimensiones mediante el truncamiento de HOSVD (SVD de orden superior) siempre conlleva mejoras tanto en la velocidad de convergencia como en la complejidad computacional, siempre y cuando las nuevas dimensiones no sean menores que el rango del tensor que se pretende descomponer.

Dentro de la amplia lista de autores presentados en el estudio, se destaca la figura de Pieter Kroonenberg. Este autor sobresale no solo por su prolífica producción académica, sino también por su continua influencia en el análisis de múltiples vías y los modelos pertenecientes a la familia Tucker. Algunas de sus principales publicaciones incluyen:

- Principal component analysis of three-mode data by means of alternating least squares algorithms, 1980, Psychometrika.
- Annotated bibliography of three-mode factor analysis, 1983, British Journal of Mathematical and Statistical Psychology.
- Three-Mode Principal Components Analysis of Semantic Differential Data: The Case of a Triple Personality, 1985, Applied Psychological Measurement.
- Comment on Kannemann's The Exact Evaluation of 2-way Cross-classifications, 1985, Biometrical Journal.

- Three-Mode Principal Component Analysis of Multivariate Longitudinal Organizational Data, 1985, Sociological Methods and amp Research.
- Exploring children's behavior in the strange situation, 1987, Advances in Psychology
- Comments on "a generalization of fisher's exact test in PXQ contingency tables using more concordant relations" Communications In Statistics, Volume B 14, 633-645, 1987, Communications in Statistics - Simulation and Computation Cross-Validation of the WISC-R Factorial Structure Using Three-mode Principal
- Components Analysis and Perfect Congruence Analysis, 1987, Applied Psychological Measurement
- An investigation of multi-attribute genotype response across environments using three-mode principal component analysis, 1989, Euphytica.
- The analysis of multiple tables in factorial ecology. III. Three- mode principal component analysis: Analyse triadique complete, 1989, Acta Ecologica/Ecologia Generalis.
- Individual Differences in Assimilation Resistance and Affective Responses in Problem Solving, 1989, Multivariate Behavioral Research.
- Three-mode principal component analysis and perfect congruence analysis for sets of covariance matrices, 1989, British Journal of Mathematical and Statistical Psychology.
- Comment: PARAFAC in three-way land, 1992, Statistical Science.
- The TUCKALS line. A suite of programs for three-way data analysis, 1994, Computational Statistics and Data Analysis
- Classifying infants in the Strange Situation with three-way mixture method of clustering, 1995, The British journal of psychology.
- Three-way cluster and component analysis of maize variety trials, 1995, Euphytica.

- Classifying infants in the Strange Situation with three-way mixture method of clustering, 1995, British Journal of Psychology.
- Combined Analysis of Categorical and Numerical Descriptors of Australian Groundnut Accessions Using Nonlinear Principal Component Analysis, 1997, Journal of Agricultural, Biological, and Environmental Statistics.
- Rules in context: A three-mode principal component analysis of Mann et al.'s data on cross-cultural differences in respect for others, 1997, Journal of Cross-Cultural Psychology.
- Dynamics of behaviour in the Strange Situation: A structural equation approach, 1997, British Journal of Psychology.
- Nonsymmetric correspondence analysis: A tool for analysing contingency tables with a dependence structure, 1999, Multivariate Behavioral Research.
- Added value of three-way methods for the analysis of mortality trends illustrated with worldwide female cancer mortality (1968-1985), 2002, Statistical Methods in Medical Research.
- Three-mode analysis of multimode covariance matrices, 2003, British Journal of Mathematical and Statistical Psychology.
- Consensus Molecular Alignment Based on Generalized Procrustes Analysis, 2003, Journal of Chemical Information and Computer Sciences.
- Grouping three-mode data with mixture methods: The case of the diseased blue crabs, 2004, Journal of Chemometrics.
- Model selection procedures in three-mode component models, 2005, Studies in Classification, Data Analysis, and Knowledge Organization.
- Motor function in Parkinson's disease and supranuclear palsy: Simultaneous factor analysis of a clinical scale in several populations, 2006, BMC Medical Research Methodology.
- Applied Multiway Data Analysis, 2007, Applied Multiway Data Analysis.

- The equivalence of Tucker3 and Parafac models with two components, 2011, Chemometrics and Intelligent Laboratory Systems.
- Combination rules for multiple imputation in three-way analysis illustrated with chromatography data, 2012, Current Analytical Chemistry.
- Assessing factorial invariance of two-way rating designs using three-way methods, 2014, Frontiers in Psychology.
- My multiway analysis: From Jan de Leeuw to TWPack and back, 2016, Journal of Statistical Software.
- The Tale of Cochran's Rule: My Contingency Table has so Many Expected Values Smaller than 5, What Am I to Do?, 2018, American Statistician.
- CATTELL'S parallel proportional profiles: The triumph of a prodigal rotation, 2021, Journal of Chemometrics.

En los últimos años, se ha observado un creciente interés en los estudios que abordan el análisis de tres vías y los modelos de la familia Tucker en el ámbito de la salud y las matemáticas. Este interés se refleja en investigaciones realizadas por diversos autores, como se evidencia en los siguientes ejemplos:

En un estudio llevado a cabo por Koleini et al. (2023), en que se realizaron comparaciones entre el análisis de varianza (ANOVA), el análisis de componentes simultáneos (ASCA+), y las descomposiciones tensoriales de Tucker3 con el objetivo de identificar y evitar modelos de Tucker que no capturan completamente la estructura en un diseño experimental. Los resultados revelaron que el ASCA+ no lograba capturar la variación debida a efectos no considerados como factores en el diseño, mientras que un modelo Tucker3 probablemente capturaría esta variación. Además, se observó que en los análisis tipo ANOVA, los efectos de los factores omitidos generalmente se confundían con los efectos estudiados, lo que destacaba la importancia de evitar la selección de modelos Tucker3 demasiado simplificados en el procedimiento ASCA+.

De manera similar, Ramos et al. (2021) destacaron diversas técnicas para el análisis de múltiples vías, como el Análisis de Componentes Principales de múltiples vías, Cuadrados Parciales de múltiples vías, Estructuración a Tres Índices de la Estadística, Tucker3, Factores Paralelos, Análisis de Componentes Independientes de múltiples

vías, Análisis de correlación canónica de conjuntos múltiples, análisis de características lentas y coordenadas paralelas. Finalmente, compararon y discutieron los beneficios y limitaciones de estos métodos.

Por otro lado, Martin-Barreiro et al. (2021) presentaron un algoritmo heurístico para calcular componentes ortogonales disjuntos que facilita el análisis de datos de tres vías y la interpretación de resultados. Destacaron que, aunque las principales técnicas de reducción de dimensionalidad, extracción de patrones y clasificación de datos mediante análisis tensorial se habían basado en el modelo de Tucker, había un importante problema de interpretabilidad de los resultados. Sus resultados evidenciaron mejoras significativas en su algoritmo, permitiendo detectar y capturar estructuras disjuntas en un arreglo de tres vías de acuerdo al modelo Tucker3.

De la misma manera, se muestran en los estudios realizados por Raljević et al. (2021), Wu et al. (2016), Gordon et al. (2021), Reitsema et al. (2023), Dos Santos et al. (2022) y Mitsushita et al. (2023). Estos estudios abordan diversos temas, como el aprendizaje automático para predecir la estabilidad del petróleo crudo, estrategias energéticamente eficientes para redes inalámbricas de área corporal, la estructura y dinámica de sistemas biomoleculares, la dinámica de las emociones en adolescentes, diagnóstico de la enfermedad de Alzheimer, y la percepción del riesgo de desastres naturales. Estos trabajos muestran la diversidad de aplicaciones y la relevancia del análisis de tres vías y el modelo Tucker en una amplia gama de campos de investigación.

Sin duda, es evidente que el análisis de tres vías y el modelo Tucker están ganando cada vez más importancia en las ciencias del comportamiento y la salud. Sin embargo, su aplicación en las ciencias sociales y la economía ha sido hasta ahora limitada. En el campo de la economía, tanto la descomposición de Tucker como el modelo Tucker3 ofrecen un potencial significativo en diversas aplicaciones.

Entre los investigadores destacados en este campo se encuentra Amaya and Pacheco (2002), quienes proporcionaron bases conceptuales tanto teóricas como analíticas para comprender el análisis de arreglos tridimensionales mediante el método Tucker3. Demostraron que a través del Tucker3 es posible reducir un arreglo tridimensional con mediciones de variables sobre individuos en diferentes ocasiones, obteniendo una representación en dimensiones menores que el arreglo de datos original. Además, presentaron un algoritmo de estimación de parámetros basado en un procedimiento iterativo con

un criterio de convergencia que consiste en una diferencia no superior a $1E-10$ entre la suma de cuadrados de los elementos del arreglo central en dos iteraciones consecutivas.

Su aplicación práctica tenía como objetivo describir la situación sociodemográfica y de salud de los 24 departamentos de Colombia durante el periodo 2004-2007. Los datos se organizaron en un arreglo de dimensiones $24 \times 12 \times 4$, donde $I = 24$ individuos representaban los departamentos de Colombia, $J = 12$ variables abarcaban aspectos como la tasa bruta de natalidad, la esperanza de vida al nacer y otros indicadores socioeconómicos, y $K = 4$ condiciones correspondían a los años 2004 a 2007.

Las conclusiones del estudio destacaron similitudes y diferencias entre el método Tucker3 y el Análisis Factorial Multivariante (AFM), especialmente en relación con el preprocesamiento de datos, la representación de individuos, variables y condiciones, y las trayectorias de individuos a través de condiciones. El método Tucker3 demostró una interpretación más clara de las trayectorias de individuos en las condiciones. En cuanto a los resultados, el PCA indicó que la mayoría de la varianza podía explicarse con pocas variables teóricas, mientras que el Tucker3 permitió dimensiones aún más reducidas, evidenciando su ventaja en comparación con tres PCA correspondientes. Sin embargo, se señaló que el procedimiento utilizado para la reducción de dimensionalidad puede influir en la confiabilidad de la evaluación visual.

Por otro lado, Ruiz, Bonilla, Moquera and Arango (2018) llevaron a cabo una evaluación de las diferencias en la estructura económica de las regiones de Colombia mediante un análisis de tres vías. Para ello, consideraron $I = 33$ unidades administrativas de Colombia, $J = 10$ categorías de desagregación del PBI (variables) que incluían actividades como agricultura, ganadería, caza, silvicultura, pesca, entre otras, durante un período que abarcó desde el año 2000 hasta 2015 (ocasiones).

Los autores tenían como objetivo comprender la dinámica económica a través de la estructura de datos tridimensionales. Para lograrlo, plantearon varios objetivos, que incluyeron la evaluación de la estructura de correlación entre los sectores de la actividad económica, la identificación de las fortalezas productivas y su evolución en el tiempo, la búsqueda de similitudes en la estructura económica entre las diferentes regiones y la evaluación de la trayectoria de la estructura económica a lo largo del tiempo.

Para abordar estos objetivos, los autores emplearon el método Tucker3 para determinar las estructuras subyacentes en las tres vías de datos. Realizaron dos tipos de matriciaciones: (i) una vía $I \times JK$ y (ii) una vía $IK \times J$, lo que resultó en la obtención de dos componentes en la primera vía y tres componentes en las otras dos, con una inercia total del $46,43\,\%$.

Los resultados revelaron que la estructura económica de las regiones de Colombia podía ser caracterizada por tres áreas de productividad distintas: (i) industrial y servicios, (ii) recursos naturales y uso del suelo, y (iii) servicios sociales y comunitarios. Además, se observó que durante el período de 2000 a 2006 se produjo una disminución en la productividad del sector minero, seguida de una recuperación en el período de 2006 a 2015, impulsada por el crecimiento del sector agrícola y ganadero.

Un hallazgo interesante fue que la primera componente principal podría utilizarse como un índice de productividad, dado su alto nivel de inercia. Al emplear este índice, se destacaron las disparidades en términos de productividad entre diversas regiones de Colombia, lo que puso de manifiesto las áreas que experimentaban un rezago productivo.

En su obra, Gallo (2015) detalla el proceso de llevar a cabo un análisis Tucker3 en datos de composición, que consisten en vectores de valores positivos que suman una unidad o una constante fija. El autor explora las relaciones entre las matrices de carga al utilizar diferentes procedimientos de preprocesamiento en el análisis. Para ello, emplea datos sobre el consumo de energía de 29 países europeos en cinco categorías (agricultura, industria, servicios, actividades de transporte y hogares) durante un período de 12 años, desde 1997 hasta 2008. El objetivo principal es examinar las proporciones del consumo de energía en las distintas categorías. El arreglo tridimensional se compone de países (individuos) $\times$ categorías (variables) $\times$ años (ocasiones).

Uno de los hallazgos destacados en su investigación es la utilidad de un análisis multivia, como el modelo Tucker3, para analizar datos de composición y comprender las relaciones entre los resultados al emplear diferentes tipos de transformaciones logarítmicas. La conclusión principal es que se debe abordar con cautela la interpretación de los resultados de Tucker, ya que el uso de herramientas de visualización estándar, como biplots conjuntos, podría llevar a interpretaciones erróneas. No obstante, la combinación de partes y composiciones en visualizaciones se revela como una herramienta

poderosa para comprender los resultados.

Por ejemplo, en el estudio realizado por Paredesa et al. (2018), se evaluó la variación en la estructura económica de las diferentes regiones de Colombia durante el período de crecimiento económico nacional de 2000 a 2015. El enfoque del análisis se centró en el Producto Interno Bruto generado por las 33 unidades administrativas del país y se desglosó en 10 categorías principales. Los resultados de su investigación sugieren que la economía de Colombia puede dividirse en tres componentes principales: Productividad Industrial, Productividad en Explotación de Recursos Naturales y Productividad del Sector Social y Comunitario.

En términos generales, el estudio demuestra que durante el período de 2000 a 2015, el crecimiento económico a nivel nacional se reflejó en todos los departamentos. A nivel interno, se destacan departamentos como Santander por su sólido desempeño en la Manufactura, Bogotá y San Andrés como centros prominentes de servicios y comercio, Casanare, Meta y Arauca por su potencial en la agricultura y la minería, y Valle del Cauca, Antioquia, Bolívar y Atlántico como economías equilibradas en manufactura, servicios y agricultura.

Por otro lado, Dell'Anno and Amendola (2015) proponen un índice de exclusión social y analizar la relación del crecimiento económico de los países europeos. Bajo el supuesto que la exclusión social es un fenómeno multidimensional, los autores hacen uso de un análisis de tres vías utilizando el modelo Tucker3. El modelo se aplica para estimar un indicador de exclusión social para 28 países europeos entre 1995 y 2010. Sus resultados sugieren que, a corto plazo, la exclusión social influye en el crecimiento económico en lugar de ser influenciada por él. Además, países con mayores niveles de exclusión social tienen tasas de crecimiento económico per cápita más altas que aquellos con mayor desigualdad de ingresos. En resumen, el análisis sugiere que la inclusión social no parece ser una fuente de crecimiento económico a corto plazo.

Frutos-Bernal et al. (2022) utilizan el modelo de descomposición de Tucker3 para encontrar patrones entre la distribución espacio-temporal de los pasajeros del metro de Barcelona. Para ello, modelan los datos en tres vías: número de pasajeros en la iésima estación en el jésimo momento del késimo día. La descomposición de Tucker3 se utiliza para descubrir grupos espaciales, patrones temporales y las relaciones entre ellos. Los resultados indican que los patrones de viaje difieren entre semana y fines de semana;

además se han identificado las horas punta y valle de cada día y se ha obtenido una clasificación de estaciones.

Como se ha evidenciado previamente, el modelo Tucker3 ha disfrutado de una amplia utilización desde la década de los ochenta. No obstante, su aplicación ha estado mayormente restringida a disciplinas específicas, tales como la psicología, la quimiometría y la detección de señales, con un alcance más limitado en el ámbito matemático.

Sin embargo, esta limitación no ha impedido que, gracias a su versatilidad y sus aplicaciones extensas, el modelo Tucker3 haya generado un creciente interés en diversos campos del conocimiento. Un ejemplo destacado es su incursión en las ciencias sociales y, en particular, en la economía, que es el enfoque central de la presente investigación. Investigaciones anteriores, como las llevadas a cabo por Amaya and Pacheco (2002), Ruiz, Mujica, Sierra, Pozo and Rodellar (2018), Gallo (2015), Paredesa et al. (2018), Dell'Anno and Amendola (2015), y Bautista Mendoza (2009), han analizado y aprovechado la potencialidad del modelo Tucker3 para el análisis e interpretación en contextos sociales y económicos

2.2. Conceptos previos

Como se mencionó previamente, el modelo Tucker3 se utiliza como una generalización en numerosas investigaciones que involucran estructuras de datos de múltiples vías. Por lo tanto, este modelo aplica de manera amplia diversas técnicas estadísticas multivariantes, entre las cuales es fundamentalmente importante distinguir el Análisis de Componentes Principales (PCA), el Análisis Factorial (AF), la Descomposición en Valores Singulares (SVD) y los Mínimos Cuadrados Alternantes (ALS). En la siguiente subsección, se realizará un análisis breve de cada una de estas técnicas. Sin embargo, para obtener una comprensión más detallada de cada una de ellas, se puede consultar la información adicional proporcionada en los anexos.

2.2.1. Análisis de componentes principales

El análisis de componentes principales (ACP) es una técnica multivariante ampliamente utilizada, introducida por Pearson en 1901 y desarrollada por Hotelling en 1933. Es un método matemático que no requiere un modelo estadístico para explicar la estructura probabilística de los errores.

La idea central del ACP es la reducción de dimensiones de un conjunto de datos altamente correlacionados, lo que da como resultado una transformación a un nuevo conjunto de variables, los componentes principales, que no están correlacionados entre ellos y capturan la misma cantidad de información posible de los datos originales. A menudo, la mayor cantidad de la variabilidad de los datos que puede ser explicada por un número menor de componentes principales. El ACP gira los ejes del espacio de datos a lo largo de las líneas de máxima varianza, y la reducción de dimensiones se realiza utilizando solo los primeros componentes principales como base establecida para el nuevo espacio.

Una vez que se han calculado los componentes principales, se pueden utilizar para diversos propósitos, como la visualización de datos en un espacio de menor dimensión, la reducción del ruido en los datos o la detección de patrones y anomalías en los mismos. Además, el ACP se utiliza con frecuencia en la exploración de datos y en el análisis de grandes conjuntos de datos en diversas áreas como la física, la biología, la economía, la psicología, entre otras.

Es importante destacar que el ACP no es una técnica de modelado estadístico, sino una técnica de reducción de dimensiones y visualización de datos. Por lo tanto, aunque puede ser útil en el proceso de modelado, su uso principal es en la exploración de datos y en la identificación de patrones y tendencias en grandes conjuntos de datos.
En resumen, el análisis de componentes principales es una técnica matemática que permite reducir la dimensionalidad de un conjunto de datos altamente correlacionados, a través de la transformación a un nuevo conjunto de variables, los componentes principales, que no están correlacionados entre si. Esta técnica se utiliza ampliamente en la exploración de datos y en el análisis de grandes conjuntos de datos en diversas áreas, y su principal objetivo es facilitar la visualización y comprensión de los datos de manera más eficiente. Para mayor referencia ver el Apéndice A.

2.2.2. Análisis factorial

El análisis factorial busca explicar y describir las relaciones de covarianza entre un conjunto de variables observables en función de variables no observables o *latentes*, denominadas factores. Estos factores no se pueden observar directamente y las variables dependen de ellos, pero también están sujetas a errores aleatorios.
A diferencia del análisis de componentes principales, que se construye para reducir la dimensión, los factores se construyen para explicar las covarianzas o correlaciones. La correlación es un aspecto central en el análisis factorial, ya que la fortaleza del modelo está determinada por la alta correlación entre las variables observables.
El análisis factorial se considera una extensión del análisis de componentes principales, pero la aproximación basada en el modelo factorial es más elaborada y se pregunta si los datos son consistentes con una estructura prescrita. La forma de extracción de los factores latentes es determinante en el análisis factorial, y aquí se considera el papel que los componentes principales juegan en la extracción de los factores latentes. Por lo tanto, el análisis factorial y el análisis de componentes principales comparten similitudes en cuanto a que ambos buscan una representación más simple de un conjunto de datos complejos. Sin embargo, el análisis factorial se enfoca en encontrar las relaciones entre variables observables y factores no observables, mientras que el análisis de componentes principales se enfoca en explicar la variabilidad total en los datos a través de un conjunto más pequeño de variables.

En cuanto al proceso de extracción de factores, el análisis factorial utiliza técnicas de estimación como el método de máxima verosimilitud y el método de mínimos cuadrados ordinarios, mientras que el análisis de componentes principales utiliza la descomposición en valores propios. Aunque ambas técnicas buscan resumir la información contenida en un conjunto de datos, la selección de una técnica u otra dependerá de la pregunta de investigación y las características del conjunto de datos. Para mayor referencia ver el Apéndice C.

2.2.3. Descomposición en valores singulares

La descomposición en valores singulares (SVD) es una técnica de factorización que generaliza el concepto de valores propios para cualquier matriz rectangular, sin necesidad de imponer restricciones sobre su dimensión. La SVD utiliza dos bases diferentes para diagonalizar una matriz: los vectores singulares a derecha e izquierda, que definen una matriz de proyección sobre el subespacio atravesado por las columnas o las filas de la matriz original. La ventaja de esta técnica es que las transformaciones ortogonales preservan normas y ángulos, lo que permite preservar las longitudes de los vectores del error.

La SVD surgió alrededor de 1875, pero su generalización a espacios de dimensión infinita fue desarrollada en el contexto de ecuaciones integrales por Smith y Weyl en 1907 y 1912, respectivamente. Su uso se popularizó a partir de la década del 60, cuando Golub y otros autores empezaron a utilizarla como herramienta para muchos algoritmos. La SVD se define como una factorización del tipo $A = U\Sigma V'$, donde A es una matriz de dimensión $m \times n$, $(m \geq n)$, U es una matriz con columnas ortogonales de $m \times m$, V es una matriz ortogonal cuadrada de $n \times n$, y Σ es una matriz diagonal de $m \times n$. Además, se explica que para cualquier matriz A de $m \times n$, la matriz $A^T A$ de $n \times n$ es simétrica y semidefinida positiva, y sus autovalores son reales y no negativos (≥ 0). Para mayor referencia ver el Apéndice D.

2.2.4. Mínimos Cuadrados Alternados

De acuerdo a Gifi (1990), el algoritmo procede alternando pasos, donde el primer paso busca minimizar una función de pérdida con respecto a $\mathbf{X}^*$ para $\mathbf{a}^*$, y en el

siguiente paso, la función de pérdida es minimizada con respecto a $\mathbf{a}^*$ para un fijo $\mathbf{X}^*$. Es decir, la característica de este algoritmo se basa en resolver problemas de optimización con más de un conjunto de parámetros; cada conjunto se estima a su vez aplicando procedimientos de mínimos cuadrados manteniendo fijos los otros conjuntos. Una vez que se han estimado todos los conjuntos, el procedimiento se repite una y otra vez hasta la convergencia (Kroonenberg and De Leeuw, 1980).

Siguiendo la línea de Gifi (1990), se pueden identificar dos variantes de este algoritmo, las cuales difieren en la forma de normalizar los datos. En una de ellas, se normaliza $\mathbf{X}^*$ manteniendo la escala de $\mathbf{a}^*$; mientras que en la otra, se normaliza $\mathbf{a}^*$ conservando la escala de $\mathbf{X}^*$.

Para el primer caso, las puntuaciones se normalizan y las puntuaciones de los objetos se restringen a la unidad ($\mathbf{X}'\mathbf{X} = 1$). Para iniciar el algoritmo, se requiere una condición inicial arbitraria para $\mathbf{a}$, tras lo cual se siguen los siguientes pasos:

1. Actualizar los valores de $\mathbf{X}$
2. Normalizar $\mathbf{X}^* = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1/2}}$
3. Actualizar los pesos ($\mathbf{a}$)
4. Prueba de convergencia / luego se regresa al paso 1, fijando $\mathbf{a}$.
5. Esto se repite en forma de bucle hasta que se llegué a un nivel de convergencia establecido a priori.

2.3. El análisis de tres vías

Hasta ahora, los métodos expuestos han considerado el análisis en dos vías: individuos y variables. No obstante, un aspecto natural es asumir una vía adicional, lo que conllevaría, naturalmente, a una generalización; esto es lo que se denomina el análisis multivía. Sin embargo, siguiendo en el paso a las tres vías, donde una tercera vía puede tomar un término genérico, que puede depender en gran medida de las diferentes disciplinas, variedad de plantas, señales, pixeles, tiempo, etc.; tales datos pueden ser arreglados en un bloque tridimensional, y este bloque referirse a una arreglo de tres vías.

Notar que la palabra vía (*way* en inglés) se refiere, de acuerdo a Kroonenberg (2008), a un aspecto más general, refiriéndose al arreglo multidimensional independientemente del contenido de los datos; mientras que la palabra modo (*mode* en inglés) es más específica y se refiere al contenido de cada una de las formas. En ese sentido, los objetos, las variables y las condiciones pueden ser los modos de una matriz de datos de tres vías.

Cuando las mismas entidades ocurren en dos diferentes vías, como es el caso en una matriz de correlación, los datos son datos bidireccionales unimodales. De esta manera, las matrices de correlación para las mismas variables están disponibles a partir de varias muestras diferentes, por lo que a menudo se habla de una matriz de datos de dos modos y tres vías, donde las variables y las muestras son los dos modos (Kroonenberg, 2008).

En una disposición de tres vías, los datos son acomodados como sujetos, primera vía, que con un índice i van a lo largo del eje vertical; mientras que las variables, segunda vía, con un índice j van a lo largo del eje horizontal; ahora, una tercera vía que se denominará *condiciones*, tomando el índice k, irá como la profundidad de la caja formada por las dos vías anteriores; luego, la cantidad total de individuos será I, de variables J y de condiciones K, como se aprecia en la figura (2.1):

Los elementos de los arreglos de tres vías, $\underline{\mathrm{X}}$, son denotados por x_{ijk}, donde los índices van desde 1 hasta I, J y K. Para matrices de vías superiores, se utilizan subíndices adicionales, elegidos como las siguientes letras del alfabeto (Kiers, 2000). Esto último

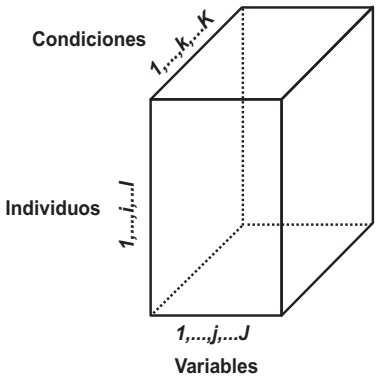


Figura 2.3: Perfil de datos de tres vías

Nota. Tomado de (Kroonenberg, 2008, 34).

es lo que Kroonenberg (2008) denomina niveles, tal como en el análisis de la varianza. De esta manera se tendrá un arreglo de tres vías $I \times J \times K$ definido como una colección de elementos:

$$\{x_{ijk} |\ \ i = 1, \cdots, I; j = 1, \cdots, J; k = 1, \cdots, K\} \tag{2.1}$$

Los datos recopilados se organizan de la siguiente manera:

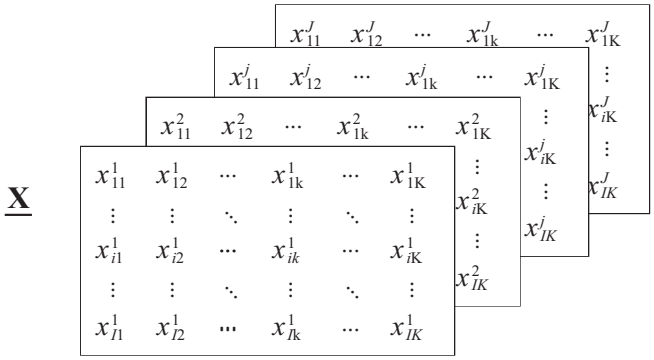


Figura 2.4: Arreglo de datos de tres vías $x_{ijk};\ i = 1, ..., I; j = 1, ..., J; k = 1, ..., K$

Nota. Tomado de (Ruiz, Mujica, Sierra, Pozo and Rodellar, 2018, 3).

El arreglo de N-vías es generalmente denominado *tensor* (De Lathauwer and De Moor, 1998). Así, la noción de *tensor* captura arreglos de diferentes tamaños: un vector de orden I es un tensor en $\mathbb{R}^I$, una matriz de $I \times J$ es un tensor en $\mathbb{R}^{I \times J}$, una matriz de tres vías $I \times J \times K$ es un tensor en $\mathbb{R}^{I \times J \times K}$, etc. En general, un arreglo $I_1 \times I_2 \times \cdots \times I_N$ de N-vías es un tensor en $\mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ (Kiers, 2000; De Lathauwer and De Moor, 1998). En la Figura 2.5 se representa una matriz de tres vías. En esta, las entidades a lo largo del eje vertical se indican con el primer índice (i), las que se encuentran a lo largo del

eje horizontal con el segundo índice (j) y las que se encuentran a lo largo del eje de profundidad con el tercer índice (k). Notar que el uso del primer y segundo índice es el mismo que el común para las matrices (dos vías).

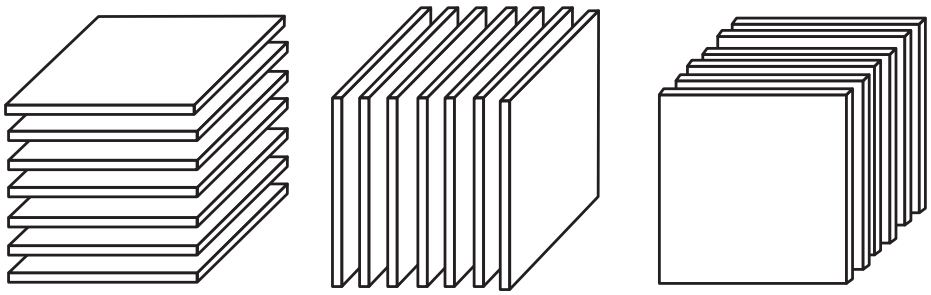


Figura 2.5: Disposición de las vías individuos, variables y condiciones, vistas como *slices* de arreglos

Nota. Tomado de (Kroonenberg, 2008, 29).

Ahora, un arreglo de tres vías puede ser considerado en términos de un conjunto de matrices. Estas matrices forman *slices* horizontales, verticales y frontales, como se observa en 2.5. Específicamente. los I *slices* [1] horizontales pertenecen a los individuos $i = 1, \cdots, I$, los J *slices* de la vía pertenecen a las variables $j = 1, \cdots, J$ y los K *slices* frontales pertenecen a las condiciones $k = 1, \cdots, K$.
Un arreglo de tres vías puede verse como un conjunto de vectores. De ello, se pueden distinguir tres conjuntos de tales vectores, denotados como fibras (*fiber* en inglés): (i) fibras verticales, (ii) fibras horizontales y (iii) fibras de "profundidad". Las fibras [2] horizontales son las filas de una matriz, mientras que las fibras verticales son las columnas. (Kiers, 2000; Kroonenberg, 2008).

En general, estos vectores se denominan *vectores de vía* n, donde n denota la vía sobre la cual se ejecutan los vectores y están asociados con espacios vectoriales particulares, que se denominarán espacios de vía n. Así, para un arreglo de tres vías, el espacio de

[1] Se toma como notación en inglés la palabra *slice* refiriéndose a los planos horizontales y verticales. No se considera su traducción explicita al español por carecer de sentido apropiado para explicar la notación de las tres vías.

[2] Si bien Kroonenberg (2008) gráfica las fibras como elementos de tres dimensiones, esto solo es una aplicación figurativa, que permite mejorar la apreciación gráfica de los elementos. Por su parte Kiers (2000) considera planos enmarcados por líneas horizontales, verticales y diagonales. No obstante, ambas expresan la misma idea.

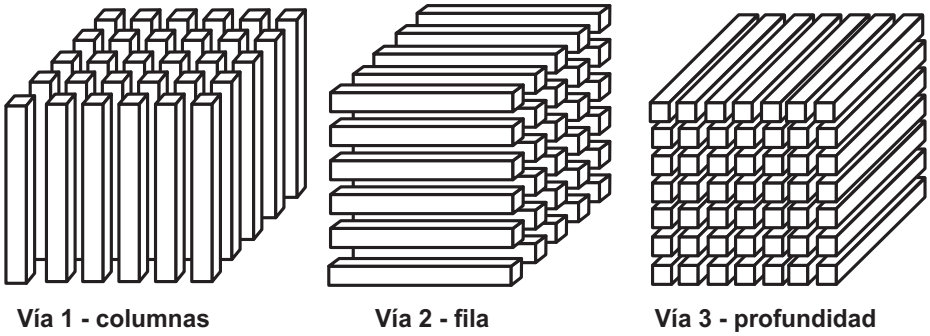


Figura 2.6: Matriz de tres vías, cortada en fibras horizontales, verticales y profundas

la vía individuos es el (sub)espacio abarcado por todas las fibras verticales. Por otro lado, el espacio de la vía de las variables es el (sub)espacio abarcado por todas las fibras horizontales. Finalmente, el espacio de la vía de las ocasiones es el (sub)espacio abarcado por todas las fibras de profundidad.

Ejemplo 2.1 *Considerar el arreglo de la matriz* $4 \times 3 \times 2$ *como planos frontales:*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \quad y \quad \begin{pmatrix} 13 & 14 & 15 \\ 16 & 17 & 18 \\ 19 & 20 & 21 \\ 20 & 23 & 24 \end{pmatrix}$$

los vectores

$$\begin{pmatrix} 2 \\ 5 \\ 8 \\ 11 \end{pmatrix} \quad y \quad \begin{pmatrix} 15 \\ 18 \\ 21 \\ 24 \end{pmatrix}$$

son dos fibras verticales (o vectores de la vía individuos) de la matriz, por su partes los vectores (1 2 3), (7 8 9) y (19 20 21) son ejemplos de tres fibras horizontales (o vectores de la vía variables), mientras que los vectores (1 13), (2 14) y (4 16) son ejemplos de fibras de profundidad (vectores de la vía condiciones). En este caso, el

espacio de la vía variables es generado por los vectores [3]:

$$\begin{pmatrix} 1 \\ 4 \\ 7 \\ 10 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 5 \\ 8 \\ 11 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 6 \\ 9 \\ 12 \end{pmatrix}, \quad \begin{pmatrix} 13 \\ 16 \\ 19 \\ 22 \end{pmatrix}, \quad \begin{pmatrix} 14 \\ 17 \\ 20 \\ 23 \end{pmatrix} \quad y \quad \begin{pmatrix} 15 \\ 18 \\ 21 \\ 24 \end{pmatrix}$$

Tal como lo señala Kiers (2000), para el caso de N-vías, se puede definir subarreglos de cualquier tamaño deseado. Por ejemplo, se puede definir un arreglo de tres vías que son un subconjunto de un arreglo de cuatro vías. Aunque, las matrices formadas no puede recibir nombres intuitivos debido a la complejidad de los arreglos, es posible usar notaciones, donde se usan subíndices de dos puntos para las vías que permanecen intactas e índices ordinarios para indicar las entidades con las que está asociado el subarreglo. Por ejemplo, para una matriz de tres vías, se denota el plano horizontal como: $\mathbf{X}_{i::}$; mientras que para el plano lateral, $\mathbf{X}_{:j:}$; y para el plano frontal, $\mathbf{X}_{::l}$. Y para un arreglo de submatrices de cuatro vías sería: $\mathbf{X}_{i::l}$ y $\mathbf{X}_{:jk:}$. Para el caso de las fibras, estas se pueden denotar como letras en minúsculas resaltadas en negro, usando la misma indexación mostrada anteriormente. Por ejemplo, $\mathbf{x}_{:jk}$, denota una fibra vertical (del vector vía A) en un arreglo de tres vías; por su parte, $\mathbf{x}_{ij:l}$ es una vía del vector C para un arreglo de cuatro vías. De la misma manera, los subarreglos de tres o más vías se denotan con mayúsculas en negrita subrayadas, nuevamente con el mismo sistema de indexación: $\underline{\mathbf{X}}_{::\mathbf{kl}:}$ denota un subarreglo de un arreglo de cinco vías $\underline{\mathbf{X}}$, la cual está asociada a la k-ésima entidad de la vía y l-ésima entidad de la vía cuatro. Finalmente, la situación más común es aquella en la que una matriz de tres vías solo se subdivide en sus planos frontales, estos se denotan simplemente como $\mathbf{X}_1, ..., \mathbf{X}_k$.

[3]Los espacios de la vía variables y condiciones se definen de manera análoga.

2.4. Estructuración de la Matriz de Datos: matriciación o matricización

La palabra *matriciación* o *matricización* es una traducción literal de la palabra inglesa *Matricization*. En el sentido de Kroonenberg (2008) y Kiers (2000) se refiere a una (re) estructuración o un despliegue de matrices. Y es que, tradicionalmente, cuando se enfrenta a una matriz de datos de tres vías, es mejor transformar el arreglo a una matriz de datos de dos vía. De hecho, esto es más fructífero y sencillo para analizar, tal como se utiliza en técnicas estándar para el análisis de dos vías; por lo que, se aprovecha mucho reunir los vectores de la vía n (generalizando), en una sola matriz.

Una forma de estructurar los datos de tres vías en dos vías es a través del promedio de las condiciones; bajo ello se tienen las puntuaciones promedio de los I individuos y las J variables. Por otro lado, es posible la estructuración de matrices a través del posicionamiento de las matrices de cada condiciones a la derecha de cada una, coincidiendo para cada individuos i en la J variables; generando con ello una matriz *amplia*. Esto es que la supermatriz con todas las fibras verticales de una matriz de tres vías se reúnen una al lado de la otra en una matriz $I \times JK$, con entidades de la vía de las variables $(j = 1, ..., J)$ anidadas dentro de los individuos de la vía condiciones $(k = 1, \cdots, K)$, Kiers (2000) denota a ello como $\mathbf{X}_a$.

Este proceso de (re)estructuración de los datos o *matriciación* de una matriz de tres vías es una analogía de *vectorizar* para reorganizar una matriz en un vector. Por otro lado, el proceso inverso se denomina reformar una matriz en una matriz de tres vías.

Otra *matriciación* es la que se forma de manera vertical, anidando la matriz de cada condición k inmediatamente debajo de cada una, en cada una de las j variables, de modo que se replican los sujetos. Esto es que se forma una supermatriz $\mathbf{X}_b$ de orden $J \times IK$, con condiciones k dentro de la vía de los individuos i. De la misma manera se puede formar otra supermatriz $\mathbf{X}_c$ de orden $K \times IJ$ con los individuos i de la primera vía anidada dentro de las variables j [4].

Lo expuesto se denota en la figura 2.7 que representa las tres formas de estructuración

[4]Son posibles otras anidaciones, pero sin más especificaciones, la matriciación pertenece a uno de los procedimientos anteriores. Para mayor referencia ver Kiers (2000) y Flores Camacho (2013)

de los datos: (i) el promedio en una matriz de los promedios de $I \times J$; (ii) la supermatriz $\mathbf{X}_a$ denominada matriz *amplia*; y, (iii) la supermatriz $\mathbf{X}_b$ denominada matriz alta. Finalmente, a manera de resumen y siguiendo a Flores Camacho (2013) se tienen:

$$X_{(I\times J\times K)} \Rightarrow X_{(I\times,JK)} \text{ matriz amplia } \mathbf{X}_a$$

$$X_{(I\times J\times K)} \Rightarrow X_{(J\times,IK)} \text{ matriz alta } \mathbf{X}_b$$

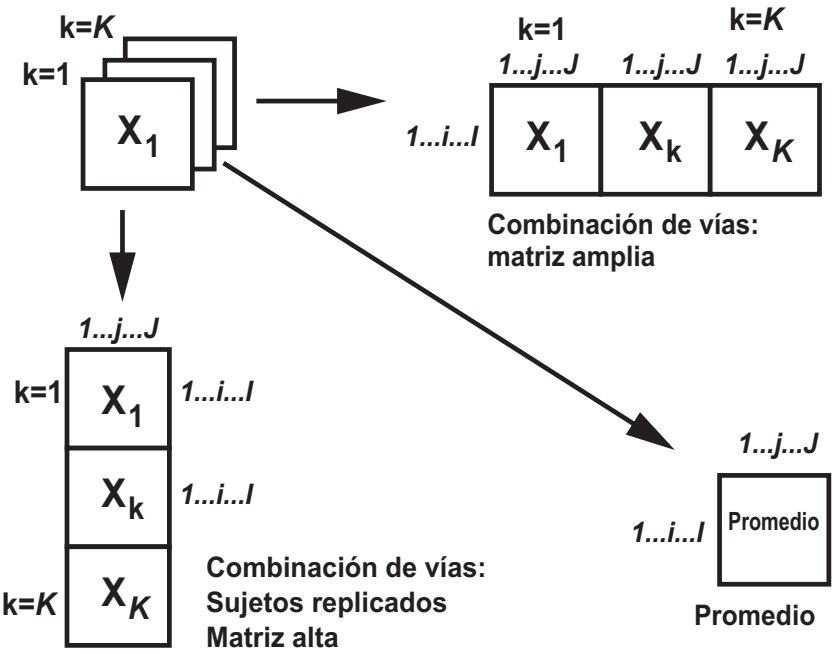


Figura 2.7: Despliegues separados en los lados frontales de un arreglo de tres vías

Nota. Tomado de (Kroonenberg, 2008, 16).

Ejemplo 2.2 *Siguiendo con el ejemplo 2.1 se puede tener:*

$$\mathbf{X}_a = \begin{pmatrix} 1 & 2 & 3 & 13 & 14 & 15 \\ 4 & 5 & 6 & 16 & 17 & 18 \\ 7 & 8 & 9 & 19 & 20 & 21 \\ 10 & 11 & 12 & 20 & 23 & 24 \end{pmatrix}$$

$$\mathbf{X}_b = \begin{pmatrix} 1 & 13 & 4 & 16 & 7 & 19 & 10 & 22 \\ 2 & 14 & 5 & 17 & 8 & 20 & 11 & 23 \\ 3 & 15 & 6 & 18 & 9 & 21 & 12 & 24 \end{pmatrix}$$

$$\mathbf{X}_c = \begin{pmatrix} 1 & 4 & 7 & 10 & 2 & 5 & 8 & \cdots & 9 & 12 \\ 13 & 16 & 19 & 22 & 14 & 17 & 20 & \cdots & 21 & 24 \end{pmatrix}$$

La matricización se puede generalizar para los datos de N vías que, por lo general puede leerse como una estructura de dos vías. Así, una matriz de N vías se representa esencialmente de la misma manera que una matriz de tres vías. Por ejemplo, $\mathbf{X}_a$ (o también podría denotarse como $\mathbf{X}_1$, si se desea denotar las vías con números) contiene todas las fibras verticales reunidas una al lado de la otra en una matriz $I \times JKLM...$, con las columnas las entidades de la vía **B** anidadas dentro de las entidades de la vía **C**, las entidades de la vía **C** anidadas dentro de las entidades de modo **D**, y así sucesivamente. Igualmente, $\mathbf{X}_b$ (o $\mathbf{X}_2$) es la matriz $J \times KLM...I$, las entidades de la columna de la vía **C** anidadas dentro de las entidades de la vía **D**, las entidades la vía **D** anidadas dentro de las entidades de la vía **E**, así sucesivamente hasta que las entidades de la vía **A** son las más externas.

Por lo tanto, $\mathbf{X}_b$ se obtiene de la matriz $\underline{\mathbf{X}}$ permutada cíclicamente de la misma manera que $\mathbf{X}_b$ se obtiene del mismo $\underline{\mathbf{X}}$. Esta equivalencia *permutacional* hace que la programación con matrices de N vías sea relativamente sencilla, incluso cuando los procedimientos son para emplear versiones matricializadas. En el caso que, si las versiones matricializadas de los arreglos reciben nombres además de símbolos, entonces se denominará $\mathbf{X}_a$ a la versión matricial de la vía **A** de $\underline{\mathbf{X}}$, $\mathbf{X}_b$ a la versión matricial de la vía **B** de $\underline{\mathbf{X}}$, etc., y $\mathbf{X}_a$ a la versión matricial de vía n de $\underline{\mathbf{X}}$.En general, existen otras posibilidades para transformar un *tensor* en una matriz. Los procedimientos anteriores deben verse como los *estándar* y preferibles, y el uso de otros siempre debe describirse explícitamente. Por ejemplo, el procedimiento de escribir una matriz $I \times J \times K \times L$ como una matriz $IJ \times KL$ se puede denotar como *matricización* combinando las vías 1 y 2 (1 anidado dentro de 2) y las vías 3 y 4 (3 anidados dentro de 4).

Un caso especial de la matricización es la vectorización, entendida como una representación en la que todos los elementos de la matriz de tres vías o de N vías puede arreglarse como un vector; es decir, colocar las columnas sucesivas de la matriz una debajo de la otra en un solo vector. De esta manera, la vectorización de la matriz **U**

en un vector $\mathbf{u}$ se denota como $\mathbf{u} = Vec(\mathbf{U})$ y:

$$\mathbf{vec}(\mathbf{U}) = [a_{1,1}, ..., a_{m,1}, a_{1,2}, ..., a_{m,2}, ..., a_{m,2}, ..., a_{1,n}, ..., a_{1,n}, ..., a_{m,n}]'$$

Donde $a_{i,j}$ representa $\mathbf{U}(\mathbf{i}, \mathbf{j})$.
La vectorización es un *isomorfismo*:

$$\mathbb{R}^{m \times n} := \mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{mn}$$

Por ejemplo, considerando la matriz siguiente matriz $\mathbf{U}$:

$$\mathbf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la vectorización de $\mathbf{U}$ es

$$\mathbf{vec}(\mathbf{U}) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

Asimismo, la vectorización también puede verse como: $vec(ab') = b \otimes a$; para cualquier vector a y b, no necesariamente del mismo orden. De la misma manera,

$$(\mathbf{vecA})'\mathbf{vecB} = \mathbf{trA}'\mathbf{B}$$

donde A y B son matrices del mismo orden, y tanto el lado derecho como el izquierdo son iguales a: $\sum_i \sum_j a_{ij} b_{ij}$.
Además de las herramientas expuestas anteriormente, en el análisis de múltiples vías a menudo se utilizan determinados productos y operadores de matrices que suelen ser muy específicos; por lo que, es necesario resumirlos antes de empezar con los modelos Tucker. Para esto a_{ij} o $a_{i,j}$ denotan los elementos (i, j) de la matriz $\mathbf{A}$, y $\mathbf{a}_l$ denota la l-ésima columna de $\mathbf{A}$.

El *producto Kronecker* es denotado con el símbolo $\otimes$ y se define como:

$$(\mathbf{U} \otimes \mathbf{V})_{ik,jl} = u_{ij} v_{kl}$$

Así que se tiene:

$$\mathbf{U} \otimes \mathbf{V} = \begin{pmatrix} u_{11}\mathbf{V} & ... & u_{1J}\mathbf{V} \\ \cdots & ... & \cdots \\ u_{I1}\mathbf{V} & ... & u_{IJ}\mathbf{V} \end{pmatrix}$$

Por ejemplo, suponiendo dos matrices $\mathbf{U}$ y $\mathbf{V}$ de 2×2, dadas por:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 & 1 \cdot 6 & 2 \cdot 5 & 2 \cdot 6 \\ 1 \cdot 7 & 1 \cdot 8 & 2 \cdot 7 & 2 \cdot 8 \\ 3 \cdot 5 & 3 \cdot 6 & 4 \cdot 5 & 4 \cdot 6 \\ 3 \cdot 7 & 3 \cdot 8 & 4 \cdot 7 & 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{pmatrix}$$

Entre las propiedades del producto Kronecker se tiene [5]:

- $\mathbf{U} \otimes (\mathbf{V} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{V} + \mathbf{U} \otimes \mathbf{C}$, suponiendo que $\mathbf{A}$ y $\mathbf{B}$ son de iguales dimensiones.
- $(\mathbf{U} + \mathbf{V}) \otimes \mathbf{C} = \mathbf{U} \otimes \mathbf{C} + \mathbf{V} \otimes \mathbf{C}$, suponiendo que $\mathbf{U}$ y $\mathbf{V}$ son de iguales dimensiones.
- $(\mathbf{kU}) \otimes \mathbf{V} = \mathbf{U} \otimes (\mathbf{kV}) = \mathbf{k}(\mathbf{U} \otimes \mathbf{V})$, donde k es un escalar.
- $(\mathbf{U} \otimes \mathbf{V}) \otimes \mathbf{C} = \mathbf{U} \otimes (\mathbf{V} \otimes \mathbf{C})$.
- El producto Kronecker no es *conmutativo*, es decir $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$. No obstante, $\mathbf{A} \otimes \mathbf{B}$ y $\mathbf{B} \otimes \mathbf{A}$ son equivalentes en permutación, lo que quiere decir que existen matrices permutación $\mathbf{P}$ y $\mathbf{Q}$, tales que: $\mathbf{A} \otimes \mathbf{B} = \mathbf{P}(\mathbf{B} \otimes \mathbf{A})\mathbf{Q}$. En particular, si $\mathbf{A}$ y $\mathbf{B}$ son matrices cuadradas, entonces $\mathbf{A} \otimes \mathbf{B}$ y $\mathbf{B} \otimes \mathbf{A}$ son de permutación similar, lo que quiere decir que $\mathbf{P} = \mathbf{Q}'$.
- Si $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ y $\mathbf{D}$ son matrices que pueden formar $\mathbf{AC}$ y $\mathbf{DB}$, entonces:

  $$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD})$$

  Este es el denominado producto mixto, porque mezcla el producto ordinario de matrices y el de Kronecker. De ello se deduce que, $\mathbf{A} \otimes \mathbf{B}$ es invertible si y solo

[5] El producto Kronecker es un caso especial del producto tensorial, por lo que es bilineal y asociativo.

si $\mathbf{A}$ y $\mathbf{B}$ son invertibles, en cuyo caso la inversa esta dada por:

$$(\mathbf{A} \otimes \mathbf{B})^{-\mathbf{1}} = \mathbf{A}^{-\mathbf{1}} \otimes \mathbf{B}^{-\mathbf{1}}$$

También se deduce que:

$$(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}'$$

- Suponiendo que $\mathbf{A}$ y $\mathbf{B}$ son matrices cuadradas de tamaño n y q, respectivamente. Sean $\lambda_1, ..., \lambda_n$ los autovalores de $\mathbf{A}$ y $\mu_1, ..., \mu_n$ los autovalores de $\mathbf{B}$, entonces los autovalores de $\mathbf{A} \otimes \mathbf{B}$ son:

  $$\lambda_i \mu_j \ , \ i = 1, ..., n; j = 1, ..., q$$

  Se deduce que la traza y determinantes de un producto Kronecker viene dado por:

  $$\mathbf{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathbf{tr}(\mathbf{A})\mathbf{tr}(\mathbf{B}) \ \ y \ \ \mathbf{det}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{det}(\mathbf{A}))^{\mathbf{q}}(\mathbf{det}(\mathbf{B}))^{\mathbf{n}}$$

- Si $\mathbf{A}$ y $\mathbf{B}$ son matrices rectangulares, entonces sus valores singulares de la siguiente manera. Suponiendo que $\mathbf{A}$ tiene r_A son los valores no nulos

  $$\sigma_{A,i}, \quad i = 1, ..., r_A$$

  De manera similar para $\mathbf{B}$

  $$\sigma_{B,i}, \quad i = 1, ..., r_B$$

  Entonces, el producto Kronecker $\mathbf{A} \otimes \mathbf{B}$ tiene $r_A r_B$ valores singulares no nulos,

  $$\sigma_{A,i}\sigma_{A,i}, \quad i = 1, ..., r_A; \ j = 1, ..., r_B$$

  Como el rango de una matriz es igual al número de sus valores singulares no nulos, entonces

  $$\mathbf{rang}(\mathbf{A} \otimes \mathbf{B}) = \mathbf{rang}(\mathbf{A})\mathbf{rang}(\mathbf{B})$$

Para representar arreglos convenientes se puede usar el producto Kronecker en algunas ecuaciones matriciales. Si se tiene las matrices $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ dadas y $\mathbf{X}$ una matriz incógnita, entonces $\mathbf{AXB}$ se puede rescribir como:

$$(\mathbf{B}' \otimes \mathbf{A})\mathbf{vecX} = \mathbf{vec}(\mathbf{AXB}) = \mathbf{vecC}$$

Considerando las propiedades del producto Kronecker que la ecuación $\mathbf{AXB} = \mathbf{C}$ tiene solución única si y sólo si $\mathbf{A}$ y $\mathbf{B}$ son inversibles.

Por otro lado, la operación de vectorización matricial se puede escribir como una suma lineal de un producto Kronecker. Sea $\mathbf{A}$ una matriz $m \times n$, la cual será vectorizada, y sea e_i el i-ésimo vector de base canónica para el espacio n dimensional, es decir: $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$. Sea $\mathbf{B}_i$ una matriz de bloques $(mn) \times m$ definida por:

$$\mathbf{B_i} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{I_m} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{e}_i \otimes \mathbf{I}_m$$

Como $\mathbf{B}_i$ consta de n matrices de bloques de tamaño $m \times m$, apiladas en columnas, y estas matrices son todas cero excepto la i-ésima, que es una matriz identidad $m \times m$. Entonces, la vectorización de $\mathbf{X}$ se puede expresar de la siguiente manera:

$$\mathbf{vec}(\mathbf{X}) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{B_i X e_i} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{e_i} \otimes \mathbf{X e_i}$$

La multiplicación de $\mathbf{X}$ por el $\mathbf{e}_i$ extrae la i-ésima columna.

Producto Khatri-Rao o producto de Kronecker por columnas, es denotado por el símbolo $\odot$, se puede calcular entre matrices del mismo orden de columnas y se define de acuerdo con:

$$(\mathbf{U} \odot \mathbf{V})_{ik,l} = u_{il} v_{kl}.$$

Por lo tanto, si $\mathbf{U}$ y $\mathbf{V}$ tienen L columnas, entonces se tiene:

$$(\mathbf{U} \odot \mathbf{V}) = (\mathbf{u_1} \otimes \mathbf{v_1} | ... | \mathbf{u_L} \otimes \mathbf{v_L})$$

Por ejemplo, suponiendo dos matrices $\mathbf{A}$ y $\mathbf{B}$ donde $I = K = 2$ y $L = 3$:

$$\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \quad y \quad \mathbf{B} = \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix}$$

Entonces: $A \odot B$ es la matriz

$$(A \odot B)_{2\times 2,3} = \begin{pmatrix} ag & bh & ci \\ aj & bk & cl \\ dg & eh & fi \\ dj & ek & fl \end{pmatrix}$$

Ahora, considerando el siguiente ejemplo: suponiendo dos matrices $\mathbf{C}$ y $\mathbf{D}$, de 3×3, considerando la partición en columnas:

$$\mathbf{C} = \left(\begin{array}{c|c|c} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 \end{array} \right) = \left(\begin{array}{c|c|c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{array} \right)$$

$$\mathbf{D} = \left(\begin{array}{c|c|c} \mathbf{D}_1 & \mathbf{D}_2 & \mathbf{D}_3 \end{array} \right) = \left(\begin{array}{c|c|c} 10 & 11 & 12 \\ 13 & 14 & 15 \\ 16 & 17 & 18 \end{array} \right)$$

De está manera:

$$\mathbf{C} \odot \mathbf{D} = \left(\begin{array}{c|c|c} \mathbf{C}_1 \otimes \mathbf{D}_1 & \mathbf{C}_2 \otimes \mathbf{D}_2 & \mathbf{C}_3 \otimes \mathbf{D}_3 \end{array} \right) = \left(\begin{array}{c|c|c} 10 & 22 & 36 \\ 13 & 28 & 45 \\ 16 & 34 & 54 \\ 40 & 55 & 72 \\ 52 & 70 & 90 \\ 64 & 85 & 108 \\ 70 & 88 & 108 \\ 91 & 112 & 135 \\ 112 & 136 & 162 \end{array} \right)$$

Producto Hadamard, también llamado producto por elemento, es denotado por el símbolo $*$, y se define como un cálculo que se realiza entre matrices del mismo orden y se define de acuerdo con:

$$(\mathbf{U} * \mathbf{V})_{ij} = u_{ij} v_{ij}$$

Nótese que el producto Hadamard se encuentra en la multiplicación de matrices de dos productos Kronecker en columnas, esto debido a que $(\mathbf{U} \odot \mathbf{V})'(\mathbf{U} \odot \mathbf{V}) = (\mathbf{U}'\mathbf{U}) * (\mathbf{V}'\mathbf{V})$ [6]. Considerando el siguiente ejemplo:

$$\mathbf{A} * \mathbf{B} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} * \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} = \begin{pmatrix} ag & bh & ci \\ dj & ek & fl \end{pmatrix}$$

Además, se tiene que $vec(\mathbf{A} * \mathbf{B}) = \mathbf{vec}(\mathbf{A}) * \mathbf{vec}(\mathbf{B})$

Finalmente, como antesala al modelo Tucker3, y considerando a Kroonenberg (2008), se presentará el modelo de análisis de componentes principales (PCA) [7] para una matriz de datos de dos vías, la cual es vista como un modelo del tipo multiplicativo. De esta manera la descomposición de $\mathbf{X}$ es descrita como la siguiente ecuación:

$$x_{ij} = \hat{x}_{ij} + e_{ij} = \sum_{s=1}^{S} a_{is} f_{js} \quad (i = 1, ..., I; j = 1, ..., J) \tag{2.2}$$

Donde s representa un componente, S corresponde al total de componentes extraídos en la solución, a_{is} es el coeficiente del i-ésimo individuo sobre el s-componente (llamado en inglés *score*), f_{js} corresponde al coeficiente de la j-ésima variables sobre el s-ésimo componente (llamado en inglés *loading*), y e_{ij} es el residual para el i-ésimo individuo en la j-ésima variables. En notación matricial:

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} = \mathbf{A}\mathbf{F}' + \mathbf{E}$$

donde $\mathbf{A} = (a_{is})$ es una matriz normalizada $I \times S$ que contiene los coeficientes de los individuos sobre los S componentes, $\mathbf{F} = (f_{js})$ es la matriz $J \times S$ que contiene los coeficientes de las variables sobre los S componentes, $\mathbf{E}$ es, por su parte, una matriz de residuos $I \times J$. La matriz $\hat{\mathbf{X}}$ es denominada como la base o la imagen estructural de la matriz de datos (Kroonenberg, 2008; Bro, 1998).

El modelo expresado en la ecuación (2.2) se descompone en sus valores singulares, tal como se vio en sección (1.5.) y (1.6.), con lo cual los coeficientes de las variables

[6] El producto Hadamard de dos matrices semidefinidas positivas es semidefinida positiva. además: $det(A * B) \geq det(A)det(B)$.

[7] Como se ha descrito en capítulos anteriores, el PCA es un modelos que describe las relaciones estructurales de las relaciones entre las entidades en los datos. De esta manera, se asume que los patrones en los datos pueden ser descritos por relaciones bilineales entre los componentes para los sujetos (filas) y sus variables (columnas), tal como se observa en la ecuación (2.2).

(f_{js}) puede descomponerse como el producto de los coeficientes normalizados de las variables (b_{js}) y la variabilidad de los componentes (g_{ss}), es decir,

$$x_{ij} = \sum_{s=1}^{S} a_{is} f_{js} + e_{ij} = \sum_{s=1}^{S} a_{is}[g_{ss} b_{js}] + e_{ij} \tag{2.3}$$

o, en notación matricial:

$$\mathbf{X} = \mathbf{AGB}' + \mathbf{E} \tag{2.4}$$

donde $\mathbf{B} = (b_{js})$ es una matriz normalizada $J \times S$ que contiene los puntajes (*scores*)de las variables sobre los S componentes y $\mathbf{G} = (g_{pp})$ es una matriz diagonal $S \times S$ que contiene los valores singulares de la matriz $\mathbf{X}$. Los componentes de $\mathbf{A}$ son denominados los vectores singulares izquierdos y los de $\mathbf{B}$ son los vectores singulares derechos.
La descomposición en valores singulares se puede denotar como la suma de matrices de rango uno, de la forma:

$$\mathbf{X} = \sum_{\mathbf{s=1}}^{\mathbf{S}} \mathbf{g_{ss}}(\mathbf{a_s} \otimes \mathbf{b_s}) + \mathbf{E}$$

$$\mathbf{X} = g_{11}(\mathbf{a_1} \otimes \mathbf{b_1}) + g_{22}(\mathbf{a_2} \otimes \mathbf{b_2}) + \cdots + g_{ss}(\mathbf{a_s} \otimes \mathbf{b_s}) + \mathbf{E}$$

donde g_{ss} corresponde al s-ésimo valor singular, a_s es el vector singular izquierdo y $b_s$es el vector singular derecho. Cada producto $\mathbf{a_s} \otimes \mathbf{b_s}$ el denominado el producto exterior, genera una matriz del mismo tamaño que la matriz de datos pero de rango uno. Luego, para obtener la representación de la matriz de datos son necesarias de S aproximaciones, cada una ponderada por su valor singular (g_{ss}).

2.5. Modelo Tucker3

Ledyard R. Tucker propuso el modelo Tucker3 en 1966 como una forma de analizar y comprimir datos multidimensionales. De hecho, fue el primero en proponer dicho concepto en su artículo *Some mathematical notes on three-mode factor analysis*. La idea detrás del modelo Tucker3 es que los datos multidimensionales pueden ser descompuestos en componentes más simples y fácilmente manejables. Es por esta razón que a Ledyard Tucker se le considera el "padre"del modelo Tucker3, porque fue el primero en proponer el concepto de este tipo de factorización, que es una de las bases teóricas del modelo Tucker3. Con ello, sentó las bases para el desarrollo de la teoría y la aplicación del modelo Tucker3 y la familia que se deriva de él en una amplia gama de áreas, incluyendo la investigación en inteligencia artificial y aprendizaje automático.
Desde entonces, la teoría del modelo Tucker3 ha evolucionado y ha sido aplicado en una amplia gama de áreas, siendo ampliamente citado e influenciando significativamente el desarrollo y la aplicación del modelo Tucker3. Después de su propuesta por Ledyard R. Tucker en 1966, el modelo Tucker3 se utilizó inicialmente en aplicaciones de investigación en el campo de la inteligencia artificial. Sin embargo, a medida que las capacidades de computación aumentaron, el modelo Tucker3 y sus variantes se convirtieron en una herramienta valiosa en una amplia variedad de aplicaciones.
Un segundo padre fundador fue Harshman, quien propuso el modelo de factores paralelos (PARAFAC) para extender el análisis de componentes principales de dos a tres vías, tal como lo propuso Tucker. Carroll es el tercer pionero, y sus contribuciones son especialmente en los campos de las diferencias individuales, la escala multidimensional y la agrupación de tres vías. Propuso una gran cantidad de modelos y algoritmos para el análisis de tres vías, como el modelo INDSCAL para la escala de diferencias individuales y el modelo de descomposición canónica de tres vías que es idéntico al modelo de Harshman. El cuarto contribuyente principal fue Kruskal, cuyo trabajo se ocupó especialmente de problemas matemáticos, en particular, las propiedades de unicidad de los modelos de múltiples vías y el rango de las matrices de múltiples vías. A lo largo de las décadas, los investigadores han desarrollado nuevos métodos y algoritmos para mejorar la eficiencia y la precisión de la factorización de tensores, todo ello en la medida que la capacidad computacional iba aumentando y con mayor énfasis con la

acumulación masiva de información; por ello, el modelo Tucker3 se utiliza comúnmente en aplicaciones de minería de datos, visión por computadora y aprendizaje automático. Además, los investigadores han desarrollado extensiones y variantes del modelo Tucker3 para abordar diferentes problemas y aplicaciones. Por ejemplo, la factorización de tensores de rango parcial, que permite representar un tensor con un número más pequeño de componentes, ha sido utilizada con éxito en aplicaciones de compresión de datos.

Es interesante notar que los modelos fundamentales de los cuales se basa Tucker3 fueron propuestos por Tucker (1963), Carroll and Chang (1970) y Harshman et al. (1970). Inicialmente Tucker lo denominó como un modelo de análisis factorial de tres vías. No obstante, cuando Kroonenberg (1983) propuso un algoritmo para el ajuste por mínimos cuadrados lo denominó análisis de componentes principales de tres vías. Asimismo, distinguieron dos casos particulares. El primero considero un modelo que solo reduce dos vías, lo cual denominaron Tucker2, y un segundo modelo en el cual se reduce una vía, al cual denominaron Tucker1. Sin embargo, el modelo original fue el modelo Tucker3 y los anteriores casos particulares. Por su parte, Carroll and Chang (1970) y Harshman et al. (1970) denominaron al modelo CANDECOMP/PARAFAC o, abreviado, modelo CP. A su vez, el modelo CP se puede escribir como matemáticamente como una variante restringida delo modelo Tucker3, teniendo ambas las misma terminología (Kiers, 2000). De hecho, el modelo CANDECOMP/PARAFAC constituye un modelo Tucker con restricciones de no negatividad tal como lo describen Kim and Choi (2007), Cichocki et al. (2009), Phan et al. (2011), Royer et al. (2012), Ishida et al. (2014), Chen et al. (2019), entre otros.
Considerando el sub-capítulo anterior, un arreglo de tres vías, $\underline{\mathbf{X}}$, es denotado por:

$$x_{ijk}|\ i = 1, ..., I; j = 1, ..., J; k = 1, ..., K$$

Donde i representa a los individuos, j a las variables y k a las ocasiones. En ese sentido, el modelo Tucker3 es una generación del modelo de componentes de dos vías, visto en la sub-sección anterior, llevado a tres vías. Por lo que, de manera análoga se representa el modelo de la ecuación (2.2) como:

$$x_{ijk} = \hat{x}_{ijk} + e_{ijk};\quad i = 1, ..., I; j = 1, ..., J; k = 1, ..., K$$

Entonces, el modelo Tucker3 es una descomposición del arreglo de los datos en tres vías, $\underline{\mathbf{X}}$, que puede ser expresado como una suma ponderada del producto exterior de tres matrices componentes, **A** (vía uno), **B** (vía dos), **C** (vía tres), más una matriz **E**, cuyos elementos e_{ijk} denota el término de error para los elemenos de x_{ijk}. Donde $\mathbf{A} = (a_{is})$ es una matriz ortonormal de dimensión $I \times P$ de los coeficientes de los sujetos, $\mathbf{B} = (b_{jp})$ es una matriz ortonormal de $J \times Q$ que expresa las puntuaciones de las variables, $\mathbf{C} = (b_{kr})$ es una matriz ortonormal de tamaño $L \times R$ que contiene los coeficientes para los componentes de la vía ocasiones; finalmente, $\underline{\mathbf{G}} = (g_{pqr})$ es un arreglo de dimensión $P \times Q \times R$. De esta última, las filas de **G** son para los sujetos idealizados y las columnas son para las variables idealizadas (Levin, 1965; Bautista Mendoza, 2009). En términos de sumatoria, también puede verse de la siguiente manera:

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk};\ i = 1, ..., I; j = 1, ..., J; k = 1, ..., K \tag{2.5}$$

o también,

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr}(a_{ip} b_{jq} c_{kr}) + e_{ijk} \tag{2.6}$$

Aquí, a_{ip}, b_{ip} y c_{ip} son los elementos de **A**, **B** y **C**, respectivamente, mientras que g_{pqr} denota los elementos del arreglo central $\underline{\mathbf{G}}$ [8] [9].

Haciendo uso de la matriciación de la vía **A**, el modelo toma la forma:

$$\mathbf{X_a} = \mathbf{A}\mathbf{G_a}(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E_a}$$

donde $\mathbf{X_a}$ es una matriz de dimensión $I \times JK$ que se obtiene de la matriciación de la vía **A** del arreglo de tres vías $\underline{\mathbf{X}}$; $\mathbf{G_a}$ es una matriz de dimensión $P \times QR$ que se deriva de la matriciación por la vía **A** del arreglo central; y, $\mathbf{E_a}$ es una matriz de residuales de dimensión $I \times JK$ (Tucker, 1966).

[8] En el caso general, para la formulación de un modelo Tucker de N vías, este puede ser escrito como:

$$x_{i_1, i_2, ..., i_N} = \sum_{p_1=1}^{P} {}_1 \sum_{p_2=1}^{P} {}_2 \cdots \sum_{p_N=1}^{P} {}_N (\prod_{n=1}^{N} a_{i_n p_n}, n) g_{p_1 ... p_N} + e_{i_1 i_2 ... i_N}$$

La ecuación anterior, es totalmente general (Levin, 1965).

[9] Nótese que si $q_{pqr} = 0$, entonces la p-ésima, q-ésima y r-ésima combinación de los componentes no contribuirán en la imagen (Kroonenberg and de Leeuw, 1993)

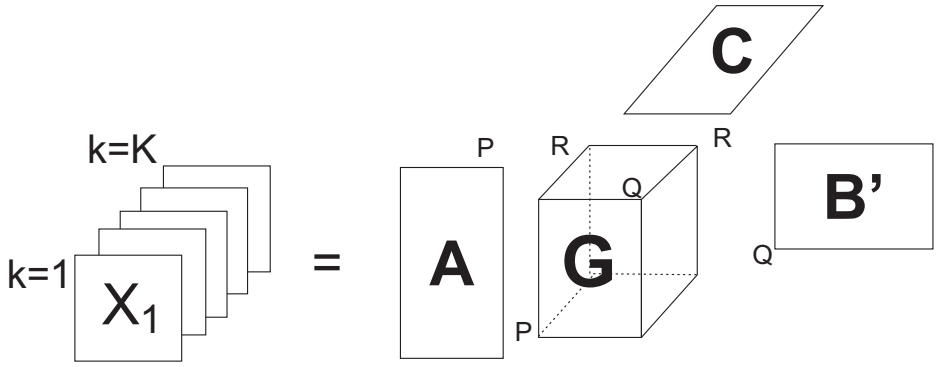


Figura 2.8: Representación del modelo Tucker3

Nota. La caja **G** se puede seccionar en R matrices frontales de orden $p \times q$ cada una. Estas son premultiplicadas por **A** y se posmultiplica por **B**. La caja resultante ahora se puede seccionar lateralmente en Q matrices de orden $P \times R$, y cada una de estas puede posmultiplicarse por **C**. Todo esto reproduce $\underline{\mathbf{X}}$. Tomado de Bautista Mendoza (2009).

De manera análoga, la matriciación de la vía **B** deriva en la siguiente forma:

$$x_{ijk} = \sum_{q=1}^{Q} b_{jq} \sum_{p=1}^{P} \sum_{r=1}^{R} g_{pr}(a_{ip}c_{kr}) + e_{ijk}$$

o matricialmente:

$$\mathbf{X_b} = \mathbf{B}\mathbf{G_b}(\mathbf{C}' \otimes \mathbf{A}') + \mathbf{E_b}$$

donde $\mathbf{X_b}$ es una matriz de dimensión $J \times IK$ que se obtiene de la matriciación de la vía **B** del arreglo de tres vías $\underline{\mathbf{X}}$; $\mathbf{G_b}$ es una matriz de dimensión $Q \times PR$ que se deriva de la matriciación por la vía **B** del arreglo central; y, $\mathbf{E_b}$ es una matriz de residuales de dimensión $J \times IK$.

Finalmente, usando la matriciación de la vía **C**

$$x_{ijk} = \sum_{r=1}^{R} c_{kr} \sum_{p=1}^{P} \sum_{q=1}^{Q} g_{pq}(a_{ip}b_{jq}) + e_{ijk}$$

o matricialmente:

$$\mathbf{X_c} = \mathbf{C}\mathbf{G_c}(\mathbf{B}' \otimes \mathbf{A}') + \mathbf{E_c}$$

donde $\mathbf{X_c}$ es una matriz de dimensión $K \times IJ$ que se obtiene de la matriciación de la vía **C** del arreglo de tres vías $\underline{\mathbf{X}}$; $\mathbf{G_c}$ es una matriz de dimensión $R \times PQ$ que se deriva de la matriciación por la vía **C** del arreglo central; y, $\mathbf{E_c}$ es una matriz de residuales de dimensión $K \times IJ$.

Tal como lo describe Bautista Mendoza (2009), el modelo Tucker3 es una descomposición en valores singulares de orden superior; de manera que, en la construcción de la imagen estructural de $\underline{\mathbf{X}}$ es necesario $P \times Q \times R$ términos de la forma $(a_p \otimes b_q \otimes c_R)$, los cuales están ponderadas por su propio g_{pqr}. Utilizando la notación de rango uno, el modelo Tucker3 tiene la forma:

$$\underline{\mathbf{X}} = \sum_{\mathbf{p=1}}^{\mathbf{P}} \sum_{\mathbf{q=1}}^{\mathbf{Q}} \sum_{\mathbf{r=1}}^{\mathbf{R}} \mathbf{g_{pqr}}(\mathbf{a_p} \otimes \mathbf{b_q} \otimes \mathbf{c_R}) + \mathbf{E}$$

donde a_p, b_q y c_r son los vectores singulares izquierdos obtenidos de las descomposiciones en valores singulares. Por su parte, $(a_p \otimes b_q \otimes c_R)$ se denomina arreglo de rango uno y q_{pqr} es el ponderador para las combinaciones de componentes. Kiers (2000) destaca una versión vectorizada del modelo Tucker3 como:

$$\mathbf{x} = (\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathbf{g} + \mathbf{e} \tag{2.7}$$

donde $\mathbf{x}$, $\mathbf{g}$ y $\mathbf{e}$ denotan la versión vectorizada de $\underline{\mathbf{X}}$, $\mathbf{G}$ y $\mathbf{E}$. La expresión (2.7) muestra de mejor manera la simetría del modelo Tucker3, enfatizando el papel que tienen los términos del núcleo (como pesos de regresión para las columnas $\mathbf{C} \otimes \mathbf{B} \otimes$).

2.5.1. Propiedades del modelo Tucker3

- Como el modelo Tucker3 posee en gran medida las mismas propiedades de la descomposición en valores singulares; en particular, permite una descomposición completa de las tres vías. De la misma manera, es posible particionar la variabilidad de las vías, ya que los componentes son ortonormales. De acuerdo a Kroonenberg (2008), el modelo tiene unicidad subespacial, pero carece de unicidad de los componentes, tal como lo describen Amigo and Marini (2013), los modelos Tucker3 no son únicos, debido a que existe libertad de rotación.

- Aunque la cantidad de componentes para las tres vías puede diferir, existen algunas restricciones integradas en el modelo. En particular, la regla del producto mínimo, el cual dice que el producto del número de componentes de dos vías siempre debe ser igual o mayor que el de la tercera vía, de modo que $P \times Q \geq R$, $P \times R \geq Q$, y $Q \times R \geq P$. La restricción es una consecuencia de la ortogonalidad de tres vías del arreglo central (Kroonenberg, 2008).

2.5.2. Ilustrando el modelo Tucker3

Para ilustrar la racionalidad del modelo Tucker3, Kroonenberg (1983) utiliza un ejemplo de dos sujetos, p_1 y p_2, dos variables y dos q_1 y q_2, en dos condiciones, r_1 y r_2. Esto implica que los datos están clasificados en tres vías: sujetos, variables y condiciones. De esta manera, lo que se busca en términos generales son tres objetivos: (i) determinar las componentes de las variables que explican la mayor parte de la variabilidad de los datos; (ii) conocer qué características generales pueden ser definidas para los sujetos; es decir, ver a los sujetos como combinaciones lineales de sujetos idealizados; finalmente, (iii) ver si las condiciones pueden ser caracterizadas por un conjunto de condiciones prototipo.

Lo anterior puede determinarse mediante el análisis de componentes principales para cada vía (sujetos, variables y condiciones). Ello implica que la variación presente de los datos se analiza de tres maneras diferentes, por tanto, los componentes extraídos deben estar relacionados de alguna manera. La pregunta que se realiza Kroonenberg (1983) es ¿cómo? Entonces, siguiendo el ejemplo, se denomina a los componentes de las variables, *variables latentes*, a las componentes de los sujetos, *sujetos idealizados*, y a los componentes de condición, *condiciones prototipo*.

La relación entre los componentes de las tres vías implica hacer las siguientes preguntas: (i) ¿El sujeto idealizado 1 y el sujeto idealizado 2 reaccionan de manera diferente a la variable latente 2 en la condición prototipo 1? (ii) ¿Es diferente la relación entre los sujetos idealizados y las variables latentes bajo las diversas condiciones del prototipo? Al realizar tres análisis de componentes separados, tales preguntas no se pueden responder de inmediato, ya que no se sabe cómo relacionar los diversos componentes; aunque, el modelo de componentes principales para las tres vías podría dar luces de cómo se pueden determinar las relaciones entre los componentes. De hecho, la matriz de tres vías que engloba estas relaciones se denomina *matriz central*, ya que se supone que contiene las características esenciales de todos los datos.

En esencia, lo que se pretende conocer son los puntajes bajo una variedad de condiciones de *prototipo*. En otras palabras, lo que se pretende saber es cómo reaccionan los sujetos *idealizados* a las variables *latentes* en diversas condiciones de *prototipo*. De esta manera, como se trata de tres vías (sujetos, variables y condiciones *reales*), lo que

se tiene que encontrar es una forma de construir lo *real* a partir del mundo *idealizado*. Una forma de hacerlo es suponer que un sujeto real es una mezcla de individuos *idealizados* y hacer una suposición similar para las variables y condiciones de modo que, las puntuaciones reales puedan considerarse como combinaciones ponderadas de las entidades *idealizadas*. En términos más técnicos, cada variable real es una combinación lineal de las variables latentes.

Lo anterior se puede estructurar mediante la construcción de la puntuación de un individuo i en una variable j bajo la condición k a partir de cantidades idealizadas conocidas. Para ello, se supondrá una disposición de 2 personas idealizadas (p_1, p_2), dos variables latentes (q_1, q_2), y dos condiciones de prototipo (r_1, r_2); con ello, las puntuaciones son las siguientes

sujeto p_1 en la variable q_1 bajo la condición r_1: $c_{p_1q_1c_1}$ *o* c_{111}

sujeto p_1 en la variable q_1 bajo la condición r_2: $c_{p_1q_1c_2}$ *o* c_{112}

sujeto p_1 en la variable q_2 bajo la condición r_1: $c_{p_1q_2c_1}$ *o* c_{121}

sujeto p_1 en la variable q_2 bajo la condición r_2: $c_{p_1q_2c_2}$ *o* c_{122}

De manera similar, se puede saber las puntuaciones del sujeto p_2 : $c_{211}, c_{212}, c_{221}$ y c_{222}. En otras palabras, conocemos todos los elementos de la matriz central. Como se mencionó anteriormente, se quiere construir la puntuación del sujeto real i, sobre una variable real j, bajo una condición real k. Esto se realizará de manera secuencial y, luego, se sistematizarán todas las partes al final. Ahora, bajo el supuesto que la puntuación de un sujeto real i sobre la variable latente q_1 bajo una condición prototipo r_1 es una combinación lineal de las puntuaciones de las personas idealizadas p_1 y p_2, los pesos estarán dados por g_{ip_1} y g_{ip_2}, de manera que:

$$s_{iq_1r_1} = g_{ip_1}c_{p_1q_1r_1} + g_{ip_2}c_{p_2q_1r_1} \tag{2.8}$$

o lo que es lo mismo

$$s_{i11} = g_{i1}c_{111} + g_{i2}c_{211} \tag{2.9}$$

De manera similar. para la variable q_2 bajo la condición r_1:

$$s_{i21} = g_{ip_1}c_{p_1q_2r_1} + g_{ip_2}c_{p_2q_2r_1} \tag{2.10}$$

$$= g_{i1}c_{121} + g_{i2}c_{221} \tag{2.11}$$

y las otras combinaciones de condición variable

$$s_{i12} = g_{i1}c_{112} + g_{i2}c_{212}, \tag{2.12}$$

$$s_{i22} = g_{i1}c_{122} + q_{i2}c_{222} \tag{2.13}$$

Los pesos g_{i1} y g_{i2} indican en qué medida los sujetos idealizados p_1 y p_2 determinan el sujeto real i. La suposición en este enfoque es que estos g_{i1} y g_{i2} son independientes de la prueba y la condición bajo la cual se mide el sujeto. Todas las interacciones entre sujetos, variables y condiciones surgen de las interacciones entre las entidades idealizadas, como se refleja en la matriz central.

El próximo paso es construir el sujeto con puntajes en una variable real j, en lugar de las variables latentes q_1 y q_2, de manera análoga al procedimiento anterior para sujetos. El puntaje del sujeto en la variable real j bajo la condición de prototipo r_1 es

$$\begin{aligned} v_{ij1} &= h_{jq_1}s_{iq_1r_1} + h_{jq_2}s_{iq_2r_1} \\ &= h_{j_1}s_{i_{11}} + h_{j_2}s_{i_{21}} \end{aligned} \tag{2.14}$$

De manera similar, en la variable real j bajo la condición de prototipo r_2 tenemos

$$v_{ij2} = h_{j1}s_{i12} + h_{j2}s_{i22} \tag{2.15}$$

donde los pesos h_{j1} y h_{j2} indican en qué medida las variables latentes determinan la variable real j. Finalmente combinamos las condiciones del prototipo. El puntaje del sujeto en la prueba j bajo la condición k se puede escribir como

$$z_{ijk} = e_{k1}v_{ijl} + e_{k2}v_{ij2} \tag{2.16}$$

donde los pesos e_{k1} y e_{k2} indican en qué medida cada condición prototipo determina la condición real k.

Reuniendo los resultados de los tres pasos obtenemos:

$$z_{ijk} = \sum_{r=1}^{2} e_{kr}v_{ijr} = \sum_{r=1}^{2} e_{kr}(\sum_{q=1}^{2} h_{jq}s_{iqr}) \tag{2.17}$$

que se puede escribir de forma compacta como

$$z_{ijk} = \sum_{r=1}^{2} e_{kr}[\sum_{q=1}^{2} h_{jq}(\sum_{p=1}^{2} g_{ip}c_{pqr})] \tag{2.18}$$

Donde $\sum_{p=1}^{2} g_{ip}c_{pqr}$ es la combinación de los sujetos p_1 y p_2. $\sum_{q=1}^{2} h_{jq}(\sum_{p=1}^{2} g_{ip}c_{pqr})]$ es la combinación lineal de las variables de q_1 y q_2. Finalmente, $\sum_{r=1}^{2} e_{kr}[\sum_{q=1}^{2} h_{jq}(\sum_{p=1}^{2} g_{ip}c_{pqr})]$ es la combinación lineal de las condiciones r_1 y r_2. o de manera más general

$$z_{ijk} = \sum_{p=1}^{s}\sum_{q=1}^{t}\sum_{r=1}^{u} g_{ip}h_{jp}e_{kr}c_{pqr}, \quad cons = t = u = 2 \tag{2.19}$$

como suele escribirse. Tenga en cuenta que el orden en que se toman las combinaciones lineales es irrelevante.

2.5.3. Interpretación del arreglo central

Al arreglo central **G** constituye el elemento fundamental el modelo Tucker3, pues representa la fuerza de las conexiones entre los conjuntos de componentes se refleja en la matriz central, la cual contiene información acerca de la importancia relativa de las combinaciones de componentes en todas las vías. En tal sentido, la interpretación que se deriva puede tener al menos siete formas no relacionadas de interpretarse (Kroonenberg, 2008):

- La fuerza de la relación entre los componentes de las diferentes vías o el peso de la combinación de los componentes.
- Pesos de regresión.
- Porcentaje de variación explicada.
- La interacción de las medidas de las tres vías.
- Las puntuaciones de elementos idealizados o latentes.
- Dirección de los cosenos (Tucker2).
- Covarianzas latentes.

No obstante, como lo describe Kroonenberg (2008):

> La construcción independiente de los espacios componentes supone una gran carga para encontrar una interpretación adecuada del conjunto central. Difícilmente se puede esperar poder interpretar fácilmente los vínculos entre los componentes, si no se sabe lo que significan estos componentes.

> Por lo tanto. Se debe aprovechar cualquier herramienta que se pueda utilizar para mejorar la interpretación de los componentes. El primero de ellos son las rotaciones de los componentes. Sin embargo, una complicación es que las rotaciones influyen directamente en los tamaños de los elementos centrales, ya que los hacen más iguales y, por tanto, más pequeños. Este es el mismo efecto que la redistribución de la varianza de los componentes rotados en PCA de dos modos. (Kroonenberg, 2008, 238).

En efecto, resulta difícil esperar entender los vínculos entre los componentes si no se comprende plenamente el significado de los componentes individuales de las vías en primer lugar. En tal sentido, podría aprovechar los métodos de rotación para mejorar la interpretación, lo cual es análogo al análisis de componentes principales; aunque ello podría influir directamente en las magnitudes de los elementos centrales.

Coeficientes de regresión y la suma de cuadrados

Como se ha mencionado la matriz central contiene la relación que enlaza los componentes de las diferentes vías. En particular, la matriz central indica el peso de una combinación de componentes, uno de cada vía. La tabla (2.1) es un ejemplo de un arreglo central en tres vías del modelo Tucker3. De manera que, el elemento g_{111} indica la fuerza de relación entre el primer componente de las tres vías ($a_{i1}b_{j1}c_{k1}$), y el elemento g_{221} indica la fuerza de relación entre el segundo componente de la primera vía y la segunda componente de la segunda vía con la primera componente de la tercera vía ($a_{i2}b_{j2}c_{k1}$).

Tabla 2.1: Notación para un arreglo central de tres vías

		3era vía					
		1		2		3	
		2da vía		2da vía		2da vía	
Componentes		1	2	1	2	1	2
1era vía	1	g_{111}	g_{121}	g_{112}	g_{122}	g_{113}	g_{123}
	2	g_{211}	g_{221}	g_{212}	g_{222}	g_{213}	g_{223}
	3	g_{311}	g_{321}	g_{312}	g_{322}	g_{313}	g_{323}

Ahora, cuando se reconstruyen los datos como:

$$\hat{x}_{ijk} = \sum_{pqr} [g_{pqr}](a_{ip}b_{jq}c_{kr}) \tag{2.20}$$

el elemento g_{pqr} es el peso del término $(a_{ip}b_{jq}c_{kr})$. Ahora, dichos elementos pueden verse también como los pesos de una regresión, donde los predictores son de la forma:

$$f_{ijk,pqr} = (a_{ip}b_{jq}c_{kr})$$

Entonces, bajo esta interpretación, los elementos de la regresión son los productos triples de los componentes, los cuales son ortonormales. Bajo ello, que se tiene la siguiente ecuación:

$$x_{ijk} = \hat{x}_{ijk} + e_{ijk} = \sum_{pqr} [g_{pqr}](a_{ip}b_{jq}c_{kr}) + e_{ijk} = \sum_{pqr} [g_{pqr}] f_{ijk,pqr} + e_{ijk} \tag{2.21}$$

Como se supone que los componentes son ortonormales, entonces su producto triple también es ortonormal. De esta manera, con predictores ortonormales, la suma de los coeficientes de la regresión al cuadrado son la suma explicada de los cuadrados de la ecuación de regresión. Cada coeficiente al cuadrado g^2_{pqr} indica la variabilidad explicada que se utiliza en lugar de su predictor. Entonces se tiene que:

$$\sum_{pqr} g^2_{pqr} = SS \;\; (ajuste) \;\; (ModeloTucker3) \tag{2.22}$$

Por lo tanto, la suma ajustada de los cuadrados puede dividirse completamente por los elementos de la matriz central. Al dividir por la suma total de los cuadrados se obtiene la proporción de la variabilidad explicada para cada combinación de componentes.
Esta interpretación se puede utilizar para evaluar si los componentes adicionales en cualquier vía introducen nuevas combinaciones considerables de componentes, o si la ganancia en la variabilidad explicada se distribuye entre una mayor cantidad de elementos adicionales de la matriz central.

Puntuaciones de los elementos idealizados

Como se vio en la ilustración del modelo Tucker3, la matriz central puede verse como los puntajes de la combinación de diferentes individuos idealizados, variables latentes, y una condición prototipo. Por tanto, los componentes son constructos no observables, cuyos niveles observados de las vías son combinaciones lineales: una persona

real es una combinación lineal de tipos de sujeto, una variable real es una combinación lineal de variables latentes y una condición real es una combinación lineal de las condiciones del prototipo.

Para visualizar lo expresado anteriormente, se considera un modelo de combinaciones lineales anidadas; que puede expresarse de la siguiente manera:

$$x_{ijk} \approx \sum_p a_{ip} \left[\sum_q b_{jq} \left(\sum_r c_{kr} q_{pqr} \right) \right]$$
$$x_{ijk} \approx \sum_p a_{ip} \left[\sum_q b_{jq} h_{pqk} \right]$$
$$x_{ijk} \approx \sum_p a_{ip} w_{pjk}$$

Donde $\sum_r c_{kr} q_{pqr}$ dice que las puntuaciones de un sujeto idealizado p en la condición k son una combinación lineal de las puntuaciones de las condiciones del prototipo R en variables latentes. Además, las puntuaciones de las variables reales son combinaciones lineales de las puntuaciones de las variables latentes Q y, finalmente, todos los sujetos reales son combinaciones lineales de los tipos de sujetos.

$\mathrm{f}_{ijk,pqr} = g_{pqr}(a_{ip} b_{jp} c_{kr})$

2.5.4. Puntuaciones (*scores*) y cargas (*loadings*) en el modelo Tucker3

Siguiendo a Kroonenberg (2008), se parte de un modelo de componentes principales de dos vías:

$$\mathbf{X} \approx \mathbf{U}\mathbf{L}'$$

donde $\mathbf{U}$ son los coeficientes de la fila (sujeto) y $\mathbf{L}$ los coeficientes de columna (variable). Luego, existen tres opciones para desarrollar un sistema con puntuaciones y cargas para el modelo Tucker3: (i) la forma de covarianza, (ii) la forma de distancia, (iii) la forma de Murakami.

La forma de covarianza parte de un modelos PCA estándar, dado por $\mathbf{U} = \mathbf{A}$ y $\mathbf{L} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G_a}$; entonces $\mathbf{U}$ es una matriz de puntuación de componentes ortonormales y $\mathbf{L}$ es la estructura ortogonal de la matriz de carga. De modo que la ortogonalidad

está dada por:

$$\mathbf{L}'\mathbf{L} = \mathbf{G}_\mathbf{a}'(\mathbf{C}' \otimes \mathbf{B}')(\mathbf{C} \otimes \mathbf{B})\mathbf{G}_\mathbf{a} = \mathbf{G}_\mathbf{a}'\mathbf{I}_\mathbf{QR}\mathbf{G}_\mathbf{a} = \sum \mathbf{G}_\mathbf{r}'\mathbf{G}_\mathbf{r} \tag{2.23}$$

Kroonenberg (2008) sugiere que sería inútil buscar la interpretación de la descomposición de $\mathbf{L}$ en una parte ortonormal $\mathbf{C} \otimes \mathbf{B}$ y una parte ortogonal $\mathbf{G}$, pues no tiene una interpretación útil.

2.5.5. Estandarización de los datos

Debido a que las variables numéricas a menudo tienen diferentes escalas y rangos de medición, sus valores no son necesariamente comparables y casi siempre se requiere algún tipo de procesamiento previo. Así, como en todo análisis multidimensional, es útil el preprocesamiento de los datos, tal como sucede en el análisis bidimensional. El preprocesamiento se refiere a la aplicación de técnicas a un conjunto de datos antes de que se ajuste un modelo de múltiples vías.

En general, en el análisis multidimensional los datos a menudo son centrados o se normalizan en las filas para eliminar las diferencias no deseadas en el nivel y la escala[10]. De la misma manera, en el análisis de tres vías, existe aún la necesidad de eliminar tales diferencias, aunque el sentido ya no es obvio o fácilmente interpretable. Por lo que, a diferencia de un análisis de dos vías en el que los datos suelen estar centrados o normalizados en las filas, en un contexto de tres vías, no es sencillo cómo preprocesar los datos, ya que hay diferentes opciones disponibles.

Sin embargo, el efecto del procesamiento es que los coeficientes de las variables en los componentes se pueden ajustar de tal manera que se pueden interpretar como correlaciones entre variables y componentes. De esta manera, el preprocesamiento es la única forma de obtener esta interpretación (Kroonenberg et al., 2009).

De hecho, Kroonenberg (2008) refiere que:

> La implicación es que las diferencias en las medias y escalas de las variables no son de interés, oscurecen la visión de las relaciones entre las variables y obstaculizan el descubrimiento de cómo los sujetos difieren en

[10]La escala generalmente se refiere a alguna medida de variabilidad, más a menudo la desviación estándar.

> sus puntuaciones en estos (grupos de) variables. En otras palabras, la información requerida sólo puede modelarse correctamente con las puntuaciones de desviación normalizadas y no con puntuaciones brutas. Una razón por la que esto es así es que las medias no contienen información sobre las diferencias individuales, pero las puntuaciones de desviación sí. Además, las variables suelen medirse en una escala de intervalo, de modo que las medias no contienen información intrínseca. y para una interpretación significativa no deben incluirse en el análisis. Además, cuando las unidades de medida de las variables son diferentes, o alternativamente las diferencias en las variabilidades son idiosincrasias de las variables más que portadoras de información sobre las diferencias entre sujetos, la escala de las variables debe eliminarse antes del análisis multiplicativo (Kroonenberg, 2008, 136).

En ese sentido, kroonenberg considera que para en análisis de múltiples vías, la información necesaria solo se puede representar adecuadamente mediante puntuaciones de desviación normalizadas en lugar de puntuaciones sin ajustar. Esto se debe a que las medias carecen de detalles sobre las diferencias individuales, mientras que las puntuaciones de desviación las incluyen. En particular, los problemas aumentan porque los datos de tres o más vías a menudo se reorganizan en matrices de dos vías antes del preprocesamiento. Esto es problemático, porque introduce una vía de columna que es una combinación de dos vías originales.

Bro and Smilde (2003) presentan dos tipos básicos de preprocesamiento: (i) el centrado o ajuste aditivo, y (ii) el escalado multiplicativo.

Por un lado, el centrado tiene que ser realizado por una transformación dentro de una vía en específico, lo cual permite eliminar las diferencias artificiales en niveles y escalas. En ese sentido, el centrado es útil para obtener datos de escala de razón (en lugar de escala de intervalo)[11], es decir, los valores observados deben ser proporcionales y un valor cero indica una falta de la propiedad que se mide.

De esta manera, el centrado elimina un término constante en todo elemento, con lo que, los datos resultantes tendrán media cero, lo que debe realizarse a través de una única vía y no a través de combinaciones (Giordani et al., 2014). El tipo de centrado se indica especificando a través de qué vía se centran los datos (Bautista Mendoza,

[11]Notar que el modelo Tucker3 necesita de datos de escala de razón

2009). De esta forma, el centrado a través de la vía **A** corresponde a promediar los datos sobre los individuos, lo cual se calcula como:

$$\tilde{x}_{ijk} = x_{ijk} - \bar{x}_{.jk}$$

Se recomienda este centrado porque elimina el puntaje promedio en cada variable para cada ocasión, de modo que todos los puntajes se desvían de los puntajes de la *persona promedio*. Esta es la persona cuyas puntuaciones están exactamente en la media de cada variable en cada ocasión (Kroonenberg et al., 2009). De hecho, diversos autores como Kruskal (1981), Harshman (1984) argumentaron que solo los centrados de fibra, son apropiados para los modelos de tres vías porque solo para tales centrados los componentes después del centrado tienen una relación simple con los componentes antes de la transformación. Dicho de otra manera, tales centramientos *no destruyen el acuerdo con el modelo*.

En la misma línea Bro and Smilde (2003) mencionan que el centrado a través de la primera vía[12] implica que cada columna de los datos originales se le calcula como la diferencia de cada elemento de la columna subtraido un valor promedio. Por tanto, se obtiene una matriz de valores medios de dos vías. Esto quiere decir que el centrado es idéntico a adecuar los datos en una estructura de dos vías y centrar estos datos de dos vías en la primera vía, tal como se muestra en la figura (2.9).

Asimismo, si se centra una matriz de estructura ($variables \times tiempo \times muestra$) restando el promedio calculado entre muestras y tiempos, entonces, de acuerdo con el ejemplo anterior, se introducen compensaciones artificiales y el modelo subsiguiente también tendrá que ajustarse a esta variación adicional. Esto podría derivar a problemas en la validación y la exploración del modelo y, por extensión, dará lugar a modelos que no proporcionen soluciones generales de mínimos cuadrados.

Kroonenberg et al. (2009) considera que en estudios longitudinales, este tipo de centrado tiene el efecto de eliminar las tendencias temporales en ocasiones, por lo que estas tendencias deben investigarse independientemente de la variación alrededor de estas tendencias. El centrado sugerido tiene la ventaja adicional de que si hay un punto cero real significativo, pero desconocido, de la escala, se elimina automáticamente de

[12]El centrado a través de la primera vía también se le denomina centrado de fibra

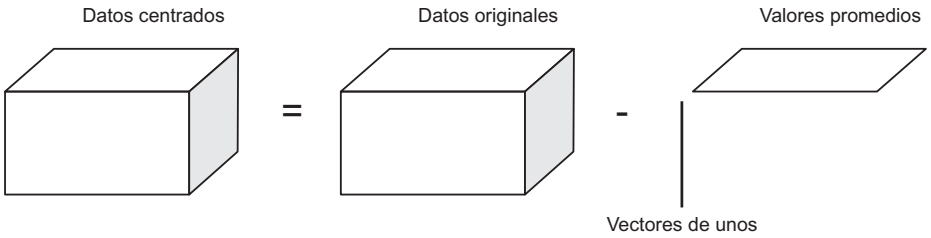


Figura 2.9: Centrado a través de la primera vía

Nota. Tomado de (Bro and Smilde, 2003, 17).

las puntuaciones. Otro efecto de este centrado es que los coeficientes componentes de los sujetos se desvían de sus medias.

Es de notar que el centrado no cambia las direcciones de la solución no centrada, sino que solo las desvía de sus medias. Por otro lado, aunque las direcciones pueden permanecer iguales, la variabilidad relativa explicada de las soluciones no centradas y centradas no guardan relación entre sí. Además, las soluciones centradas y no centradas pueden ser muy diferentes, y solo deben compararse después de rotarlos al máximo (Kroonenberg, 2008; Kroonenberg et al., 2009).

Consecuentemente, Kroonenberg (2008) hace notar que:

> La consecuencia de eliminar cualquier media es que la cantidad de variación explicada por un análisis subsiguiente será más pequeña y, a veces, dramáticamente más pequeña que el análisis no centrado, porque la suma de los cuadrados causada por las medias distintas de cero suele ser la más grande presente en los datos. El centrado de fibra o a través de la vía A para las variables puede, por ejemplo, interpretarse geométricamente como moviendo el centroide de los datos al origen, y así la suma de cuadrados causada por las medias de las variables (que a menudo no es significativa, debido a diferentes escalas de medición o debido a datos de intervalo) se elimina. Por lo tanto, la pérdida en la variación explicada (artificial) no debe lamentarse en tal caso (Kroonenberg, 2008, 121).

Lo descrito por Kroonenberg (2008) va en línea con los argumentado por Kruskal (1981), Harshman (1984). No obstante, también es necesario considerar que la estimación de mínimos cuadrados es sensible a los valores atípicos, por lo que se puede imaginar fácilmente que una o unas pocas observaciones atípicas podrían influir en

gran medida en las medias. En ese sentido, de eliminarse estos valores atípicos, será difícil interpretar los patrones multiplicativos subyacentes de los datos debido al sesgo. Una solución a esos casos puede ser usar medidas robustas, de manera que se reemplace la media por las medianas u otras medidas robustas. La cuestión de la robustez es la más importante, porque a veces surgen interacciones (de tres vías) no tanto debido a patrones sistemáticos sino debido a observaciones periféricas.

Por su parte, el escalado o normalización divide todo elemento por un término constante, con lo cual se eliminan diferencias de escala entre las variables. Bautista Mendoza (2009) sugiere escalar cualquier ocurrencia,k, de una variable por la misma cantidad Escalar dentro de los elementos de la vía B (variables) lleva a calcular el siguiente factor:

$$v_j = \sqrt{\sum_{i=1}^{I}\sum_{k=1}^{K} x_{ijk}^2 / IK}$$

Con ello, los datos escalados se derivan como:

$$\tilde{x}_{ijk} = \frac{x_{ijk}}{v_j}$$

El objetivo principal de la normalización es ponderar diferencialmente ciertas partes (principalmente variables y/o sujetos) de los datos para cambiar la forma en que influyen en el resultado de un análisis. Esto puede hacerse porque las escalas de medida de las variables no son comparables, o con el objetivo de igualar sus varianzas para igualar su influencia en la solución (Kroonenberg, 2008; Bro and Smilde, 2003).

Por su parte Kiers (2000) sugiere que puede ser útil centrar en varias vías diferentes sucesivamente o normalizar dentro de varias vías sucesivamente. Por ejemplo, centra primero en A y luego en B (o viceversa, lo que da el mismo resultado). De manera similar, la normalización primero dentro de A y luego dentro de B se puede denotar como normalización dentro de A y dentro de B, respectivamente. Aquí es importante dar el orden de las normalizaciones, porque una normalización afecta a la otra, y el orden inverso generalmente conduce a un resultado diferente.

Por otro lado, Kroonenberg (2008) sugiere que para datos de perfil de múltiples vías de sujetos (objetos) por variables por condiciones (situaciones y/o puntos de tiempo),

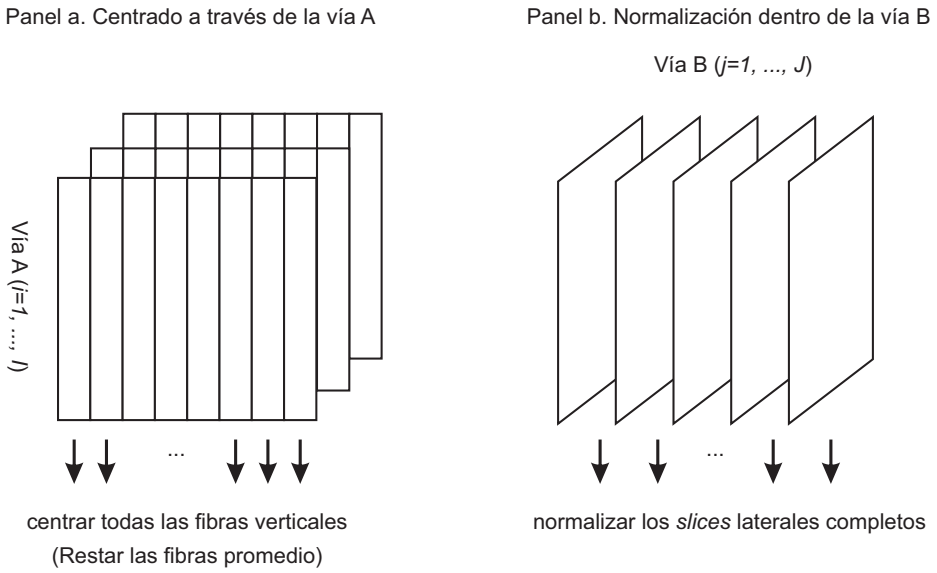


Figura 2.10: Visualización de centrado y la normalización

Nota. Tomado de (Kiers, 2000, 119).

generalmente se usará preprocesamiento de centrado de columna (es decir, fibra) (entre sujetos) y normalización de corte lateral por variable. Notar que, después del centrado de fibra, los datos contendrán todas las diferencias individuales, es decir, la interacción sujeto por variable, la interacción sujeto por condición y la interacción triple sujeto por (variable por condición)[13].

También es importante considera que el paso de normalización no hace que los datos sean consistentes con el modelo, pero permite evitar que los resultados se vean afectados por diferencias de rango entre entidades de uno o más vías (Giordani et al., 2014; Kiers, 2000).

De acuerdo a Bautista Mendoza (2009), el modelo Tucker3 requiere en su etapa previa un único procedimiento de matriciación, lo cual corresponde a la vía en la cual se desea aplicar el centrado; en general, se realiza sobre la vía A o la primera vía, generando $\mathbf{X}_a$ de dimensión $I \times JK$. De hecho, en el modelo Tucker3 el centrado puede realizarse por cualquier vía. Sobre la matriz centrada, el método realiza matriciación por la

[13]Cuando las escalas de calificación de tres vías no se manejan como datos de perfil sino como datos de escala de calificación, en su mayoría organizados como conceptos (situaciones, condiciones) por escalas por sujetos, se debe considerar seriamente el doble centrado, en particular, el centrado de fibra entre conceptos (es decir, , por variable) y centrado de fibra a través de escalas (es decir, por fila), siempre que las escalas no sean bipolares. El resultado será que los datos contendrán principalmente la interacción concepto por escala y que las diferencias individuales solo estarán presentes en la interacción de tres vías (Kroonenberg, 2008; Kiers, 2000; Kroonenberg et al., 2009)

vía variables o segunda vía ($J \times IK$) y escala dividiendo cada fila por la desviación estándar correspondiente. En ese sentido, si se parte de la matriz $\mathbf{X}_a$, esto corresponde a centrar a través de individuos y escalar a través de variables (Bautista Mendoza, 2009).

2.5.6. Estimación de parámetros del modelo Tucker3

Para la estimación de los parámetros, Tucker (1966) propone métodos algébricos para tal resolución, los cuales son muy bien descritos por Bautista Mendoza (2009). Como el modelo Tucker3 es la descomposición en cuatro matrices denominadas $\mathbf{A}, \mathbf{B}, \mathbf{C}$ y el arreglo central $\underline{\mathbf{G}}$ puede existir situaciones en que la solución sea exacta, en esos casos las propuestas de Tucker son útiles. No obstante, ante situaciones en que las soluciones sean aproximadas, Kroonenberg and De Leeuw (1980) describen algoritmos vías mínimos cuadrados alternantes para estimar dichos parámetros.
De acuerdo a Bautista Mendoza (2009), el algoritmo de mínimos cuadrados alternantes se basa:

> El algoritmo de mínimos cuadrados alternantes se basa en dividir los parámetros en grupos. Cada conjunto de parámetros se estima en el sentido de mínimos cuadrados condicionalmente sobre los parámetros restantes. La estimación de los parámetros se repite iterativamente hasta que los valores de los parámetros estimados no difieran significativamente (Bautista Mendoza, 2009, 6).

De acuerdo a Bro (1998), el algoritmo Tucker3 genera las matrices $\mathbf{B}$ y $\mathbf{C}$ como punto de partida del algoritmo. El arreglo central $\underline{\mathbf{G}}$ es estimado minimizando la función de pérdida e itera hasta la convergencia. Bautista Mendoza (2009) lo resume de las siguiente manera:

1. Determinar el criterio de parada o convergencia del algoritmo.
2. Inicializar $\mathbf{B}$ y $\mathbf{C}$.
3. $[\mathbf{A}, \mathbf{S}, \mathbf{V}] = dvs(\mathbf{X}_{(I \times JK)}(\mathbf{C} \otimes \mathbf{B}), P)$.
4. $[\mathbf{B}, \mathbf{S}, \mathbf{V}] = dvs(\mathbf{X}_{(J \times IK)}(\mathbf{C} \otimes \mathbf{A}), Q)$.

5. $[\mathbf{C}, \mathbf{S}, \mathbf{V}] = dvs(\mathbf{X}_{(K\times IJ)}(\mathbf{B} \otimes \mathbf{A}), R)$.

6. Ir al paso 3 hasta cumplir con el criterio de convergencia.

7. Calcular $\mathbf{G}_a = \mathbf{A}'\mathbf{X}_a(\mathbf{C} \otimes \mathbf{B})$.

Donde, $[\mathbf{A}, \mathbf{S}, \mathbf{V}] = dvs(\mathbf{X}_{(I\times JK)}(\mathbf{C} \otimes \mathbf{B}), P)$ es la descomposición en valores singulares de $\mathbf{X}_{(I\times JK)}(\mathbf{C} \otimes \mathbf{B})$, con lo cual se genera la matriz $\mathbf{A}$, con: (i) los primeros P vectores singulares izquierdos, (ii) la matriz $\mathbf{V}$ con los primeros P vectores singulares derechos y (iii) la matriz diagonal $\mathbf{C}$, cuyos elementos de la diagonal corresponden a los P primeros valores singulares. la matriz $\mathbf{X}_{(I\times JK)}$ es la matriz de datos $\mathbf{X}_a$ generada por matriciación del arreglo tridimensional por la vía A.

Como se menciono anteriormente, el algoritmo necesita de un mecanismo de iniciación de las matrices $\mathbf{B}$ y $\mathbf{C}$ para obtener la matriz $\mathbf{A}$ estimada. Por su parte, las matrices $\mathbf{B}$ y $\mathbf{C}$ con las que se inicia corresponden a los vectores singulares izquierdos de la descomposición en valores singulares de la matriciación del arreglo por la vía B ($\mathbf{X}_{(J\times IK)}$) y C ($\mathbf{X}_{(K\times IJ)}$), reteniendo los primeros Q y R vectores singulares izquierdos, respectivamente.

Finalmente, el número de vectores singulares retenidos para la solución P, Q y R se seleccionan considerando la regla del producto mínimo que presenta Kroonenberg (2008):

$$P \times Q \geq R, \quad P \times R \geq Q, \quad Q \times R \geq P$$

2.5.7. Función de pérdida

En el modelo Tucker3 se ajusta en el sentido de mínimos cuadrados alternantes, como la suma de cuadrados de las aproximaciones del modelos a los datos; por tanto, minimizando:

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(x_{ijk} - \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} g_{pqr}(a_{ip}b_{jq}c_{kr}))^2$$

Sobre $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ y $\underline{\mathbf{G}}$.

Ahora, considerando la matriciación sobre las vías $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$, la función de pérdida toma la siguiente forma:

$$||\mathbf{X}_a - \mathbf{A}\mathbf{G}_a(\mathbf{C}' \otimes \mathbf{B}')||^2$$

$$||\mathbf{X}_b - \mathbf{B}\mathbf{G}_b(\mathbf{C}' \otimes \mathbf{A}')||^2$$

$$||\mathbf{X}_c - \mathbf{C}\mathbf{G}_c(\mathbf{B}' \otimes \mathbf{A}')||^2$$

Como se sabe, $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ son matrices ortonormales y $\mathbf{G}_a$, $\mathbf{G}_b$ y $\mathbf{G}_c$ representan las matriaciones del arreglo central por las vías $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$, respectivamente (Bautista Mendoza, 2009).

2.5.8. Estabilidad de la estimación parámetros

Al igual que en los modelos de dependencia, una forma de evaluar el ajuste general es calcular la variabilidad explicada para destacar si el ajuste del modelo es aceptable. En el modelo Tucker3, el ajuste general se puede dividir en cuatro maneras: (1) por componente ortonormal, (2) por nivel de cada modo, (3) por elemento de la matriz central y (4) por punto de datos individual.

Debido a que no existe un modelo estocástico para múltiples vías, el único camino para establecer la estabilidad de las estimaciones es tomar muestras de arranque de los residuos, aunque ello presuponga la validez del conjunto de datos de tres vías[14]. En este contexto, el bootstrap emerge como una herramienta esencial en el análisis de diversas vías. Esta técnica se sustenta en la simulación de muestreo aleatorio con reposición de la muestra original de datos, generando así múltiples conjuntos de datos simulados. Estos conjuntos se utilizan para calcular diversos estimadores, permitiendo medir la variabilidad o incertidumbre en los resultados. En otras palabras, el bootstrap imita el proceso de muestreo que originó la muestra de datos real.

En términos fundamentales, el principio rector del procedimiento bootstrap es que "la distribución observada es la mejor estimación de la distribución de la población". Siguiendo este principio, resulta evidente que el muestreo de la distribución observada es la mejor aproximación posible, a excepción del muestreo directo de la distribución de la población misma.

Por ende, para obtener una estimación de un parámetro de población, se realiza un muestreo repetido con reposición de la muestra observada, asignando igual probabilidad a todos los puntos de datos. Para cada una de estas muestras, se calcula el valor

[14]En los modelos de múltiples vías, generalmente no se realizan suposiciones respecto a la distribución de los datos, por lo que los errores no se calculan de la misma manera; de hecho, Kroonenberg (2008) considera que ni siquiera se sabe si la teoría de las distribuciones serían apropiadas

del estimador del parámetro, y la media de estas estimaciones se convierte en la mejor estimación del parámetro de la población. La desviación estándar de estas estimaciones representa su error estándar

De acuerdo a ello, la calidad del análisis de arranque puede ser evaluada a través de métricas como la precisión o el error, que miden el grado de similitud entre la matriz original y la matriz de Tucker3 estimada. En particular, las matrices de componentes de cada uno de los análisis de arranque deben transformarse de tal manera que sus orientaciones sean óptimamente iguales antes de que se puedan calcular los errores estándar de arranque.

La orientación óptima en el modelo Tucker3 se refiere a la dirección en la que los datos deben ser transformados para que sean mejor representados por los tres factores de la descomposición. Esta dirección óptima se determina mediante el cálculo de la matriz de covarianza de los datos y el uso de técnicas de optimización. El objetivo es encontrar la orientación que maximiza la varianza explicada por los factores, lo que mejora la capacidad del modelo para representar los datos y mejorar la precisión de las estimaciones.

De acuerdo a Kiers (2004), se pueden encontrar dos procedimiento para encontrar la orientación óptima para las soluciones de arranque, usando para ello transformación a una solución objetivo: (1) un procedimiento basado en optimizar las congruencias de los componentes de las soluciones de arranque a la forma básica del original de tres vías, y (2) un procedimiento basado en transformaciones o rotaciones hacia las matrices componentes originales. De acuerdo a Kroonenberg (2008), ambas propuestas arrojan intervalos que son bastante buenas aproximaciones a los intervalos de confianza del 95 % en la mayoría de los casos. Por otro lado, una forma de validar los resultados del análisis de tres vías es mediante el uso de variables externas. En especial si se cuenta con información complementaria a la original. La idea detrás de esta validación externa es evaluar el desempeño del modelo en datos desconocidos para determinar su capacidad de generalización y su habilidad para predecir valores futuros.

El proceso de validación externa es importante para garantizar que el modelo funcione correctamente y sea confiable en aplicaciones prácticas; en ese sentido, el objetivo de la validación externa es evaluar la generalización del modelo a nuevos conjuntos de datos. Si el modelo tiene un buen desempeño en la validación externa, es probable

que tenga un buen desempeño en nuevos conjuntos de datos similares. Sin embargo, si el modelo tiene un mal desempeño en la validación externa, es probable que tenga un mal desempeño en nuevos conjuntos de datos y, por lo tanto, no sea útil para la predicción o la toma de decisiones. Como tal, el proceso incluye los siguientes pasos, tal como lo describe Kiers (2004):

- Selección de un conjunto de datos independiente para la validación externa.
- Aplicación del modelo a los datos de validación externa para predecir los valores de la variable de interés.
- Cálculo del error o desviación entre los valores previstos y los valores reales en los datos de validación externa.
- Análisis y comparación del error con otros modelos o con una aproximación estadística simple, como la media o la mediana, para evaluar la capacidad predictiva del modelo.

2.5.9. Inercia

La inercia es una forma de medir la importancia relativa de cada componente en la representación de los datos. En un análisis de componentes principales, se suele utilizar la inercia para decidir cuántos componentes se deben incluir en el modelo final. Por ejemplo, se pueden incluir solo aquellos componentes que expliquen un porcentaje significativo de la varianza en los datos. En ese sentido, la inercia retenida o variabilidad del arreglo original explicado por el modelo Tucker3 corresponde al porcentaje de ajuste del modelo, definido en términos de los elementos del arreglos central:

$$\frac{g_{pqr}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2}$$

Por otro lado, el modelo Tucker3 genera la contribución de los factores o componentes que explican cada una delas vías y la contribución de la combinación de componentes de las diferentes vías. De ello se tiene que la contribución de los componentes extraídos por una de las vías se obtiene considerando la suma de cuadrados de los elementos

del arreglo central relacionados con ese componente sobre la suma de cuadrados totales. De esta manera, la contribución de los componentes de la vía **A** se expresa como:

$$\frac{\sum_{q=1}^{Q}\sum_{r=1}^{R} g_{qr}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2}$$

Para los componentes de la vía **B**:

$$\frac{\sum_{p=1}^{P}\sum_{r=1}^{R} g_{pr}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2}$$

y finalmente, para la vía **C**:

$$\frac{\sum_{p=1}^{P}\sum_{q=1}^{Q} g_{pq}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2}$$

La proporción de variabilidad explicada por cada combinación de componentes se obtiene tomando el cuadrado del valor central g_{pqr} que identifica la combinación del componente p de la vía **A**, el componente q de la vía **B** y el componente r de la vía **C**, dividido por la suma de cuadrados de todos los valores, lo cual está dado por:

$$\frac{g_{pqr}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2}$$

Como Tucker3 extrae tres sistemas de ejes, se requiere generar la contribución de la combinación de componentes a la inercia total, a partir del arreglo central; esto le permite identificar e interpretar las combinaciones de los componentes más influyentes (Bautista Mendoza, 2009).

Criterio dif-fit de Timmerman

Conforme a la recomendación de Timmerman and Kiers (2000), en el contexto del modelo Tucker3, el criterio se emplea para evaluar el ajuste incremental al comparar modelos de descomposición de datos multidimensionales. En este sentido, se selecciona el modelo con la menor Suma de Cuadrados del Error (SCE).

Para comparar modelos con diferente número de componentes ($S = P + Q + R$), se introduce el criterio Dif-Dif, definido en términos de la SCE como el cociente:

$$b_s = \frac{dif_S}{dif_{S*}}$$

El cálculo de este criterio se realiza siguiendo el siguiente procedimiento:

- Se define la SCE de un modelo con S componentes como V_S y la SCE de un modelo con $S-1$ componentes como V_{S-1}.
- Se obtiene $dif_S = V_{S-1} - V_S$.
- Se calcula $b_S = \frac{dif_S}{dif_{S*}}$; aquí, dif_{S*} representa la mayor diferencia encontrada después de dif_S.
- Se selecciona el modelo S con el valor más bajo de b_S.

Ahora, la visualización de la evolución del dif-fit se denomina *Scree plot multiway*, el cual Timmerman considera como una versión multivariante de la versión de Cattell, el gráfico de sedimentación.

Scree plot multiway

El *scree plot multiway* es un gráfico utilizado para visualizar la inercia de un modelo de Tucker3. El gráfico muestra la cantidad de varianza en los datos explicada por cada componente del modelo. En un *scree plot*, se representa en el eje Y la inercia o la cantidad de varianza explicada por cada componente y en el eje X se representa el número de componentes. Con un *scree plot*, es fácil identificar el punto en el que la inercia disminuye significativamente, lo que indica que los componentes adicionales no explican una gran cantidad de varianza en los datos y, por lo tanto, no son necesarios para una buena representación de los datos, tal como se observa en la figura (2.11).

2.5.10. Casos particulares del modelo Tucker 3

En esta subsección se especificarán dos modelos particulares derivados del modelo Tucker3: Tucker 1 y Tucker2. Estos modelos pueden considerarse casos extremos del modelo original, al restringir a una o dos vías; es decir, existe la factibilidad de realizar

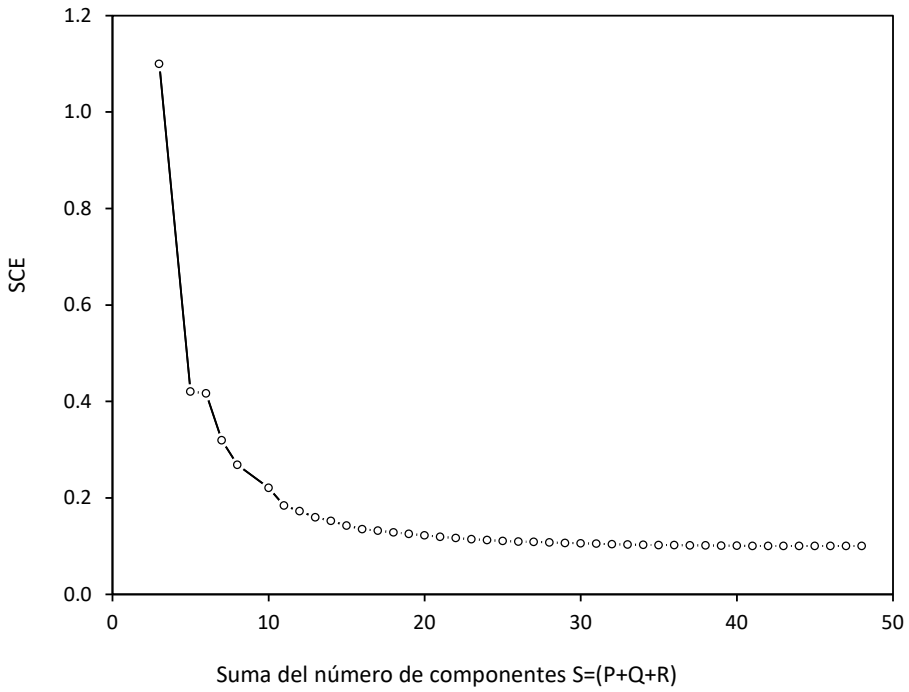


Figura 2.11: Scree plot multiway

el análisis sobre solo una o dos vías de los datos. En ese sentido, partiendo del modelo siguiente:

$$\mathbf{X_k} = \mathbf{A_k}\mathbf{G_k}\mathbf{B_k'} + \mathbf{E_k} \tag{2.24}$$

el cual parte de la ecuación (2.24) que, finalmente es una descomposición en valores singulares para cada *slice* frontal del arreglo de tres vías. Donde $\mathbf{A_k}$, $\mathbf{B_k}$, $\mathbf{G_k}$ y $\mathbf{E_k}$ son los vectores izquierdos, vectores derechos, la matriz central y la matriz de residuos, respectivamente, definidos en cada *slice* frontal k.

Modelo Tucker1

Respecto a la expresión (2.24), una restricción natural es exigir que las matrices componentes para una de las vías sean las mismas: $\mathbf{A_k} = \mathbf{A}$ o $\mathbf{B_k} = \mathbf{B}$. En el primer caso ($\mathbf{A_k} = \mathbf{A}$), las diferencias individuales son constantes en todas las ocasiones, k. De este modo, se supone que la matriz de puntuación (*scores*) es común en todas las

condiciones. Esto deriva a la siguiente expresión [15]:

$$\mathbf{X_k} = \mathbf{AG_kB'_k} + \mathbf{E_k} \tag{2.25}$$

En el segundo, se asume que la matriz de cargas (*loading*) es común para todos los k, de manera que $\mathbf{B_k} = \mathbf{B}$, ello conduce a la siguiente expresión:

$$\mathbf{X_k} = \mathbf{A_kG_kB'} + \mathbf{E_k} \tag{2.26}$$

Tanto (2.25) y (2.26) son los modelos Tucker1, en el que el primero se refiere al número de formas en que los componentes se determinan de forma independiente. En ese sentido, los modelos pueden ser resueltos mediante PCA en la matriz de tres vías matriciada en dos vías (Kroonenberg, 2008).

De la misma manera, el posible considerar un modelo en que los pesos relativos de los componentes son idénticos; es decir, $\mathbf{G_k} = \mathbf{G}$. Esta situación es la clásica descomposición en valores singulares estándar[16], pero esta vez con la misma descomposición para cada uno de los k *slices* frontales. Si se considera tomar el promedio a través de la tercera vía, esto lleva a los siguientes resultados:

$$\mathbf{Z_k} = \mathbf{AGB'} + \mathbf{E_k} = \mathbf{AF'} + \mathbf{E_k} \tag{2.27}$$

De acuerdo a Kroonenberg (2008), este modelo pudó haber sido llamado modelo Tucker0, pero Kiers (1988) lo denominó SumPcA. Lo anterior puede verse también como:

$$z_{ijk} = \sum_{p=1}^{s} g_{ip}c_{pjk}$$

donde $i = 1, ..., I$, $j = 1, ...J$ y $k = 1, ..., K$, y con la restricción sobre $\mathbf{G}$, el modelo se convierte en:

$$\mathbf{Z} = \mathbf{GC}$$

donde $\mathbf{Z}$ y $\mathbf{C}$ son matrices de tres vías escritas como combinaciones de las mismas.

[15] Si se considera la notación $\mathbf{X^{(I\times JK)}} = \mathbf{AG^{(P\times QR)}}(\mathbf{C}\otimes\mathbf{B})' + \mathbf{E^{(I\times JK)}}$, el modelo Tucker1 supone que: $\mathbf{B} = I_k$ y $\mathbf{C} = \mathbf{I_k}$. Esto define al modelo $\mathbf{X^{(I\times JK)}} = \mathbf{AG^{(P\times JK)}} + \mathbf{E}$, que corresponde a PCA en una versión oportunamente desplegada de $\underline{\mathbf{X}}$ (Amigo and Marini, 2013).

[16] Los modelos de Tucker1 tienen las mismas propiedades técnicas que la descomposición en valores singulares, pero lo que es realmente diferente es que no hay solo uno sino k términos de error $\mathbf{E_k}(Kroonenberg,\ 2008)$.

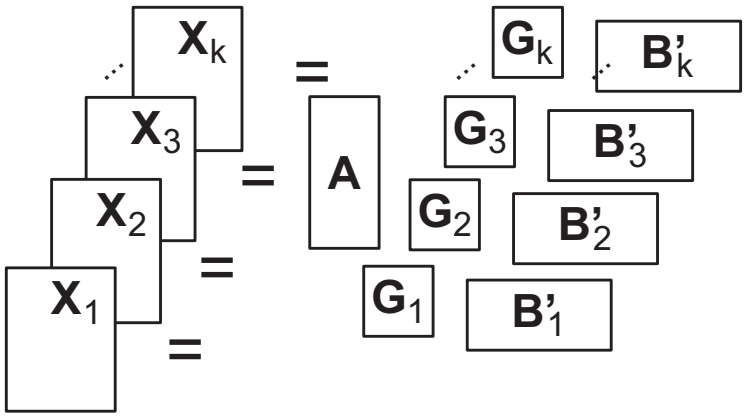


Figura 2.12: Modelo Tucker1 con $\mathbf{A_k} = \mathbf{A}$, pero diferentes $\mathbf{B_k}$

Nota. Tomado de (Kroonenberg, 2008, 52).

Modelo Tucker2

Al restringir de manera simultanea $\mathbf{A_k} = \mathbf{A}$ y $\mathbf{B_k} = \mathbf{B}$ (ver figura 2.13), mientras se mantiene la ortonormalidad de los componentes, se provoca una ruptura de una de las propiedades de la descomposición en valores singulares, en particular, la vinculación única de las columnas de $\mathbf{A}$ y $\mathbf{B}$. Esto deriva a que $\mathbf{G_k}$ no sea necesariamente diagonal, pero todos los elementos de este, g_{pqk}, pueden ser diferentes de cero.
Si se denomina $\mathbf{H_k^{(P\times Q)}} = \mathbf{G_k}$, la derivación lleva a que los números de columnas de $\mathbf{A}$ y $\mathbf{B}$ no son necesariamente iguales, de manera que el número de componentes de $\mathbf{A}$ estará determinado por P y de $\mathbf{B}$ por Q. En resumen, se obtiene la siguiente expresión:

$$\mathbf{X_k} = \mathbf{A}\mathbf{H_k}\mathbf{B}' + \mathbf{E_k} \tag{2.28}$$

donde $\mathbf{H_k}$ representa la fuerza relativa de los vínculos entre las columnas de $\mathbf{A}$ y $\mathbf{B}$. Por ejemplo, si la matriz de datos de ocasión k , $\mathbf{X_k}$, se recupera aproximadamente por la suma de los productos de la segunda columna de $\mathbf{A}$ y la quinta columna de $\mathbf{B}$, y la tercera columna de $\mathbf{A}$ y la primera columna de $\mathbf{B}$, entonces h_{25k} y h_{31k} son distintos de cero y los elementos restantes de $\mathbf{H_k}$ son cero. En notación de sumatoria, la ecuación (2.28) se expresa como:

$$x_{ijk} = \sum_{q=1}^{P}\sum_{q=1}^{Q} h_{pqh}(a_{ip}b_{jq}) + e_{ijk}$$

En su forma básica, la matriz $\mathbf{A}$ es ortonormal de $I \times P$ columnas, $\mathbf{B}$ es una matriz ortonormal de $J \times Q$ columnas, donde no necesariamente $P = Q$, por su parte, $\mathbf{H_k}$ es

denominada el *slice* frontal de la matriz central extendida, $\underline{\mathbf{H}}$ que es una matriz de tres vías de dimensión $P \times Q \times K$. Dado que $\mathbf{A}$ y $\mathbf{B}$ son ortonormales, el cuadrado de cada elemento h_{pqk} expresa el tamaño de la variabilidad, que se explica por la combinación de la p-ésima columna de $\mathbf{A}$ y la q-ésima columna de $\mathbf{B}$.

El modelo Tucker2 presenta algunas propiedades de la descomposición en valores singulares; en particular,permite una descomposición completa de la matriz de tres vías falibles. Asimismo, dadas las componentes ortonormales, es posible realizar una partición completa de la variabilidad debida a las componentes. Esta propiedad se extiende a la partición de la variabilidad por nivel de cada vía y por elemento de la matriz central (Kroonenberg, 2008). El modelo Tucker2 tiene diversas aplicaciones, entre las cuales se destacan tres. En primer lugar, se puede utilizar en aquellos análisis de datos para los que no se puede definir una condensación natural de la tercera vía, como por ejemplo el análisis multivariado de series de tiempo. Esto porque en general, no se puede atribuir ningún significado útil a los componentes de una vía de tiempo (Kroonenberg and De Leeuw, 1980). En segundo lugar, el modelo se puede aplicar en escalas de diferencias individuales con matrices de similitud asimétricas; aunque esto es aplicable, también, con Tucker3. Finalmente, el modelo Tucker2 se puede utilizar para probar la idoneidad de varios modelos de diferencias individuales en el escalado multidimensional (Kroonenberg and De Leeuw, 1980; Kroonenberg, 2008).

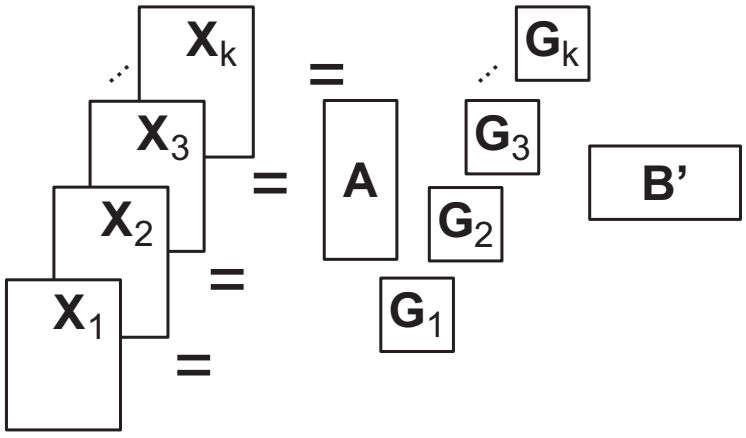


Figura 2.13: Modelo Tucker1 con $\mathbf{A_k} = \mathbf{A}$ y $\mathbf{B_k} = \mathbf{B}$

Nota. Tomado de (Kroonenberg, 2008, 54).

2.5.11. Algoritmo Tucker3

El algoritmo utilizado para la estimación de los parámetros y componentes en las tres vías del modelo se fundamenta en el método de mínimos cuadrados alternantes. En este enfoque, se emplea la descomposición en valores singulares de diferentes arreglos matriciales, ejecutando el procedimiento iterativamente hasta alcanzar un nivel de convergencia predefinido. Diversos autores, como Andersson and Bro (2000), Kroonenberg (1983), Amaya and Pacheco (2002) y Bautista Mendoza (2009), han desarrollado código para este procedimiento y el procesamiento de datos, utilizando varios lenguajes como MATLAB, TUCKALS3, SAS y R. Todos ellos establecen criterios diversos para la convergencia del algoritmo.

En el presente estudio, se basa en el trabajo de Amaya and Pacheco (2002) y Bautista Mendoza (2009), quienes proporcionan la base de la programación. La estructura tridimensional del modelo Tucker3 se expresa como sigue:

$$\mathbf{X}_a = \mathbf{A}\mathbf{G}_a(\mathbf{C}^T \otimes \mathbf{B}^T) + \mathbf{E}$$

Cada componente de esta ecuación ha sido previamente descrito y explicado en la sección metodológica. Con base en lo anterior, el algoritmo comienza estableciendo un criterio de convergencia, y luego sigue el siguiente procedimiento: Iniciar con las matrices $\mathbf{B}$ y $\mathbf{C}$

1. $[\mathbf{A}, \mathbf{S}, \mathbf{V}] = svd(\mathbf{X}_a(\mathbf{C} \otimes \mathbf{B}), P)$
2. $[\mathbf{B}, \mathbf{S}, \mathbf{V}] = svd(\mathbf{X}_b(\mathbf{C} \otimes \mathbf{A}), Q)$
3. $[\mathbf{C}, \mathbf{S}, \mathbf{V}] = svd(\mathbf{X}_c(\mathbf{B} \otimes \mathbf{A}), R)$
4. Ir al paso 1 hasta llegar a la convergencia
5. $\mathbf{G}_a = \mathbf{A}^T\mathbf{X}_a(\mathbf{C} \otimes \mathbf{B})$

Donde $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = svd(\mathbf{X}, P)$ es la descomposición en valores singulares (SVD) reducida de $\mathbf{X}$ reteniendo los primeros P componentes. La matriz $\mathbf{X}$ es descompuesta como $\mathbf{X} = \mathbf{USV^T}$, donde la matriz $\mathbf{U}$ contiene los primeros P vectores singulares izquierdos, $\mathbf{V}$ los primeros P vectores singulares derechos, y donde $\mathbf{S}$ es un matriz diagonal cuyos elementos diagonales corresponden a la raíz cuadrada de los primeros

P valores singulares. Además, a representa un arreglo $I \times JK$, $b = J \times IK$ y $c = K \times IJ$. Como lo describen Amaya and Pacheco (2002):

> "Luego de haber determinado el criterio de convergencia del algoritmo, se procede a generar las estimaciones iniciales de las matrices **B** y **C**. La estimación de la matriz componente **B** se obtiene al seleccionar los Q primeros vectores singulares izquierdos de la SVD de $\mathbf{X}_b$, y la estimación inicial de la matriz componente **C** corresponde a los R primeros vectores singulares izquierdos de la SVD de $\mathbf{X}_c$. Con base en estas estimaciones iniciales, se da paso al proceso iterativo aplicando SVD sobre los arreglos matriciales $\mathbf{X}_a(\mathbf{C} \otimes \mathbf{B})$, $\mathbf{X}_b(\mathbf{C} \otimes \mathbf{A})$ y $\mathbf{X}_c(\mathbf{B} \otimes \mathbf{A})$ para obtener las estimaciones de las tres matrices componentes al seleccionar respectivamente los P, Q y R vectores singulares izquierdos de la descomposición. Adicionalmente, con base en las tres matrices componentes estimadas, se obtiene el arreglo central en alguna de sus formas de despliegue"(Bro (1998) referenciado por Amaya and Pacheco, 2002, 48).

Además, como lo considera Bautista Mendoza (2009), el número de vectores singulares retenidos para la solución P, Q y R se seleccionan considerando la siguientes expresiones:

$$P \times Q \geq R, \quad P \times R \geq Q, \quad Q \times R \geq P$$

Capítulo 3

Metodología

3.1. Tipo de investigación

Esta investigación es de tipo correlacional, pues analiza la relación entre diversas variables económicas sin establecer causalidad. Se recopila información de múltiples variables en un conjunto de departamentos y se emplea la correlación para identificar relaciones subyacentes.

El modelo Tucker3 es una técnica multivariante que reduce la dimensionalidad en estudios de tres vías, permitiendo el análisis simultáneo de múltiples variables. Como generalización de métodos como la descomposición en valores singulares, descompone los datos en tres matrices que reflejan la estructura dimensional. Su ortogonalidad garantiza la máxima captura de variabilidad, convirtiéndolo en una extensión del análisis factorial y de componentes principales.

El análisis considera un arreglo tridimensional compuesto por departamentos, variables macroeconómicas y tiempo. El modelo Tucker3 permite reducir la dimensionalidad e identificar patrones en estas dimensiones, facilitando el estudio de la evolución temporal y geográfica de las actividades económicas.

Dado que Tucker3 no tiene una solución exacta, se emplea el método de mínimos cuadrados alternantes con descomposición en valores singulares. La implementación se realiza en *R*, siguiendo las metodologías de Bautista Mendoza (2009), Amaya and Pacheco (2002) y Giordani et al. (2014).

3.2. Muestra y variables

El modelo Tucker3 es una técnica multivariante utilizada en esta investigación para identificar patrones y estructuras subyacentes en la dinámica económica del Perú. Se emplea un arreglo de tres vías compuesto por:

- **Individuos (vía A):** Representados por los departamentos del país (24 departamentos y 1 provincia constitucional), con Lima dividida en Lima Metropolitana y Lima Provincias, totalizando 27 unidades de análisis.
- **Variables macroeconómicas (vía B):** Se consideran 13 indicadores recopilados por el Instituto Nacional de Estadística e Informática (INEI), incluyendo el Producto Bruto Interno y el Valor Agregado Bruto de diversos sectores económicos.
- **Tiempo (vía C):** Comprendido entre los años 2007 y 2021, lo que abarca 14 periodos de observación.

El arreglo de tres vías $\underline{\mathbf{X}}$ tiene dimensiones $27 \times 13 \times 14$, donde x_{ijk} representa el valor de la variable j en el departamento i en el año k.

En resumen, la muestra de esta investigación está conformada por indicadores macroeconómicos departamentales en el Perú, analizados en un marco temporal de 14 años para identificar relaciones y patrones en la economía regional.

3.3. Obtención de datos

Los datos utilizados en este estudio provienen del *Sistema de Información Regional para la Toma de Decisiones* (SIRTOD) del *Instituto Nacional de Estadística e Informática* (INEI), entidad técnica que rige el Sistema Estadístico Nacional. El INEI es responsable de normar, planificar, dirigir y supervisar las actividades estadísticas oficiales en el país, asegurando su integración y coherencia a través de normativas técnicas comunes.

Como ente rector, el INEI mide y proyecta hechos económicos y sociales, produciendo estadísticas oficiales a través de censos, encuestas, indicadores e índices, cuentas nacionales y regionales, entre otros. Además, proporciona información pública conforme a la Ley 27806 de Transparencia y Acceso a la Información. Entre los principales sistemas de consulta del INEI destacan:

- **REDATAM:** base de datos estadísticos.
- **Sistema de estadísticas de criminalidad y seguridad ciudadana.**
- **SIRTOD:** información regional para la toma de decisiones.
- **ODS-INEI:** monitoreo de indicadores de desarrollo sostenible.
- **SIGE:** información geográfica para emprendedores.
- **SISPEMA:** evaluación de actividades estadísticas.
- **PENDES:** monitoreo del plan estratégico nacional para el desarrollo estadístico.

3.4. Procesamiento de los datos

En esta sección, se describirá cómo se aplicó el modelo Tucker3 a los datos recolectados para investigar para realizar un análisis factorial dinámico y con ello evaluar la dinámica macroeconómica

Los datos utilizados en este estudio provienen del INEI. La muestra incluye 24 departamentos con 13 variables económicas y 14 años (ocasiones). Se llevaron a cabo varios procesos de limpieza de datos para asegurar la calidad de los datos, incluyendo la eliminación de registros duplicados, la corrección de errores de entrada y la eliminación de valores atípicos. También se llevó a cabo una revisión manual de los datos para asegurar la integridad de los mismos.

Para garantizar que ninguna variable tuviera un impacto desproporcionado en el análisis, los datos se normalizaron usando el método descrito por Kiers (2000). La normalización se llevó a cabo mediante la transformación de cada variable a una escala común. Luego, se aplicó el modelo Tucker3 a los datos normalizados de cada variable utilizando R versión 4.2.2. El modelo se ajustó a los datos mediante un proceso iterativo de optimización de la función de pérdida de acuerdo a los algoritmos usados por Bautista Mendoza (2009) y Bro (1998).

Los resultados del análisis incluyen las matrices de tres vías y la matriz central. Se presentaron gráficos y tablas que ilustran la estructura subyacente de los datos, así como la importancia relativa de cada variable en la explicación de la estructura. Estos gráficos incluyen una matriz de correlación y un gráfico de barras de la importancia de las variables. Los resultados del análisis con el modelo Tucker3 proporcionaron una comprensión detallada de la estructura de los datos y permitieron aplicar de manera efectiva en la estructura dinámica. Esto es útil para analizar la dinámica macroeconómica del país dentro de un conjunto de ocasiones que, en la presente investigación, abarcaron entre 2007 y 2021.

3.5. Técnicas de recolección de datos

La presente investigación hace uso exhaustivo de datos secundarios, de manera que esta técnica de recolección de datos implica el uso de datos previamente recopilados por el INEI. Estos datos provienen de fuente pública del Sistema Nacional de Estadística, específicamente del Sistema Nacional Regional para la Toma de Decisiones (SIRTOD). Como el objetivo está dirigido a utilizar el modelo Tucker3 para analizar la dinámica económica del Perú entre 2007 y 2021. Los datos económicos usados son recopilados por el INEI, que publica periódicamente los datos de las cuentas nacionales del país a través del SIRTOD. Es importante destacar que el uso de los datos secundarios es una opción eficiente y práctica, pues permite ahorrar tiempo y recursos en comparación a otras técnicas. Asimismo, como los datos secundarios son de carácter público, estos gozan de rigurosidad en su publicación, por lo que es posible afirmar que no existen sesgos por parte de la fuente pública.

En resumen, el uso de los datos secundarios para aplicarlos al modelo Tucker3 es una herramienta muy útil para el estudio de la dinámica económica de un país. La combinación de la revisión y selección de información relevante y la aplicación del modelo Tucker3 permite una comprensión profunda y detallada de la dinámica económica y ayuda a identificar patrones y tendencias relevantes para la investigación.

Capítulo 4

Dinámica macroeconómica en el Perú, 2007-2021

La dinámica económica es un proceso complejo y en constante cambio que abarca transformaciones en la economía, incluyendo producción, consumo, distribución de riqueza y crecimiento económico. Esto implica factores como oferta y demanda, inflación, comercio, inversión y políticas gubernamentales. Existe interdependencia entre variables y sectores, como la inversión que afecta a la demanda, producción y el crecimiento. Incertidumbres externas como crisis políticas y fluctuaciones también influyen. La tecnología y políticas gubernamentales constantemente cambian el panorama. Comprender esta dinámica es crucial para tomar decisiones sobre inversión, carrera y gastos. Eventos económicos impactan a nivel general, así que conocer la dinámica ayuda a tomar decisiones informadas. También permite evaluar sectores económicos y entender conexiones globales. Políticas públicas influyen en la dinámica, afectando empleo, salarios y calidad de vida. Además, la economía global interconectada hace vital entender cómo dinámicas locales afectan a otras regiones y países.

En especial, entender la dinámica económica puede ayudar a comprender las políticas públicas por diversas maneras, aquí algunas de estas razones:

- Comprender los objetivos y prioridades del gobierno: Las políticas gubernamentales son a menudo un reflejo de los objetivos y prioridades del gobierno, por ejemplo, la creación de empleo, el crecimiento económico o la estabilidad financiera. Conociendo la dinámica económica, es posible entender mejor qué está motivando a los políticos a tomar determinadas medidas.

- Evaluar la efectividad de las políticas gubernamentales: Conociendo la dinámica económica, es posible evaluar la efectividad de las políticas gubernamentales y determinar si están logrando sus objetivos. Por ejemplo, si una política gubernamental está destinada a estimular el crecimiento económico, es posible evaluar si está funcionando observando el crecimiento económico.

- Participación activa en la democracia: Conociendo la dinámica económica, es posible tener una opinión informada sobre las políticas gubernamentales y participar activamente en democracia. Por ejemplo, es posible votar o participar en debates políticos sobre las políticas económicas.

- Comprender las implicaciones a largo plazo: Las políticas gubernamentales a menudo tienen implicaciones a largo plazo, y es importante entender cómo pueden afectar la dinámica económica a largo plazo. Conociendo la dinámica económica, es posible evaluar las implicaciones a largo plazo de las políticas gubernamentales

La dinámica macroeconómica es una rama de la economía que se centra en el estudio de los comportamientos y fluctuaciones de las grandes variables económicas a nivel agregado, como el producto, el empleo, los precios, la inversión, el gasto público y el comercio internacional. Se enfoca en entender cómo estas variables interactúan entre sí, cómo se afectan mutuamente y cómo influyen en el crecimiento económico y la estabilidad. Se utiliza para desarrollar teorías y modelos que expliquen el comportamiento de la economía en su conjunto y para diseñar políticas económicas efectivas. La dinámica macroeconómica está estrechamente relacionada con la composición sectorial de una economía, que refiere a la distribución de la producción y el empleo entre los diferentes sectores, como la agricultura, la industria y los servicios. Los cambios en la producción, el empleo, los precios y la demanda agregada pueden afectar la composición sectorial. Por lo tanto, es importante considerarla al analizar la dinámica económica de un país.

La relación entre la dinámica macroeconómica y la composición sectorial está influenciada por muchos elementos subyacentes, incluyendo la evolución de la tecnología, la demanda de los consumidores, la disponibilidad de recursos, la competitividad global, los cambios en las políticas gubernamentales y los ciclos económicos. Estos factores pueden afectar el crecimiento y el desarrollo de los diferentes sectores económicos y,

por lo tanto, influir en la dinámica macroeconómica y la composición sectorial.
Las tendencias de crecimiento son un aspecto importante de la dinámica económica en general, ya que representan la evolución de la economía a lo largo del tiempo. Una tendencia de crecimiento puede ser positiva o negativa, dependiendo de si la economía está expandiéndose o contrayéndose. Las tendencias de crecimiento pueden ser influenciadas por muchos factores, incluyendo la inversión, el gasto gubernamental, la productividad, la demografía y el comercio.
La relación entre la dinámica macroeconómica y las unidades administrativas territoriales puede ser estudiada a través de la comparación de indicadores macroeconómicos a nivel regional. Esto puede ayudar a identificar las diferencias económicas entre regiones y proporcionar información sobre su dinámica económica a lo largo del tiempo. Esta información puede ser utilizada por los tomadores de decisiones para formular políticas que fomenten el crecimiento económico equilibrado en todo el país.
En esta investigación, se estudiará cómo el Producto Bruto Interno y el valor agregado de las actividades económicas del país están altamente correlacionados. En las siguientes secciones se desarrollará su evolución histórica a nivel nacional y departamental.

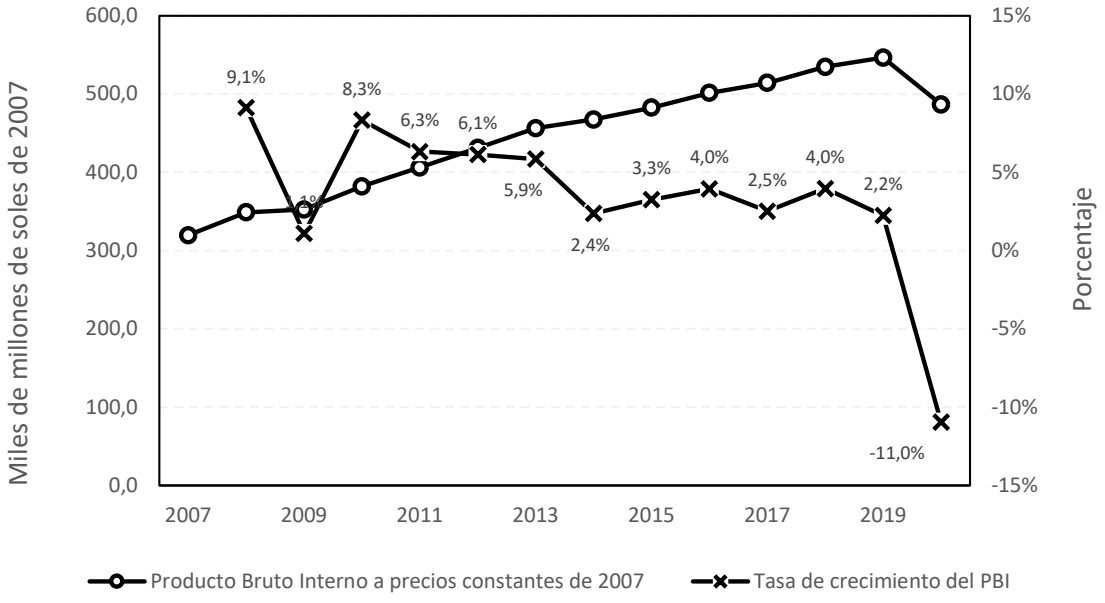
4.1. Análisis de las actividades económicas

El PBI en el Perú ha presentado un perfil de crecimiento que ha constituido uno de los mayores en América Latina y el Caribe durante los últimos años. En 2008, el crecimiento del PBI alcanzó 9,1 %, pasando de 319 mil millones de soles[1] a 349 mil millones de soles. No obstante, esta dinámica se fue desacelerando a medida que se fue cerrando el 2021. En particular, el año 2021 fue sumamente agravada por la crisis sanitaria mundial, que afectó en el crecimiento, presentando un decrecimiento de 11 %, pero se destaca el crecimiento positivo que mantuvo la economía peruana durante el 2007 y 2021. Por lo que, obviando el 2021, el menor nivel de crecimiento fue en 2019, el cual se caracterizó por una afectación de crisis sociales internacionales y nacionales, como por ejemplo los ocurridos en Chile, Bolivia, Venezuela. En el Perú del 2019, la continua crisis política se manifestó en la dinámica del crecimiento económico. Otro factor que afectó fue la guerra comercial entre China y los Estados Unidos, que generó

[1] A precios constante de 2007

un aumento de las tarifas arancelarias y un mayor proteccionismo de ambos países.

Figura 4.1: Producto Bruto Interno a precios constantes de 2007



En 2020, la crisis sanitaria por el coronavirus SARS-CoV-2 impactó notablemente la salud pública, generando un alto impacto en vidas humanas. En el caso de la economía, su efecto fue demoledor. Mientras Perú alcanzó un nivel de PBI de 546,6 mil millones de soles en 2019, en 2020 pasó a ser 486,7 mil millones, un retroceso de 13 puntos porcentuales entre 2019 y 2020. Esto se ve reflejado en la evolución de los diversos sectores económicos, siendo actividades económicas como el turismo la más golpeada durante la crisis.

La transformación de la estructura productiva peruana ha experimentado escasos cambios entre 2007 y 2021. Es posible destacar que ciertas actividades económicas se han intensificado, lo que refleja una tendencia definida en la economía del país. El Valor Agregado Bruto (VAB) se erige como un indicador esencial de la producción a nivel nacional, y específicamente a nivel departamental[2]. De esta manera, la configuración de la estructura productiva se encuentra moldeada por este nivel.

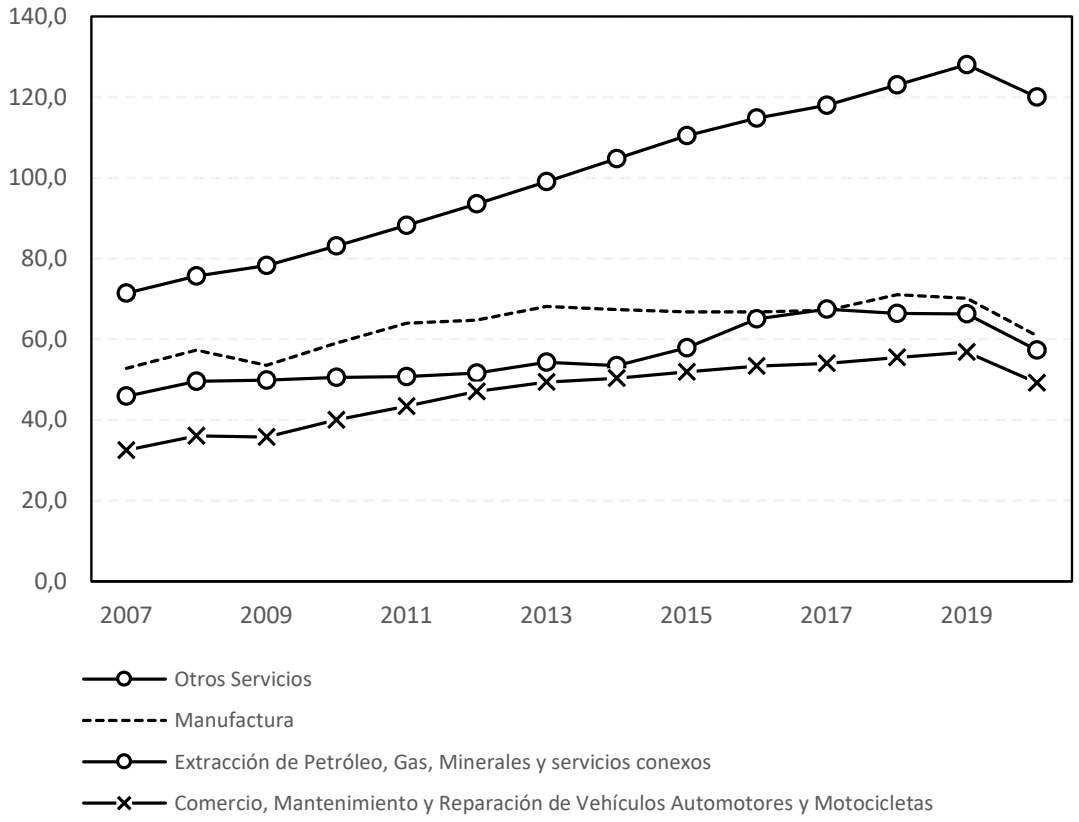
En el Perú, los servicios constituyen uno de los sectores más importantes en función al VAB. En 2007, el VAB generado por los servicios constituyó el 24,4 % del total

[2]El INEI considera el VAB como un indicador sustituto del crecimiento económico a nivel departamental, ya que refleja el nivel de producción en una determinada región geográfica.

generado en ese año, y para el 2020, este porcentaje pasó a ser el 27,1 %; un aumento de 2,7 puntos porcentuales, lo cual se evidencia como una tendencia clara en el tiempo entre este periodo de tiempo, tal como se aprecia en la figura (4.2). Luego, le sigue los demás con una tendencia menos clara, de la cual es posible destacar a la manufactura que, en 2007 represento 52,8 % del VAB total, y para 2020 fue el 13,7 % del total del VAB; es decir, se redujo 4,3 puntos porcentuales, lo cual se evidenciaba como una tendencia decreciente entre 2018 y 2020, pero cuya caída más fuerte fue en 2020 como consecuencia de la crisis sanitaria y sus repercusiones. Otro sector importante es la Extracción de Petróleo, Gas, Minerales y servicios conexos, que en 2020 representó 12,9 % del VAB total, menor en 2,7 puntos porcentuales a su valor en 2007 (ver figura (4.2)).

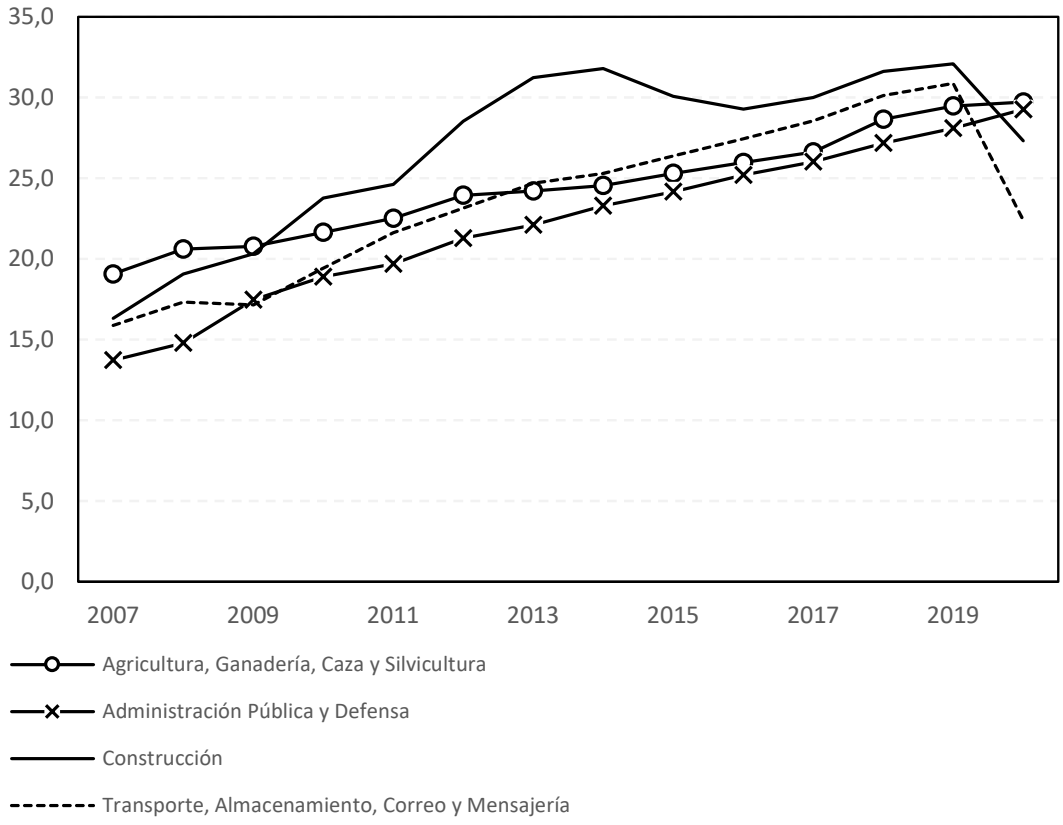
Sectores como el Comercio, Mantenimiento y Reparación de Vehículos Automotores

Figura 4.2: Sectores con mayor Valor Agregado Bruto en la estructura sectorial, 2007-2021 (Porcentaje)



y Motocicletas, Agricultura, Ganadería, Caza y Silvicultura, Construcción y la Electricidad, Gas y Agua no tuvieron variación significativa entre 2007 y 2020, pues se incrementaron en 0 %, 0,2 %, 0,6 % y 0,3 %, respectivamente en la estructura sectorial. Es decir, su incremento en la composición sectorial ha sido poco significativo, pero positivo su crecimiento sobresale. En mayor detalle, se observa que la construcción mantuvo una dinámica de crecimiento muy acelerada, especialmente entre 2011 y 2014; pero entre 2007 y 2020 paso de producir 16,32 miles de millones de soles de VAB a 32,19 miles de millones de soles de VAB, un incremento de alrededor del 95 %, el cual se freno bruscamente en 2020 por el efecto de la pandemia, alcanzando un 27,31 miles de millones de soles de VAB (ver figura (4.3)). Una caída similar se produjo en

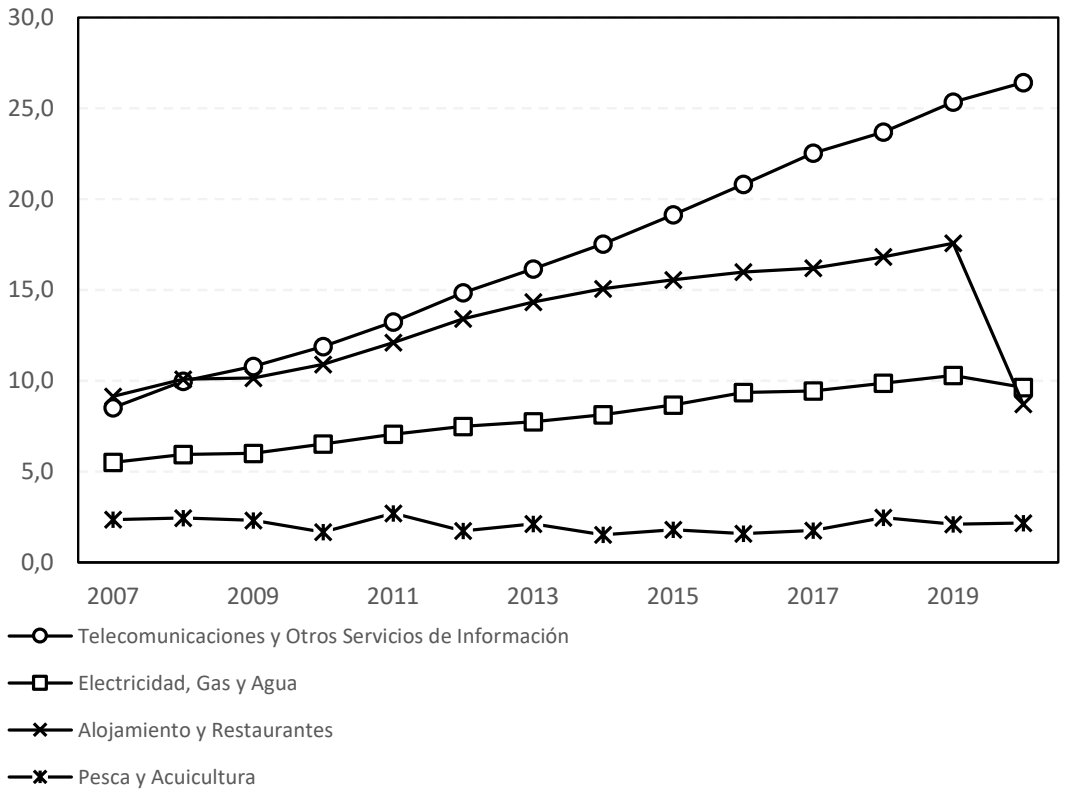
Figura 4.3: Sectores con mayor Valor Agregado Bruto en la estructura sectorial, 2007-2021 (Porcentaje)



el Transporte, Almacenamiento, Correo y Mensajería, que cayó de producir 30,88 % millones de soles de VAB en 2019 a 22,41 % de millones de soles de VAB en 2020, una retroceso de del 27,4 % entre dichos años. Por otro lado, un sector con un gran incremento es el de Telecomunicaciones y Otros Servicios de Información, que en 2007

tuvo un participación de 2,9 % y en 2020 fue de 6 %; es decir, paso de producir 8,52 mil millones de soles en VAB en 2007 a 26,41 mil millones de soles en 2020. Este mantenimiento de su tendencia se debió en gran medida a la pandemia, pues contrario a los demás sectores, el confinamiento incentivo más este sector en 2020 a través del trabajo virtual(ver figura (4.4)). No obstante, quizá el sector más golpeado fue el rubro de Alojamiento y Restaurantes, pues con una dinámica creciente, pasó de producir 9,14 mil millones de soles en VAB en 2007 a 17,57 mil millones de soles en 2019; un incremento de aproximadamente de 93 %, pero cuya caída en 2020 fue de 8,87 mil millones de soles de VAB, es decir en 2020 su generación de producción cayó en 50 %. Como se puede apreciar en el texto anterior, varios sectores de la economía peruana

Figura 4.4: Sectores con mayor menor Agregado Bruto en la estructura sectorial, 2007-2020 (Porcentaje)



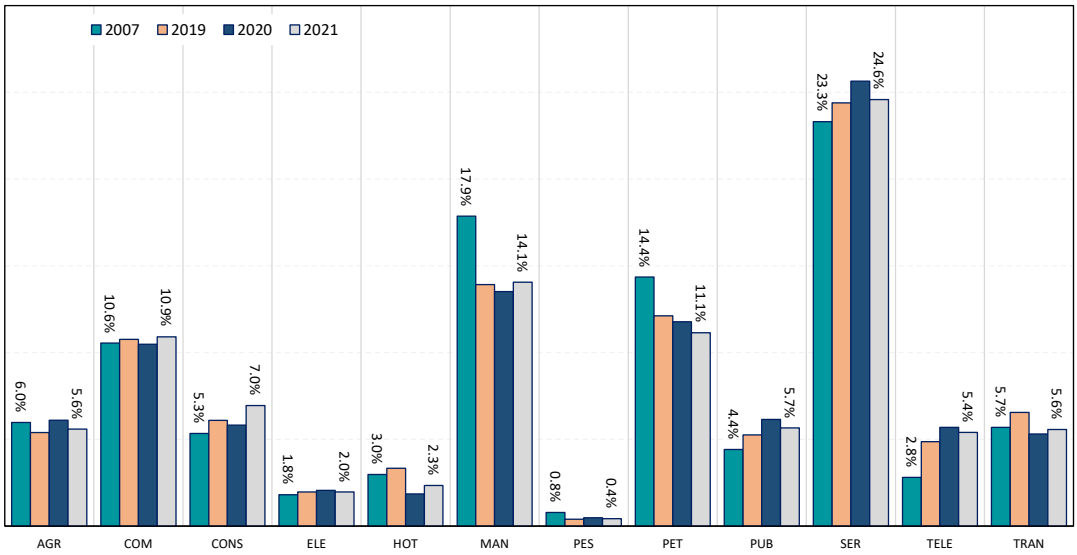
han experimentado cambios en su posición dentro de la estructura productiva. Algunos han ganado terreno, mientras que otros han retrocedido. Por ejemplo, en 2007, la pesca representaba aproximadamente el 0,8 % del valor agregado total, pero para el año 2020, esta cifra había disminuido alrededor del 0,5 %.

Otro punto relevante es la disminución de la importancia de las actividades extractivas,

como la industria petrolera, minera y del gas. En 2007, estas actividades contribuyeron con un 14,4 % al valor agregado, pero para el 2020, esta participación se redujo al 13,5 %. En contraste, la agricultura se ha mantenido constante, representando alrededor del 6 %.

Por otro lado, la manufactura ha experimentado una tendencia a la baja en su contribución a la estructura productiva, pasando del 17,9 % en 2007 al 13,5 % en 2020. En contraposición, los sectores de servicios, telecomunicaciones y construcción han experimentado un crecimiento en su participación dentro del valor agregado total. Lo mencionado puede apreciarse en la figura siguiente:

Figura 4.5: Estructura sectorial, 2007-2021 (Porcentaje)



4.2. Aplicación del modelo

Como se ha expuesto previamente, los datos para la aplicación se estructuran como una matriz de dimensiones $(25 \times 13 \times 15)$. En otras palabras, abarca los valores correspondientes a los 24 departamentos del país ($I = 24$), cada uno con $J = 13$ variables económicas. Estas variables económicas representan los distintos sectores económicos. Además, estos valores están definidos en función de $K = 15$ condiciones, las cuales corresponden a los años comprendidos entre 2007 y 2021.

Los datos en la matriz fueron sometidos a un proceso de estandarización, que incluye el centrado y el escalado. Este procedimiento se llevó a cabo según la metodología detallada en la subsección (2.5.5). En el proceso de centrado, los datos de los individuos se ajustaron utilizando una matriz de dimensiones (24×180). Luego, en la fase de escalado, se aplicó el ajuste sobre la matriz centrada utilizando la dimensión 12×360, correspondiente a la dirección B (variables). Esta etapa de escalado se alinea con la sugerencia propuesta por Bro and Smilde (2003) para el escalado interno de variables.Además, se anticipa que Tucker 3 llevará a cabo un Análisis de Componentes Principales tridimensional, en el cual las diversas ocurrencias de una variable se tratarán como una entidad única.

No obstante, como se analizó en el capítulo anterior y como se destacará más adelante, la inclusión de Lima y Callao tiene un impacto significativo en el ajuste del modelo. Esta influencia se basa en la considerable proporción que representan en el Producto Interno Bruto nacional, lo cual introduce una distorsión en la explicación de la variabilidad. Esto es destacado por Pravdova et al. (2001), quienes muestran que los datos discordantes tienen efectos sobre el algoritmo de mínimos cuadrados alternantes. Por lo tanto, con el fin de abordar esta situación, se ha llevado a cabo una nueva estructuración basada en una matriz de dimensiones $24 \times 13 \times 15$, excluyendo explícitamente a Lima y Callao. A través de esta reconfiguración, se procede a realizar una comparación exhaustiva de los resultados obtenidos de ambas aplicaciones.

4.2.1. Dimensionalidad de la solución

Mediante una iteración que se repitió 100 veces, el algoritmo de Mínimos Cuadrados Alternantes alcanzó un umbral de convergencia de $1e-10$. Este proceso llevó a

determinar la dimensionalidad de la solución, guiada por el gráfico de distribución de valores singulares (Scree plot multiway) ilustrado en la figura 4.6, junto con el criterio *Dif-Fit* propuesto por Timmerman y Kiers y referenciado por Kroonenberg (2008). Los parámetros resultantes de esta elección fueron $P = 2$, $Q = 2$ y $R = 1$ ($S = 5$), lo que generó un ajuste que alcanzó el 99,2 %.

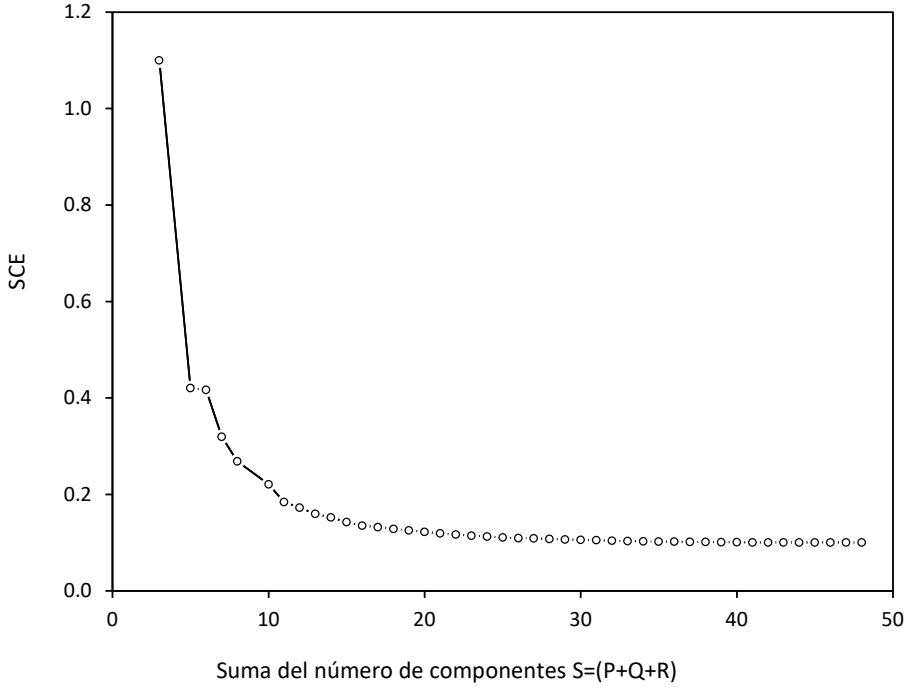


Figura 4.6: Scree plot multiway considerando Lima y Callao

Ahora, al considerar la estructura que excluye a Lima y Callao y al mantener los mismos parámetros de iteración para lograr la convergencia, se obtiene una solución que conserva la disposición de los componentes en todas las dimensiones: $P = 2$, $Q = 2$, $R = 1$ y $S = 5$. Al analizar el Scree plot multiway, se observa que esta estructura se mantiene intacta. Además, al evaluar el criterio *Dif-Dif*, se nota que la variación se suaviza y la convergencia es más rápida en el modelo que excluye a Lima y Callao (Figura 4.7). Aquí, *b(Total)* representa el criterio *Diff-diff* al incluir a Lima y Callao, mientras que *b(Sin Lima y Callao)* lo considera excluyéndolos.

Este nuevo cálculo condujo a una aplicación que captura el 86,6 % de la variabilidad total, lo que comparado con el primer modelo (que incorporaba a Lima y Callao) representó una diferencia de 12,6 puntos porcentuales. Por ende, esta variación se puede atribuir de manera significativa a la influencia de Lima y Callao.

Figura 4.7: Comparación de los resultados del criterio *Diff-Diff*

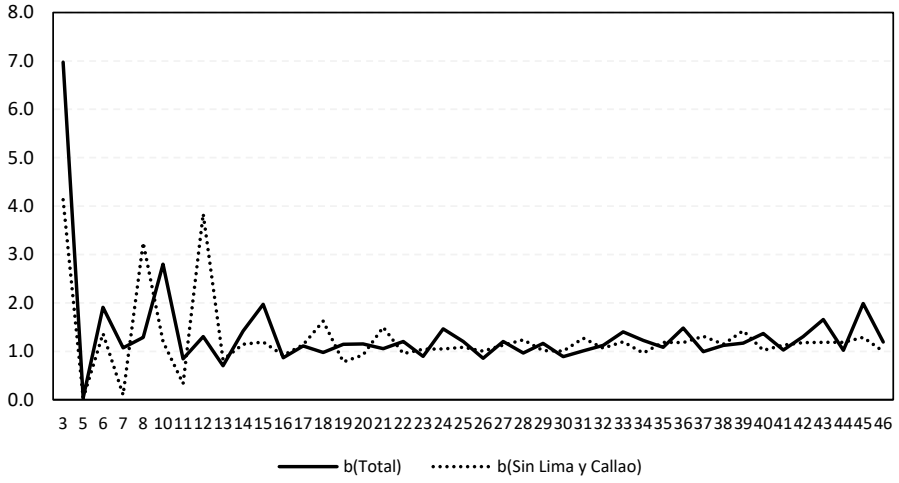


Tabla 4.1: Contribución de componentes a la inercia total de los dos modelos

	Con Lima y Callao		**Sin Lima y Callao**		**Diferencia**	
	Comp. 1	**Comp. 2**	**Comp. 1**	**Comp. 2**	**Comp. 1**	**Comp. 2**
Total	99.2		86.6		12.6	
A	97.4	1.8	66.5	20.1	30.9	-18.4
B	97.4	1.8	66.5	20.1	30.9	-18.4
C	99.2	—	86.6	—	12.6	—

En cuanto a la contribución, el modelo que incorpora a Lima y Callao presenta dos componentes en la dimensión A que representaron el 97,4 % y el 1,8 % respectivamente. En la dimensión B, las dos componentes contribuyeron con el 97,4 % y el 1,8 % respectivamente. Por último, en la dimensión C, la contribución fue del 99,2 %. En todas las vías, destaca el primer componente por su relevante aporte a la explicación de la variabilidad total. No obstante, la estimación que excluye a Lima y Callao evidencia una reducción en la contribución de la primera componente en todas las dimensiones, al mismo tiempo que incrementa la importancia de la segunda componente en las dimensiones A y B. Para obtener más detalles, la tabla (4.1) presenta una comparación entre ambos modelos, considerando las diferencias entre ellos.

Igualmente, en la tabla (4.2) se presenta una comparación entre las triadas en ambas aplicaciones. Como se puede apreciar, al excluir Lima y Callao de la muestra, las estimaciones tienden a equilibrar la contribución en ambas componentes, tanto en

Tabla 4.2: Contribución de triadas de componentes a la inercia total

	Con Lima y Callao		**Sin Lima y Callao**		**Diferencia**	
	B1C1	**B2C1**	**B1C1**	**B2C1**	**B1C1**	**B2C1**
A1	97.4	0.0	66.5	0.0	30.9	0.0
A2	0.0	1.8	0.0	20.1	0.0	-18.4

la vía A como en la B. En el modelo que considera a Lima y Callao, se presentan dos triadas distintas de cero: A1B1C1 y A2B2C1, las cuales contribuyen con un 97,4 % y un 1,8 %, respectivamente. La exclusión de Lima y Callao en la estimación provoca cambios en estos valores. En el caso de la triada A1B1C1, la contribución disminuye a un 66,5 %, lo que representa una reducción de 30,9 puntos porcentuales. En cuanto a la triada A2B2C1, su contribución aumenta al 20,1 %, lo que significa un incremento de 18,4 puntos porcentuales.

4.2.2. Estimación y representación de los componentes

INDIVIDUOS

En ambas aplicaciones, se advierte una solución en la que se determinó la combinación $S = 5$. Esto se logró a través de un conjunto de $I = 24$ individuos (incluyendo Lima y Callao) y $I = 23$ individuos (excluyendo Lima y Callao), junto con $P = 12$ variables y $K = 5$ ocasiones. Esto pone de manifiesto que la presencia de Lima y Callao influyó en la inercia, pero no en la estructuración de la solución.

En el primer modelo, la representación de los componentes muestra el tamaño de la representación de Lima y Callao. Como se puede ver en la figura (4.8), hay una gran distancia entre Lima y Callao y el resto de los departamentos del país. Estos últimos se agrupan en dos grandes grupos, separados entre sí. Esto demuestra la influencia de Lima y Callao en la representación de los componentes. Por lo tanto, se llevará a cabo un análisis más exhaustivo de la aplicación en la que se prescinde de Lima y Callao.

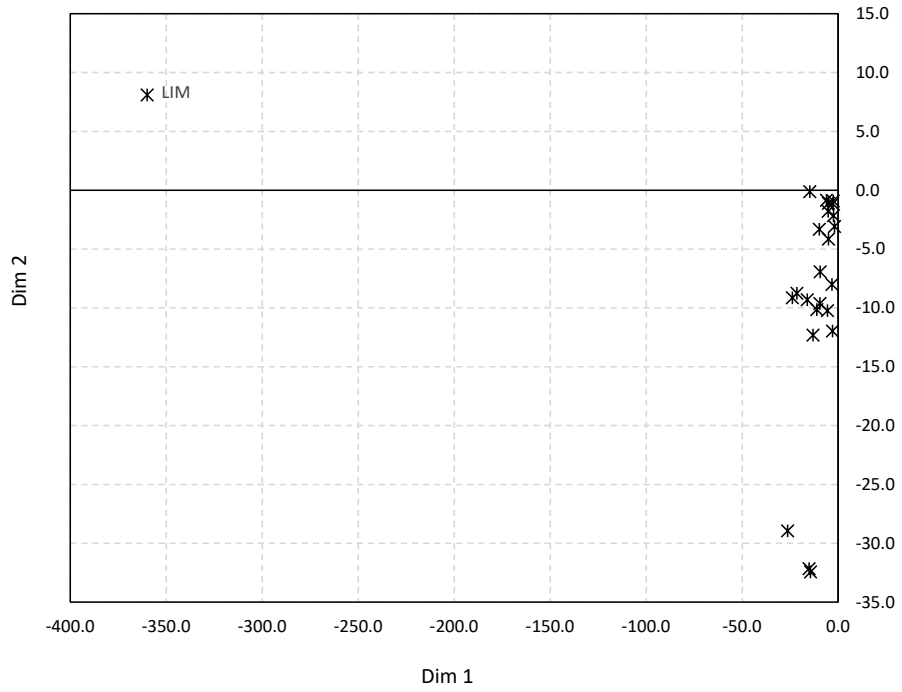


Figura 4.8: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Considerando Lima y Callao)

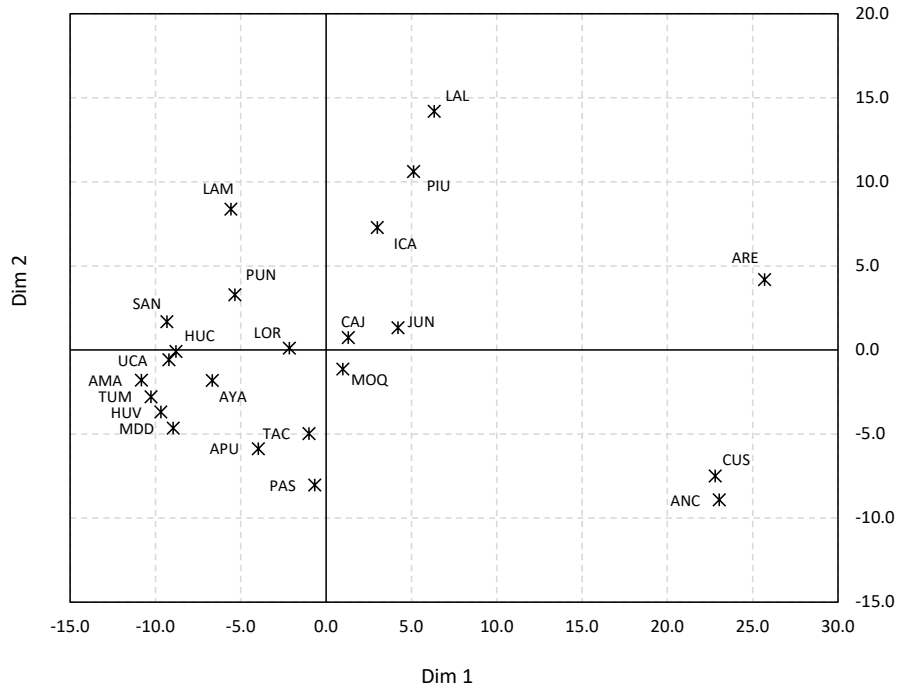


Figura 4.9: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Excluyendo Lima y Callao)

A pesar de la exclusión de Lima y Callao, se puede observar que la relación entre la disposición de los departamentos en el plano de los dos componentes no varía en torno a la relación entre ellos. La figura (4.9) muestra claramente la disposición de los departamentos. De esta se observa a grandes rasgos dos grupos: uno formado por Arequipa, Cusco y Áncash; y el otro constituido por el resto de departamentos. Asimismo, se denotan similitudes entre Lambayeque, La Libertad, Piura e Ica.

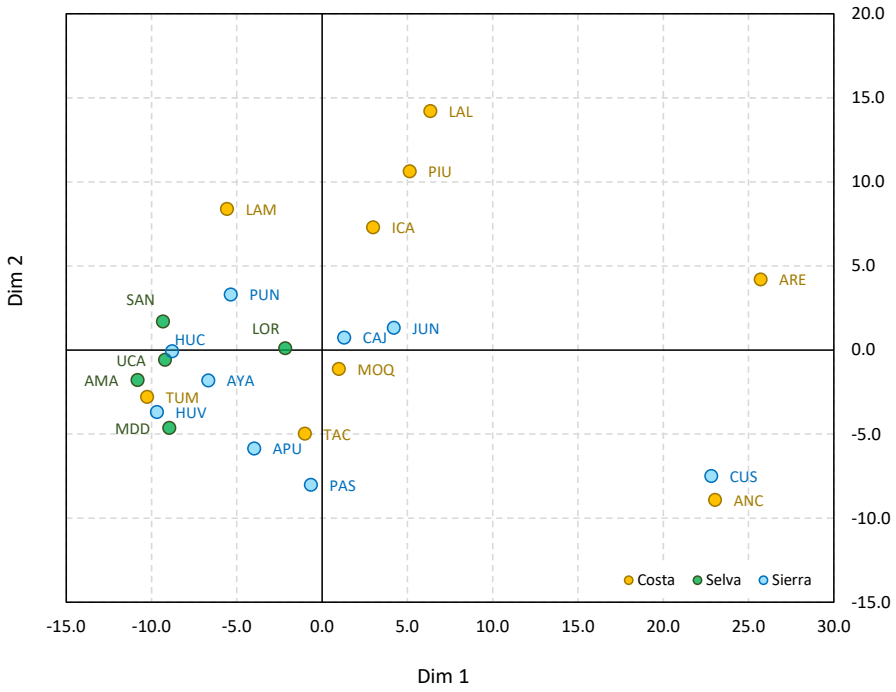


Figura 4.10: Departamentos del Perú en plano factorial (1,2) en Tucker 3 (Excluyendo Lima y Callao)

Cuando se realiza una caracterización de los departamentos, emergen aspectos sustanciales de su relación con los componentes. De esta manera, al considerar su ubicación en las regiones naturales, su relación se percibe con mayor claridad en la figura (4.10). Se establece un vínculo evidente entre departamentos costeros como Lambayeque, La Libertad, Piura e Ica, ya que comparten una ubicación en la costa. Esto subraya la importancia de sus actividades económicas. Esta correlación también se puede identificar entre Cusco y Áncash. Por otro lado, los departamentos de la selva mantienen una relación estrecha entre sí; no obstante, como se analizará más adelante, esta interconexión está relacionada con el tamaño de su producto interno, su población y su extensión territorial.

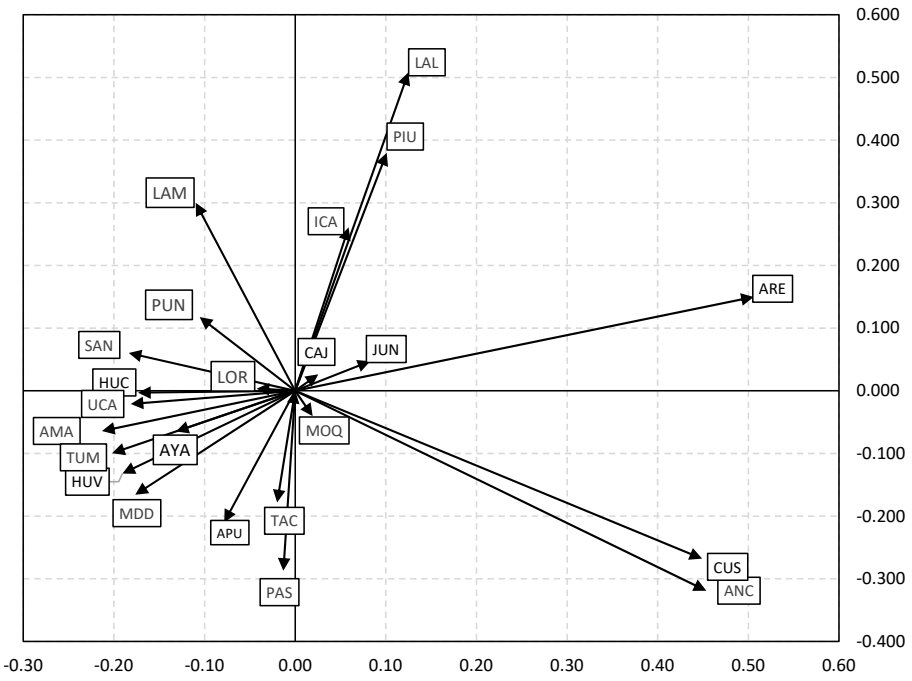


Figura 4.11: Distancias de los Departamentos del Perú en plano factorial (1,2)

VARIABLES

En relación a las 12 variables que han sido evaluadas durante el transcurso de los 15 años, se han identificado dos componentes principales. En el modelo sin la inclusión de Lima y Callao, estos componentes representan un 66,5 % y un 20,1 % de la variabilidad total, respectivamente. De manera similar a la observación en el análisis de individuos que incluye a Lima y Callao, se destaca un elemento distintivo en términos de los componentes de las variables. En este caso, la actividad económica que engloba al petróleo, la minería y el gas emerge como el factor diferenciador (ver figura 4.12). Al aislar esta variable, se revela una relación lineal entre las dos componentes. Esto pone de manifiesto que el grupo conformado por la manufactura y los servicios se diferencia del grupo que comprende actividades como electricidad, hoteles, pesca, telecomunicaciones, sector público y transporte.

La pertenencia a un sector económico es una característica importante para entender la relación entre las diferentes actividades económicas. En el sector secundario, por ejemplo, la manufactura tiene una mayor importancia en comparación con la construcción y la electricidad. En el sector terciario, las actividades relacionadas con los servicios presentan similitudes, aunque el comercio se diferencia ligeramente (figura 4.13). Por otro lado, las actividades del sector primario se muestran aisladas entre sí. Estos resultados evidencian la importancia relativa de cada actividad en la dinámica económica regional y destacan una gran dispersión entre actividades como los servicios y las actividades extractivas (figura 4.14).

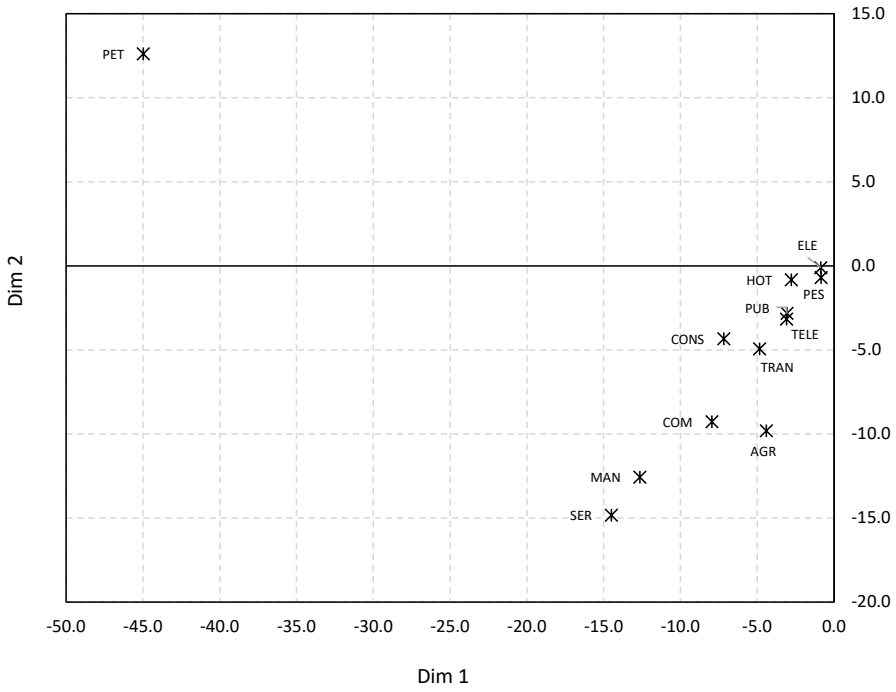


Figura 4.12: Variables sobre el plano (1,2) en Tucker 3 (Excluyendo Lima y Callao)

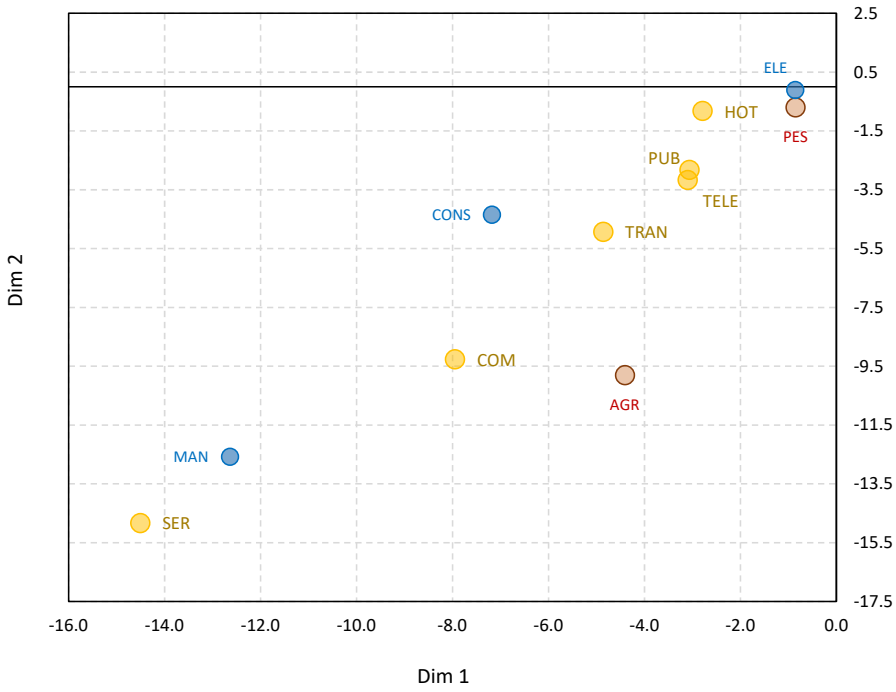


Figura 4.13: Variables sobre el plano (1,2) en Tucker 3 (Aislando la actividad económica de Petróleo, Minería y gas)

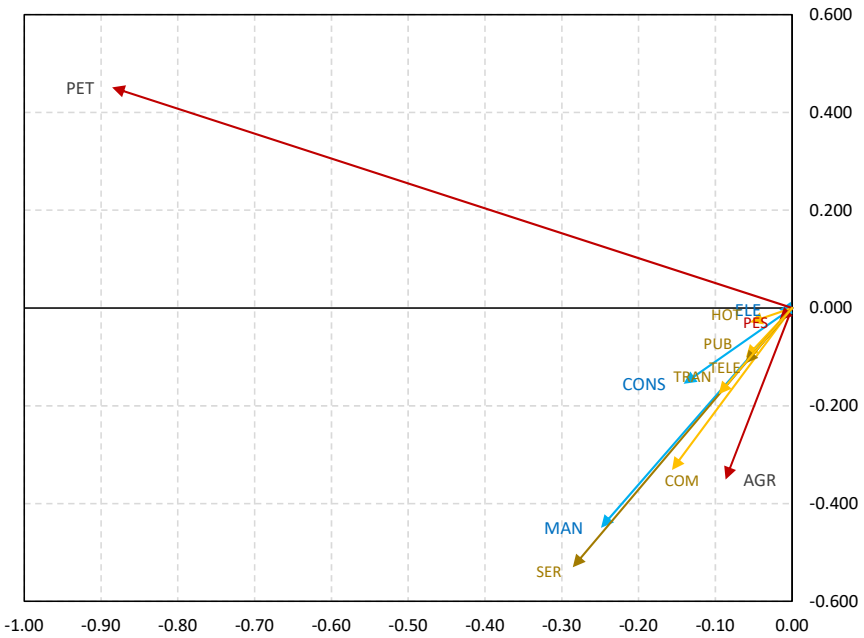


Figura 4.14: Distancias de las variables en el plano (1,2)

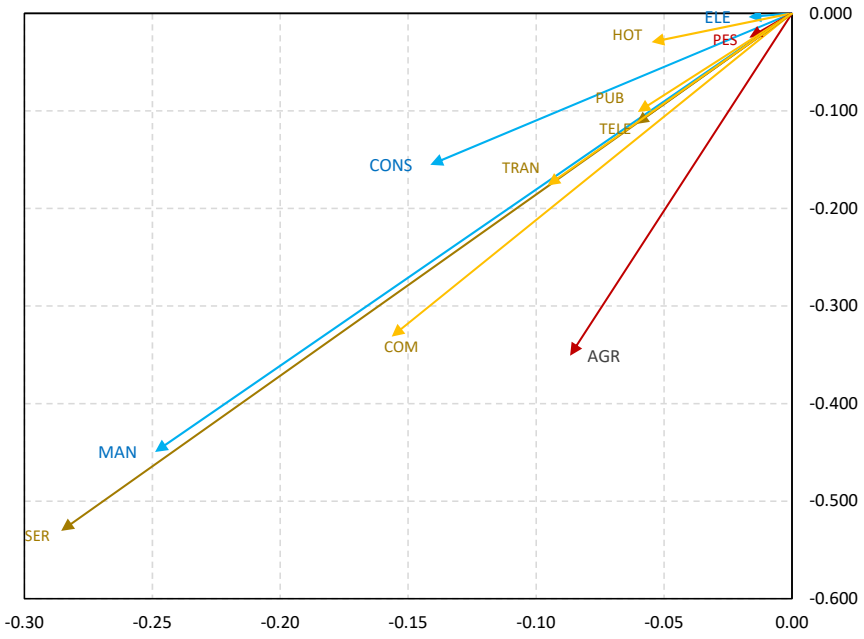
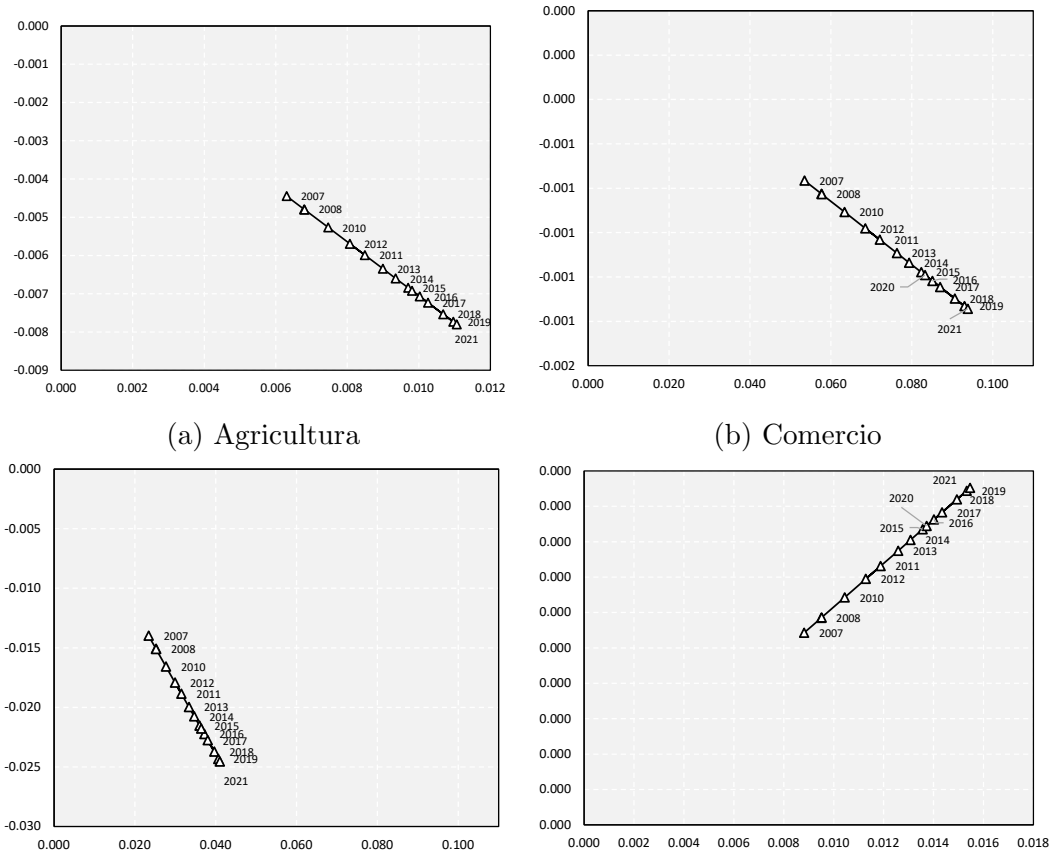


Figura 4.15: Distancias de las variables en el plano (1,2) (sin considerar Petróleo, Minería y Gas

4.2.3. Trayectorias

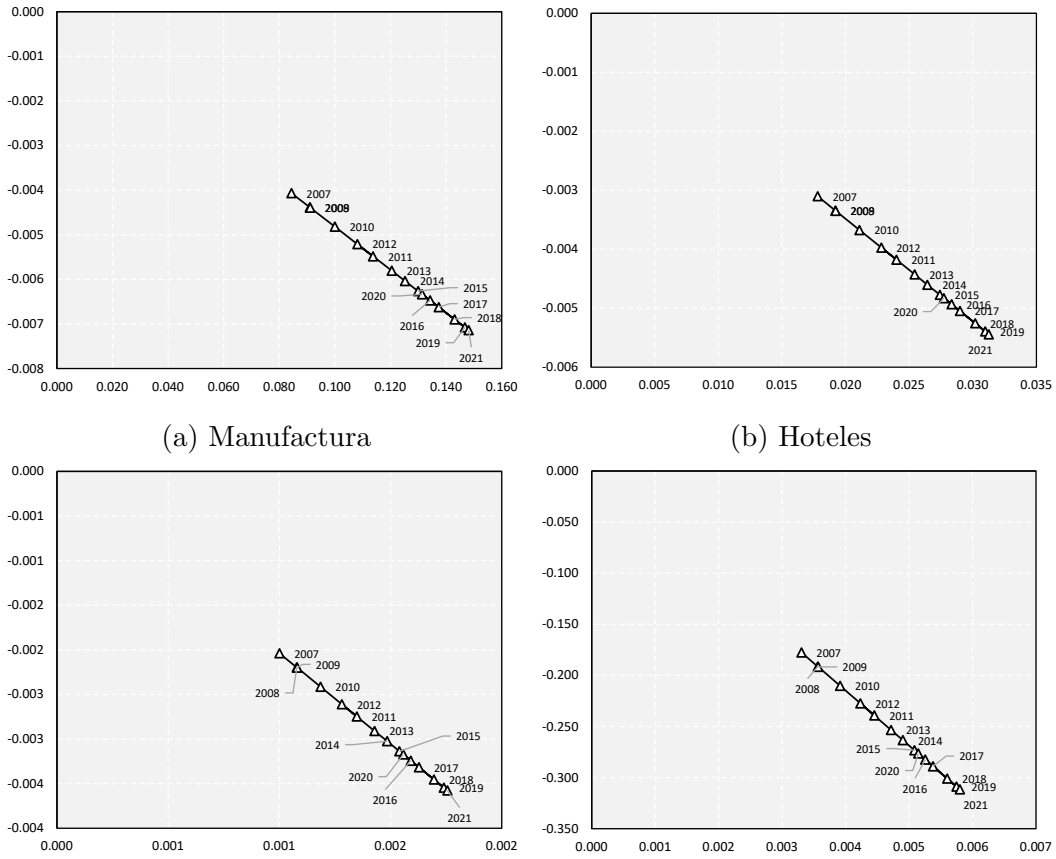
En el análisis de Tucker 3, se observa una tendencia creciente y estable en la representación conjunta de cada variable a lo largo de los 15 años. En particular, no se aprecian diferencias entre las variables en los años 2008 y 2009. Posteriormente, la tendencia se mantuvo en aumento hasta el año 2019. Sin embargo, en el año 2020 se produjo un retroceso hasta los valores registrados en 2015, siendo aún más pronunciado en el caso de la actividad hotelera, que retrocedió hasta el valor obtenido en 2014. Este fenómeno puede ser explicado por el impacto de la crisis financiera en ese año. Finalmente, en el año 2021 se observó un alto crecimiento, retomando los niveles más altos de la tendencia (ver figuras 4.16, 4.17 y 4.18).



(a) Agricultura (b) Comercio

(c) Comercio (d) Electricidad

Figura 4.16: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2)

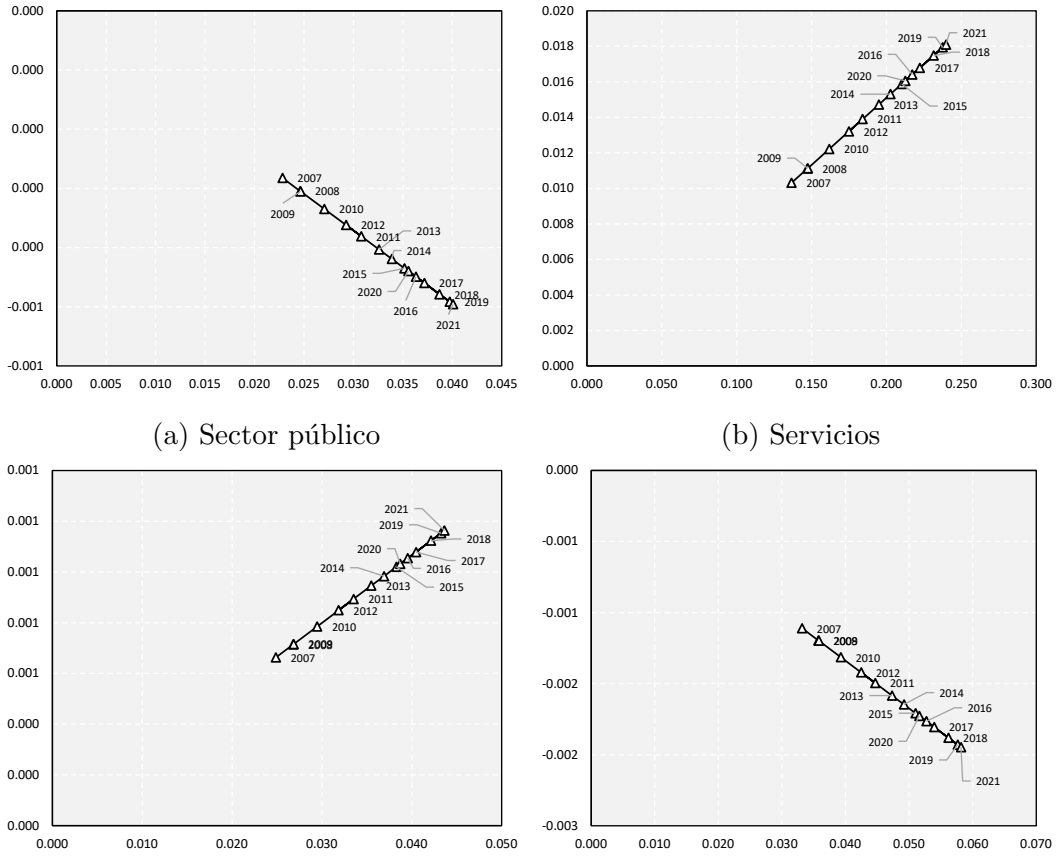


(a) Manufactura (b) Hoteles

(c) Pesca (d) Petróleo, Mineria y gas

Figura 4.17: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2)

El análisis de las trayectorias de los departamentos muestra que la estructura económica regional no ha experimentado cambios significativos entre el periodo 2007 y 2021. Las trayectorias se mantienen estables, lo que refleja la dinámica económica sectorial a nivel departamental (ver figura 4.19 y 4.20). Al considerar cuatro ejemplos específicos (Amazonas, Áncash, Arequipa y Junín), se observa un patrón común en su dinámica. El crecimiento de su economía regional muestra una tendencia positiva hasta el año 2018, para luego retroceder paulatinamente (ver figuras 4.16,4.17 y 4.18). En particular, el año 2020 presenta el peor perfil, ya que retrocede hasta los valores registrados en 2015.



(a) Sector público

(b) Servicios

(c) Telecomunicaciones

(d) Servicios

Figura 4.18: Trayectorias de las variables en el periodo 2007-2021 en Tucker 3, plano(1,2)

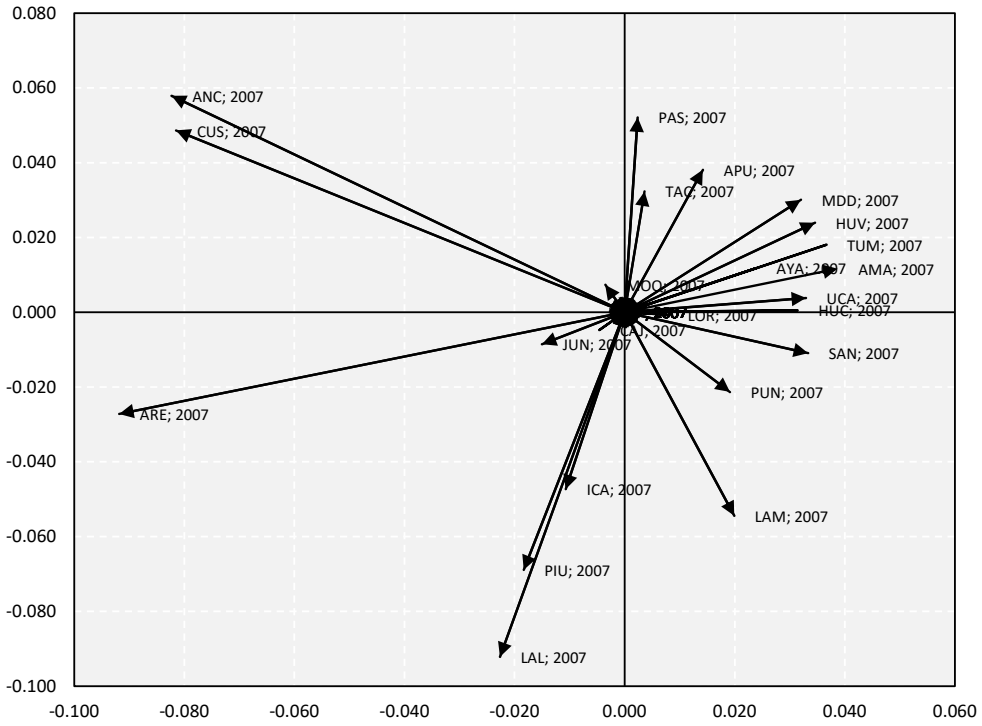


Figura 4.19: Trayectorias de los departamentos en el plano (X,Y) en 2007

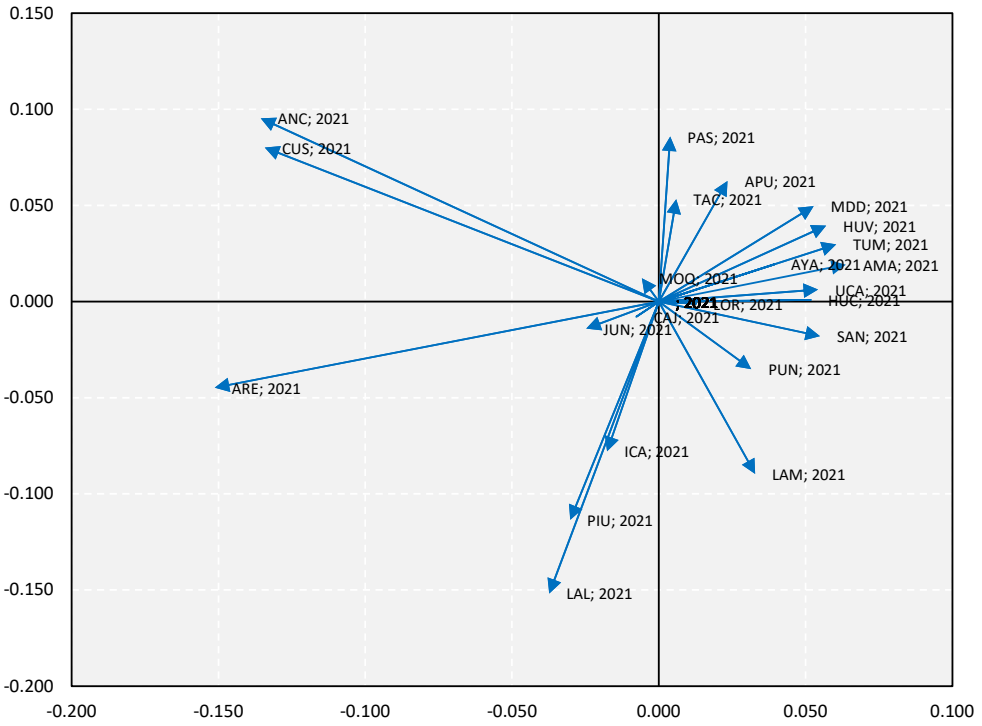


Figura 4.20: Trayectorias de los departamentos en el plano (X,Y) en 2021

Capítulo 5

Resultados y discusión

La revisión exhaustiva de la literatura relacionada con el análisis de tres vías, en particular, el enfoque del modelo Tucker 3, ha resultado fundamental para el estudio, ya que ha permitido explorar de manera profunda tanto los aspectos teóricos como las aplicaciones prácticas de este método en un contexto tridimensional, donde una de las dimensiones clave es el tiempo.

Como se ha evidenciado, la aplicación del modelo Tucker 3 presenta pocas aplicaciones prácticas en el ámbito de la economía; aunque no es parte explicar las razones de tales sesgos, lo cierto es que el modelo Tucker 3 ha sido usado con mayor frecuencia en el análisis de datos masivos, inteligencia artificial, procesamiento de imágenes, etc. En ese sentido, este análisis ha revelado una visión más completa de las características, ventajas y desventajas inherentes al modelo Tucker 3. En términos generales, el modelo Tucker 3 ha demostrado ser altamente efectivo para ajustar y comprender datos que se distribuyen en un arreglo tridimensional que representa a individuos (en este caso, Departamentos), variables y ocasiones (representando el tiempo). Esto va en la línea de lo expuesto por Amaya and Pacheco (2002), Bautista Mendoza (2009), Ruiz, Mujica, Sierra, Pozo and Rodellar (2018) en su aplicación a la economía, y en particular, a las relaciones económicas regionales y sociales.

El modelo ajustado se ha configurado con un total de $S = 8$ componentes, distribuidos en $P = 2$ componentes para la dimensión A, $Q = 2$ para la dimensión B, y $R = 1$ para la dimensión C. Este ajuste ha sido capaz de explicar satisfactoriamente el 86,6 % de la variabilidad total presente en los datos. Sin embargo, es importante destacar que se consideró la exclusión de los datos correspondientes a Lima y Callao

debido a su influencia desproporcionada en el conjunto de datos. Si se hubiera incluido esta información, se habría aumentado el porcentaje de variabilidad explicada en un 12,6 puntos porcentuales. No obstante, la inclusión de estos datos habría dificultado la capacidad para discernir las relaciones y patrones entre los otros individuos. A pesar de esta exclusión, es notable que la estructura de componentes se mantuvo sólida, lo que respalda la conclusión de que el algoritmo determinó la misma dimensión óptima en ambos casos.

Los resultados derivados de la solución del modelo han permitido tener una mayor comprensión de las tres vías, así como de su interacción. Esto es fundamental, ya que permite analizar la dinámica macroeconómica de las variables económicas (VBA) de los departamentos del Perú en un período de tiempo definido. En particular, ha permitido distinguir tendencias marcadas entre los individuos, las variables y el tiempo.

En cuanto a los individuos, se observa una tendencia hacia la afinidad entre diferentes grupos, como los departamentos de Arequipa, Cusco y Áncash por un lado, y Lambayeque, Piura, Ica y La Libertad por otro. Esta relación está claramente marcada en los individuos y muestra su relación con su posición geográfica y su estructura productiva. Mientras tanto, el resto de los departamentos se agrupan en una nube de puntos, destacando los departamentos de la selva por su afinidad geográfica. Es evidente que la ubicación y la geografía determinan la preponderancia de las actividades económicas en una región determinada, lo que resulta en relaciones entre ellos debido a la afinidad estructural de sus actividades económicas. Sin embargo, es importante señalar que existen otras características que pueden explicar estas asociaciones y complementar las relaciones. Aunque esta investigación no tiene como objetivo analizarlas, se presentan como parte de los resultados.

El análisis de las variables indica que, a pesar de la disminución en la participación dentro de la estructura de las actividades económicas, el petróleo, la minería y el gas siguen siendo las actividades más destacadas en el periodo analizado. Su importancia es tal que se distingue notablemente de otros sectores que han ganado relevancia en los últimos años, como los servicios. De hecho, las actividades relacionadas con el sector terciario de la economía muestran una tendencia marcada hacia una mayor asociación con otras actividades como la pesca, agricultura y construcción, formando un bloque diferenciado de los servicios y la manufactura. Se podría inferir que esta dispersión se

debe a su importancia relativa dentro de la estructura productiva, caracterizada por un aumento sostenido de los servicios en la actividad económica del país entre 2007 y 2021.

Durante el período comprendido entre 2007 y 2021, la estructura productiva experimentó una notable estabilidad en líneas generales, aunque con ciertas variaciones destacables. Algunos sectores económicos perdieron parte de su influencia mientras que otros aumentaron su importancia relativa. Esto se tradujo en que la componente relacionada con esta vía representara aproximadamente el 99,2 % del total.

No obstante, al profundizar en el análisis de las trayectorias de las variables, se destaca un crecimiento constante que se vio bruscamente interrumpido en 2020 debido a la pandemia. Esta interrupción provocó un retroceso que hizo retroceder aproximadamente cuatro años en términos de progreso económico.

En lo que respecta a las trayectorias de los individuos, se observó un deterioro más pronunciado, particularmente a partir de 2018, y alcanzó su punto más bajo en 2020. Esta tendencia de declive marcado en los indicadores de desempeño individual fue especialmente acentuada durante este período.

Capítulo 6

Conclusiones

El modelo Tucker 3 como una generalización que abarca el análisis de componentes principales, el análisis factorial y la descomposición en valores singulares para el análisis de tres dimensiones, ha demostrado ser una herramienta eficaz para evaluar la dinámica macroeconómica de Perú durante el período comprendido entre 2007 y 2021. Este modelo ha posibilitado un análisis exhaustivo de las tendencias de datos tridimensionales, que abarcan departamentos, variables económicas y tiempo. Ha proporcionado perspectivas valiosas al revelar similitudes y diferencias notables entre los diversos departamentos del país y las variables económicas estudiadas. Además, debido a la relativa estabilidad en la estructura de las actividades económicas durante este período, el modelo ha sido especialmente efectivo para sintetizar información compleja y esclarecer tendencias emergentes de manera precisa y concisa.

Durante la fase de análisis y exploración de esta investigación, se realizaron descripciones exhaustivas de los principios teóricos que fundamentan el modelo Tucker 3, junto con la presentación del soporte analítico que lo respalda. La dimensionalidad se determinó meticulosamente mediante la aplicación del algoritmo de mínimos cuadrados alternantes, con un criterio de convergencia establecido en $1e-10$. Este proceso permitió establecer los parámetros y componentes correspondientes a las dimensiones A, B y C, así como a la estructura central. De esta manera, se logró cumplir cabalmente con los objetivos específicos y el objetivo general de analizar la estructura macroeconómica de los departamentos a través del enfoque del modelo Tucker 3. Al considerar la dimensión temporal, es relevante destacar que este modelo se puede concebir como un modelo factorial dinámico aplicado al análisis de tres dimensiones.

La aplicación del modelo Tucker evidencia una tendencia hacia la afinidad entre grupos, como los departamentos de Arequipa, Cusco y Áncash, así como Lambayeque, Piura, Ica y La Libertad. Esta afinidad se relaciona con la ubicación geográfica y la estructura productiva de estos departamentos. Por otro lado, otros departamentos se agrupan de manera menos definida, con los departamentos de la selva destacando por su proximidad geográfica. Por tal sentido, es evidente que la ubicación y la geografía juegan un papel crucial en la determinación de las actividades económicas en una región, lo que conduce a relaciones basadas en la afinidad estructural de estas actividades. Aunque esta investigación no se centra en otras características que puedan explicar estas asociaciones, se reconocen como parte de los resultados.

6.1. Sugerencias para investigaciones futuras

- **Extensión del modelo Tucker3 a nuevas dimensiones:** Se recomienda explorar la inclusión de una cuarta dimensión en el modelo, incorporando factores como el comercio internacional, la inversión extranjera o aspectos demográficos. Esto permitiría capturar relaciones aún más complejas dentro de la economía peruana.

- **Comparación con otros métodos de análisis multivariado:** Sería de interés realizar estudios comparativos entre el modelo Tucker3 y otros enfoques multivariantes como el Análisis de Componentes Principales Evolutivo (ACP Evolutivo) o el método STATIS. Dichas comparaciones podrían evaluar la estabilidad de las soluciones obtenidas y su aplicabilidad en distintos contextos económicos.

- **Aplicaciones en predicción y series de tiempo:** Si bien esta investigación se centró en la descripción de patrones estructurales, futuras investigaciones podrían incorporar técnicas de predicción basadas en la evolución temporal de las variables macroeconómicas, integrando el modelo Tucker3 con metodologías de series de tiempo o modelos de Machine Learning.

- **Incorporación de Lima y Callao en los análisis:** Dado que en esta investigación se excluyeron estas regiones debido a su influencia desproporcionada,

un estudio específico sobre el impacto de Lima y Callao en la estructura macroeconómica nacional podría aportar información clave sobre la concentración económica y su efecto en el resto del país.

- **Uso de una versión robusta del modelo Tucker3:** Se sugiere investigar variantes robustas del modelo Tucker3 que permitan minimizar la influencia de valores atípicos y datos ruidosos. Esto podría lograrse mediante enfoques de estimación robusta o técnicas de regularización para mejorar la estabilidad y la interpretabilidad de los resultados.

- **Análisis de efectos de políticas económicas:** Se sugiere aplicar el modelo Tucker3 para evaluar el impacto de distintas políticas económicas en la evolución de las actividades productivas regionales, considerando medidas como incentivos fiscales, subsidios o cambios en el marco regulatorio.

- **Exploración en otros ámbitos socioeconómicos:** Dado que el modelo Tucker3 ha sido poco utilizado en economía, se recomienda extender su aplicación a otras áreas del análisis socioeconómico, como estudios de desigualdad, exclusión social o movilidad laboral, aprovechando su capacidad para modelar interacciones complejas.

- **Desarrollo de herramientas computacionales para la implementación del modelo:** Una línea de investigación relevante sería la creación de paquetes en R o Python que permitan automatizar la aplicación del modelo Tucker3 en análisis económicos, facilitando su uso por parte de investigadores y analistas de datos.

Estas propuestas buscan fortalecer el uso del análisis de tres vías en la economía y ampliar las posibilidades de aplicación del modelo Tucker3 en la investigación macroeconómica y social.

Bibliografía

Amaya, J. L. and Pacheco, P. N. (2002). Análisis factorial dinámico mediante el método tucker3, *Revista Colombiana de Estadística* **25**(1): 43–57.

Amigo, J. M. and Marini, F. (2013). Multiway methods, *Data handling in science and technology*, Vol. 28, Elsevier, pp. 265–313.

Anderson, T. W. (1963). Asymptotic theory for principal component analysis, *The Annals of Mathematical Statistics* **34**(1): 122–148.

Andersson, C. A. and Bro, R. (1998). Improving the speed of multi-way algorithms:: Part i. tucker3, *Chemometrics and intelligent laboratory systems* **42**(1-2): 93–103.

Andersson, C. A. and Bro, R. (2000). The n-way toolbox for matlab, *Chemometrics and intelligent laboratory systems* **52**(1): 1–4.

Andersson, C. A. and Henrion, R. (1999). A general algorithm for obtaining simple structure of core arrays in n-way pca with application to fluorometric data, *Computational statistics & data analysis* **31**(3): 255–278.

Appellof, C. J. and Davidson, E. R. (1981). Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents, *Analytical Chemistry* **53**(13): 2053–2056.

Basford, K., Kroonenberg, P. and DeLacy, I. (1991). Three-way methods for multi-attribute genotype× environment data: an illustrated partial survey, *Field Crops Research* **27**(1-2): 131–157.

Bautista Mendoza, G. R. (2009). Comparación de los métodos tucker 3 y análisis factorial múltiple para el análisis de datos tres vías, *Departamento de Estadística* .

Beffy, J. (1992). Application de l'analyse en composantes principales à trois modes pour l'étude physico-chimique d'un écosystème lacustre d'altitude: perspectives en écologie, *Revue de statistique appliquée* **40**(1): 37–56.

Bentler, P. M. and Lee, S.-Y. (1978). Statistical aspects of a three-mode factor analysis model, *Psychometrika* **43**(3): 343–352.

Bentler, P. M. and Lee, S.-Y. (1979). A statistical development of three-mode factor analysis., *British Journal of Mathematical and Statistical Psychology* .

Bloxom, B. (1968). A note on invariance in three-mode factor analysis, *Psychometrika* **33**(3): 347–350.

Bretherton, C. S., Smith, C. and Wallace, J. M. (1992). An intercomparison of methods for finding coupled patterns in climate data, *Journal of climate* **5**(6): 541–560.

Bro, R. (1997). Parafac. tutorial and applications, *Chemometrics and intelligent laboratory systems* **38**(2): 149–171.

Bro, R. (1998). Multi-way analysis in the food industry, *Models, Algorithms, and Applications. Academish proefschrift. Dinamarca* .

Bro, R. and Kiers, H. A. (2003). A new efficient method for determining the number of components in parafac models, *Journal of Chemometrics: A Journal of the Chemometrics Society* **17**(5): 274–286.

Bro, R. and Smilde, A. K. (2003). Centering and scaling in component analysis, *Journal of chemometrics* **17**(1): 16–33.

Buffo, F. and Verdiell, A. (2004). La descomposición en valores singulares: un enfoque geométrico y aplicaciones, *Revista de Educación Matemática* **19**(1): 33–13.

Carroll, J. D. and Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition, *Psychometrika* **35**(3): 283–319.

Cattell, R. B. (1952). Factor analysis: an introduction and manual for the psychologist and social scientist.

Chen, B., Sun, T., Zhou, Z., Zeng, Y. and Cao, L. (2019). Nonnegative tensor completion via low-rank tucker decomposition: model and algorithm, *IEEE Access* **7**: 95903–95914.

Chow, S.-M., Zu, J., Shifren, K. and Zhang, G. (2011). Dynamic factor analysis models with time-varying parameters, *Multivariate Behavioral Research* **46**(2): 303–339.

Cichocki, A., Zdunek, R., Phan, A. H. and Amari, S.-i. (2009). *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*, John Wiley & Sons.

Comon, P., Luciani, X. and De Almeida, A. L. (2009). Tensor decompositions, alternating least squares and other tales, *Journal of Chemometrics: A Journal of the Chemometrics Society* **23**(7-8): 393–405.

Coppi, R. and Bolasco, S. (1989). *Multiway data analysis*, North-Holland Publishing Co.

Cuadras, C. M. (1996). *Nuevos métodos de análisis multivariante*, CMC Edicions Barcelona, Spain.

De Lathauwer, L. and De Moor, B. (1998). From matrix to tensor: Multilinear algebra and signal processing, *Institute of mathematics and its applications conference series*, Vol. 67, Citeseer, pp. 1–16.

Dell'Anno, R. and Amendola, A. (2015). Social exclusion and economic growth: An empirical investigation in european economies, *Review of Income and Wealth* **61**(2): 274–301.

Dos Santos, R. F., Paraskevaidi, M., Mann, D. M., Allsop, D., Santos, M. C., Morais, C. L. and Lima, K. M. (2022). Alzheimer's disease diagnosis by blood plasma molecular fluorescence spectroscopy (eem), *Scientific Reports* **12**(1): 16199.

Estienne, F., Matthijs, N., Massart, D., Ricoux, P. and Leibovici, D. (2001). Multi-way modelling of high-dimensionality electroencephalographic data, *Chemometrics and Intelligent Laboratory Systems* **58**(1): 59–72.

Fallas-Monge, J. J., Chavarrıa-Molina, J. and Soto-Quiros, P. (2021). Descomposicion en valores singulares de una matriz: Un repaso por los fundamentos teoricos y sus aplicaciones en el procesamiento de imagenes, *Revista de Investigación Operacional* **42**(2): 148–173.

Flores Camacho, C. F. (2013). Análisis factorial dinámico para la proyección de las exportaciones en el perú.

Frutos-Bernal, E., Martín del Rey, Á., Mariñas-Collado, I. and Santos-Martín, M. T. (2022). An analysis of travel patterns in barcelona metro using tucker3 decomposition, *Mathematics* **10**(7): 1122.

Gallo, M. (2015). Tucker3 model for compositional data, *Communications in Statistics-Theory and Methods* **44**(21): 4441–4453.

Gandour, J. T. and Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation, *Language and speech* **21**(1): 1–33.

Geladi, P. (1989). Analysis of multi-way (multi-mode) data, *Chemometrics and intelligent laboratory systems* **7**(1-2): 11–30.

Gifi, A. (1990). Nonlinear multivariate analysis, *(No Title)* .

Giordani, P., Kiers, H. A. and Del Ferraro, M. A. (2014). Three-way component analysis using the r package threeway, *Journal of Statistical Software* **57**: 1–23.

Golub, G. H. and Van Loan, C. F. (2013). *Matrix computations*, JHU press.

Gordon, F., Elcoroaristizabal, S. and Ryder, A. G. (2021). Modelling förster resonance energy transfer (fret) using anisotropy resolved multi-dimensional emission spectroscopy (armes), *Biochimica et Biophysica Acta (BBA)-General Subjects* **1865**(2): 129770.

Harshman, R. A. (1984). Data preprocessing and the extended parafac model, *Research Methods for Multi-mode Data Analysis.* pp. 216–284.

Harshman, R. A. and Lundy, M. E. (1984). The parafac model for three-way factor analysis and multidimensional scaling, *Research methods for multimode data analysis* **46**: 122–215.

Harshman, R. A. et al. (1970). Foundations of the parafac procedure: Models and conditions for an.[ex]planatory" multimodal factor analysis.

Henrion, R. (1998). Simultaneous simplification of loading and core matrices in n-way pca: application to chemometric data arrays, *Fresenius' journal of analytical chemistry* **361**(1): 15–22.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components., *Journal of educational psychology* **24**(6): 417.

Ishida, A., Fujii, K., Yamamoto, N., Murakami, J., Saito, S. and Kano, N. (2014). Application of nonnegative tucker decomposition in medical data analysis, *Advanced Materials Research*, Vol. 971, Trans Tech Publ, pp. 1874–1883.

Johnson, R. A., Wichern, D. W. et al. (2014). *Applied multivariate statistical analysis*, Vol. 6, Pearson London, UK:.

Jolliffe, I. T. (2002). Principal component analysis, *Technometrics* **45**(3): 276.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis, *Psychometrika* **23**(3): 187–200.

Kanatani, K. (2021). Linear algebra for pattern processing: Projection, singular value decomposition, and pseudoinverse, *Synthesis Lectures on Signal Processing* **12**(1): 1–155.

Kent, J., Bibby, J. and Mardia, K. (1979). *Multivariate analysis*, Academic Press Amsterdam.

Kiers, H. A. (1988). Comparison of "anglo-saxon" and "french" three-mode methods, *Statistique et analyse des données* **13**(3): 14–32.

Kiers, H. A. (2000). Towards a standardized notation and terminology in multiway analysis, *Journal of Chemometrics: A Journal of the Chemometrics Society* **14**(3): 105–122.

Kiers, H. A. (2004). Bootstrap confidence intervals for three-way methods, *Journal of Chemometrics: A Journal of the Chemometrics Society* **18**(1): 22–36.

Kim, Y.-D. and Choi, S. (2007). Nonnegative tucker decomposition, *2007 IEEE conference on computer vision and pattern recognition*, IEEE, pp. 1–8.

Koleini, F., Hugelier, S., Lakeh, M. A., Abdollahi, H., Camacho, J. and Gemperline, P. J. (2023). On the complementary nature of anova simultaneous component analysis (asca+) and tucker3 tensor decompositions on designed multi-way datasets, *Journal of Chemometrics* p. e3514.

Kroonenberg, P. and de Leeuw, J. (1993). Principal component analysis of three-mode data.

Kroonenberg, P. M. (1983). *Three-mode principal component analysis: Theory and applications*, Vol. 2, DSWO press.

Kroonenberg, P. M. (2008). *Applied multiway data analysis*, John Wiley & Sons.

Kroonenberg, P. M. (2016). My multiway analysis: from jan de leeuw to twpack and back, *Journal of Statistical Software* **73**: 1–22.

Kroonenberg, P. M. and De Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms, *Psychometrika* **45**(1): 69–97.

Kroonenberg, P. M., Harshman, R. A. and Murakami, T. (2009). Analysing three-way profile data using the parafac and tucker3 models illustrated with views on parenting, *Applied Multivariate Research* **13**(1): 5–41.

Kruskal, J. B. (1981). Multilinear models for data analysis, *Behaviormetrika* **8**(10): 1–20.

Lee, S.-Y. and Fong, W.-K. (1983). A scale invariant model for three-mode factor analysis, *British Journal of Mathematical and Statistical Psychology* **36**(2): 217–223.

Leurgans, S. and Ross, R. T. (1992). Multilinear models: applications in spectroscopy, *Statistical Science* pp. 289–310.

Levin, J. (1965). Three-mode factor analysis., *Psychological Bulletin* **64**(6): 442.

Martin-Barreiro, C., Ramirez-Figueroa, J. A., Nieto-Librero, A. B., Leiva, V., Martin-Casado, A. and Galindo-Villardón, M. P. (2021). A new algorithm for computing disjoint orthogonal components in the three-way tucker model, *Mathematics* **9**(3): 203.

Mitsushita, K., Murakoshi, S. and Koyama, M. (2023). How are various natural disasters cognitively represented?: a psychometric study of natural disaster risk perception applying three-mode principal component analysis, *Natural Hazards* **116**(1): 977–1000.

Molenaar, P. (1985). A dynamic factor model for the analysis of multivariate time series, *Psychometrika* **50**(2): 181–202.

Molenaar, P. C. (1987). Dynamic factor analysis in the frequency domain: Causal modeling of multivariate psychophysiological time series, *Multivariate Behavioral Research* **22**(3): 329–353.

Naik, G. R. (2017). *Advances in principal component analysis: research and development*, Springer.

Paredesa, J. E. R., Lozanoa, J. M. B., Restrepoa, J. M. and Londonob, D. A. (2018). Análisis de datos en tres vıas para la evaluación de la dinámica económica de los departamentos en colombia durante el periodo 2000-2015. three way data analysis for the evaluation of the economic dynamic of departments of colombia during.

Peña, D. (2002). Análisis multivariante de datos, *Mc Graw Hill* .

Phan, A. H., Tichavskỳ, P. and Cichocki, A. (2011). Damped gauss-newton algorithm for nonnegative tucker decomposition, *2011 IEEE Statistical Signal Processing Workshop (SSP)*, IEEE, pp. 665–668.

Pla, L. E. (1986). *Análisis multivariado: método de componentes principales*, number 519.535 P696, OEA, Washington, DC (EUA). Secretaría General. Programa Regional de ....

Pravdova, V., Estienne, F., Walczak, B. and Massart, D. (2001). A robust version of the tucker3 model, *Chemometrics and Intelligent Laboratory Systems* **59**(1-2): 75–88.

Raljević, D., Vuković, J. P., Smrečki, V., Pajc, L. M., Novak, P., Hrenar, T., Jednačak, T., Konjević, L., Pinević, B. and Gašparac, T. (2021). Machine learning approach for predicting crude oil stability based on nmr spectroscopy, *Fuel* **305**: 121561.

Ramos, M., Ascencio, J., Hinojosa, M. V., Vera, F., Ruiz, O., Jimenez-Feijoó, M. I. and Galindo, P. (2021). Multivariate statistical process control methods for batch production: a review focused on applications, *Production & Manufacturing Research* **9**(1): 33–55.

Reitsema, A. M., Jeronimus, B. F., van Dijk, M., Ceulemans, E., van Roekel, E., Kuppens, P. and de Jonge, P. (2023). Distinguishing dimensions of emotion dynamics across 12 emotions in adolescents' daily lives., *Emotion* **23**(6): 1549.

Rella, J. C. (2020). en galego: Modelo factorial dinámico para a economía galega, *journal of economics* **120**(1): 387–422.

Royer, J.-P., Thirion-Moreau, N. and Comon, P. (2012). Nonnegative 3-way tensor factorization taking in to account possible missing data, *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, IEEE, pp. 71–75.

Ruiz, J. E., Bonilla, J. M., Moquera, J. and Arango, D. (2018). Análisis de datos en tres vías para la evaluación de la dinámica económica de los departamentos en colombia durante el periodo 2000-2015, *II Congreso Colombiano de Estadística* .

Ruiz, M., Mujica, L. E., Sierra, J., Pozo, F. and Rodellar, J. (2018). Multiway principal component analysis contributions for structural damage localization, *Structural health monitoring* **17**(5): 1151–1165.

Sanguansat, P. (2012). *Principal component analysis*, BoD–Books on Demand.

Smilde, A. K. (1992). Three-way analyses problems and prospects, *Chemometrics and Intelligent Laboratory Systems* **15**(2-3): 143–157.

Soldevilla Olivares, M. J. (2004). *Descomposición del valor singular y sus aplicaciones*, PhD thesis, Universidad Nacional de Ingeniería.

Timmerman, M. E. and Kiers, H. A. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima, *British journal of mathematical and statistical psychology* **53**(1): 1–16.

Trefethen, L. N. and Bau III, D. (1997). *Numerical linear algebra*, Vol. 50, Siam.

Tucker, L. R. (1958). An inter-battery method of factor analysis, *Psychometrika* **23**(2): 111–136.

Tucker, L. R. (1963). Implications of factor analysis of three-way matrices for measurement of change, *Problems in measuring change* **15**(122-137): 3.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis, *Psychometrika* **31**(3): 279–311.

Tucker, L. R. et al. (1964). The extension of factor analysis to three-dimensional matrices, *Contributions to mathematical psychology* **110119**.

Turk, M. A. and Pentland, A. P. (1991). Face recognition using eigenfaces, *Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition*, IEEE Computer Society, pp. 586–587.

Vidal, R., Ma, Y. and Sastry, S. (2005). Generalized principal component analysis (gpca), *IEEE transactions on pattern analysis and machine intelligence* **27**(12): 1945–1959.

Wallace, J. M., Smith, C. and Bretherton, C. S. (1992). Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies, *Journal of climate* **5**(6): 561–576.

Wu, D., Yang, B., Wang, H., Wu, D. and Wang, R. (2016). An energy-efficient data forwarding strategy for heterogeneous wbans, *IEEE access* **4**: 7251–7261.

Zuur, A., Tuck, I. and Bailey, N. (2003). Dynamic factor analysis to estimate common trends in fisheries time series, *Canadian journal of fisheries and aquatic sciences* **60**(5): 542–552.

Apéndice A

Apéndice. Análisis de componentes principales

El análisis de componentes principales (ACP) es probablemente el más antiguo y mejor conocido de las técnicas multivariantes. Introducido en 1901 por Pearson y desarrollado por Hotelling alrededor de 1933 fue, al igual que otros métodos multivariantes, ampliamente usado con el desarrollo computacional. Se trata, pues, de un método matemática que no requiere un modelo estadístico para explicar la estructura probabilística de los errores (Jolliffe, 2002; Turk and Pentland, 1991; Pla, 1986).

La idea central del ACP es la reducción de las dimensiones de un conjunto de datos, los cuales están altamente correlacionados. No obstante, Johnson et al. (2014) y Naik (2017) consideran que otro aspecto importante es la interpretación que resulta de revelar los factores ocultos que subyacen al conjunto de variables. Entonces, el resultado es una transformación a un nuevo conjunto de variables, los componentes principales, que no están correlacionados entre ellos y que capturan la mayor cantidad de información posible de los datos originales.

Si bien se requieren p componentes para reproducir la variabilidad de todo el sistema, frecuentemente la mayoría de la variabilidad puede ser calculada a través de un número menor de k componentes principales. Por ende, si hay (casi) tanta información en los k componentes como en las p variables originales, luego las k componentes principales pueden entonces remplazar a las p variables originales, y el conjunto de los datos originales, consistente con las n mediciones en las k componentes principales.

Vidal et al. (2005) denominan al ACP como un ajuste de un subespacio afín de baja

dimensión a un conjunto de puntos de datos en un espacio de alta dimensión. Esto implica que el PCA gira los ejes del espacio de datos a lo largo de las líneas de máxima varianza. De esta manera, el eje de mayor varianza se denomina primer componente principal; otro eje, que es el siguiente de mayor varianza, se denomina segundo componente principal, y así sucesivamente. Así, la reducción de las dimensiones se realiza utilizando solo los primeros componentes principales como base establecida para el nuevo espacio: el subespacio tiende a ser pequeño y puede eliminarse con una mínima pérdida de información (Sanguansat, 2012).

Entonces, suponiendo un conjunto de valores de p variables aleatorias en n elementos de una población, los cuales se encuentran en una matriz $\mathbf{X}$ de dimensión $n \times p$. En esta estructura, las columnas contienen a las variables, mientras que las filas a los individuos. De esto, lo que interesa son las varianzas, covarianzas o correlaciones de las p variables (Peña, 2002).

De acuerdo a lo anterior, los componentes principales son una combinación lineal particular de p variables aleatorias $X_1, X_2, ..., X_p$. Geométricamente, esa combinación lineal representa la selección de un nuevo sistema de coordenadas, las cuales son obtenidas por la rotación del sistema original con las $X_1, X_2, ..., X_p$ como los ejes de coordenadas. Los nuevos ejes representan las direcciones con máxima variabilidad y proporciona una descripción más simple y parsimoniosa de la estructura de covarianzas (Sanguansat, 2012; Johnson et al., 2014).

Como se mencionó anteriormente, los componentes principales depende solamente de la matriz covarianzas, Σ, o la matriz de correlaciones, ρ de los $X_1, X_2, ..., X_p$ (Jolliffe, 2002; Johnson et al., 2014). Asimismo, su desarrollo no requiere una suposición normal multivariada, tal como lo destaca Pla (1986); sin embargo, la suposición de multinormalidad permite estudiar la significancia estadística y, con ello, es posible utilizar una muestra observada para las pruebas de hipótesis que permitan conocer la estructura poblacional original (Pla, 1986).

Considerando que el vector aleatorio $X' = [X_1, X_2, ..., X_p]$ tiene una matriz de covarianzas Σ con autovalores $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$, se puede considerar la siguiente

combinación lineal:

$$\begin{aligned} Y_1 &= \mathbf{a_1'X} = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p \\ Y_2 &= \mathbf{a_2'X} = a_{21}X_1 + a_{22}X_2 + ... + a_{2p}X_p \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ Y_p &= \mathbf{a_p'X} = a_{p1}X_1 + a_{p2}X_2 + ... + a_{pp}X_p \end{aligned} \tag{A.1}$$

Del sistema anterior, de acuerdo a Johnson et al. (2014) se obtiene:

$$Var(Y_i) = \mathbf{a_i'\Sigma a_i} \qquad i = 1, 2, ..., p \tag{A.2}$$

$$Cov(Y_i, Y_k) = \mathbf{a_i'\Sigma a_k} \qquad i, k = 1, 2, ..., p \tag{A.3}$$

Los componentes principales son esas combinaciones lineales incorrelacionadas, $Y_1, Y_2, ..., Y_p$, cuya varianza en (A.2) es la mayor posible. Así, la primera componente principal es la combinación lineal con la máxima varianza. Es decir, maximiza $Var(Y_1) = \mathbf{a_1'\Sigma a_1}$. En este punto, es claro que $Var(Y_1) = \mathbf{a_1'\Sigma a_1}$ puede ser incrementada multiplicando cualquier constante $\mathbf{a_1}$, para evitar esa indeterminación, es conveniente restringir a los vectores de coeficientes de longitud unitaria. Por lo tanto, siguiendo a Johnson et al. (2014), se define:

La primera componente principal = la combinación lineal $\mathbf{a_1'X}$que maximiza

$$Var(Y_1) = \mathbf{a_1'\Sigma a_1} \text{ sujeto a } \mathbf{a_1'a_1} = 1$$

La segunda componente principal = la combinación lineal $\mathbf{a_2'X}$que maximiza

$$Var(Y_2) = \mathbf{a_2'\Sigma a_2} \text{ sujeto a } \mathbf{a_2'a_2} = 1 \text{ y}$$

$$Cov(\mathbf{a_1'X}, \mathbf{a_2'X}) = 0$$

en el el $i-$ésimo elemento

La $i-$ésima componente principal = la combinación lineal $\mathbf{a_i'X}$ que

$$\text{maximiza } Var(Y_i) = \mathbf{a_i'\Sigma a_i}$$

$$\text{sujeto a } \mathbf{a_2'a_2} = 1 \text{ y}$$

$$Cov(\mathbf{a_1'X}, \mathbf{a_2'X}) = 0 \text{ para} k < i$$

Ahora, dado que Σ es la matriz de covarianza asociada al vector aleatorio $\mathbf{X}$, esta tendrá pares de autovalores y autovectores de la siguiente manera, $(\lambda_1, \mathbf{e_1}), (\lambda_2, \mathbf{e_2}), ..., (\lambda_p, \mathbf{e_p})$,

donde $\lambda_1 \geq \lambda_2 \geq; ...; \lambda_p \geq 0$; por lo que, la i-ésima componente principal está dada por

$$Y_i = \mathbf{e_i'X} = e_{i1}X_1 + e_{i2}X_2 + ... + e_{ip}X_p \tag{A.4}$$

donde se obtiene

$$\begin{aligned} Var(Y_i) &= \mathbf{e_i'\Sigma e_i} = \lambda_i \qquad\qquad i = 1, 2, ..., p \\ Cov(Y_i, Y_k) &= \mathbf{e_i'\Sigma e_k} = 0 \qquad\qquad i \neq k \end{aligned} \tag{A.5}$$

Asimismo, se tiene que:

$$\sigma_{11} + \sigma_{22} + ... + \sigma_{pp} = \sum_{i=1}^{p} Var(X_i) = \lambda_1 + \lambda_2 + ... + \lambda_p = \sum_{i=1}^{p} Var(Y_i) \tag{A.6}$$

De lo anterior es posible definir:

$$\begin{aligned} \text{varianza total de la población} &= \sigma_{11} + \sigma_{22} + ... + \sigma_{pp} \\ &= \lambda_1 + \lambda_2 + ... + \lambda_p \end{aligned} \tag{A.7}$$

Consecuentemente, se define la varianza explicada por la k-ésima principal:

$$\begin{aligned} &\text{(Proporción total de varianza poblacional} \\ &\text{debido al } k\text{-ésimo componente principal)} = \frac{\lambda_k}{\lambda_1 + \lambda_2 + ... + \lambda_p} \end{aligned} \tag{A.8}$$

para $k = 1, 2, ..., p$

Ahora, si $Y_1 = \mathbf{e_1'}$, $Y_2 = \mathbf{e_2'}$,..., $Y_p = \mathbf{e_p'}$ son los componentes principales obtenidos de la matriz de covarianza Σ, entonces, los coeficientes de correlación entre los componentes Y_i y las variables X_k están dados por:

$$\rho_{Y_i, X_k} = \frac{e_{ik}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \qquad\qquad i, k = 1, 2, ..., p \tag{A.9}$$

Aquí $(\lambda_1, \mathbf{e_1}), (\lambda_2, \mathbf{e_2}), ..., (\lambda_p, \mathbf{e_p})$ son los pares de los autovalores y autovectores de Σ, tal como se vio anteriormente. Suponiendo que los componentes principales derivados de variables aleatorias son normales multivariadas ($\mathbf{X}$ se distribuye como $N_p(\mu, \mathbf{\Sigma})$), la densidad de $\mathbf{X}$ es constante en los elipsoides centrados en μ

$$(\mathbf{x} - \mu)'\mathbf{\Sigma^{-1}}(\mathbf{x} - \mu) = c^2 \tag{A.10}$$

cuyos ejes son $\pm c\sqrt{\lambda_i}\mathbf{e_i}$, para $i = 1, 2, ..., p$. Donde los $(\lambda_i, \mathbf{e_i})$ son los pares de autovalores y autovectores de Σ. Un punto que se encuentra en el i-ésimo eje del

elipsoide tendrá coordenadas proporcionales a $\mathbf{e_i'} = [e_{i1}, e_{i2}, ..., e_{ip}]$ en el sistema de coordenadas, que tiene como origen μ y los ejes son paralelos a los ejes originales $x_1, x_2, ..., x_p$. Por conveniencia se considerará $\mu = 0$, de manera que se puede escribir (A.10) de la siguiente manera:

$$c^2 = \mathbf{x'\Sigma^{-1}x} = \frac{1}{\lambda_1}(\mathbf{e_1'x})^2 + \frac{1}{\lambda_2}(\mathbf{e_2'x})^2 + ... + \frac{1}{\lambda_p}(\mathbf{e_p'x})^2$$

Donde $\mathbf{e_1'x}, \mathbf{e_2'x}; ...; \mathbf{e_p'x}$ son reconocidos como componentes principales de $\mathbf{x}$. Ajustando con $y_1 = \mathbf{e_1'x}$, $y_2 = \mathbf{e_2'x}$, ..., $y_p = \mathbf{e_p'x}$ se tiene

$$c^2 = \frac{1}{\lambda_1}y_1^2 + \frac{1}{\lambda_2}y_2^2 + ... + \frac{1}{\lambda_p}y_p^2 \tag{A.11}$$

La ecuación (A.11) define un elipsoide, ya que $(\lambda_1, \lambda_2, ..., \lambda_p)$ son positivos, en un sistema de cordenadas con ejes $y_1, y_2, ..., y_p$ tendiendo en las direcciones de $\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_p}$, respectivamente. Si λ_1 es el mayor autovalor, entonces el eje mayor conduce en la dirección $\mathbf{e_1}$. Los ejes menores restantes se encuentran en las direcciones definidas por $\mathbf{e_2}, ..., \mathbf{e_p}$.

A modo de resumen, las componentes principales $y_1 = \mathbf{e_1'x}, y_2 = \mathbf{e_2'x}, ..., y_p = \mathbf{e_p'x}$ se encuentran en las direcciones de los ejes del elipsoide de densidad constante. Por lo tanto, cualquier punto en el i-ésimo eje del elipsoide tiene coordenadas $\mathbf{x}$ proporcionales a $\mathbf{e_i'} = [e_{i1}, e_{i2}, ..., e_{ip}]$ y, necesariamente, las coordenadas del componente principal son de la forma $[0, ..., 0, y_i, 0, ..., 0]$. Asimismo, de acuerdo a Peña (2002), los componentes principales como nuevas variables presentan las siguientes propiedades:

- Conservan la variabilidad inicial. La suma de las varianzas de los componentes principales es igual a la suma de las varianzas de las variables originales, tal como se mostró en la ecuación (A.6).

- La proporción de variabilidad explicada por un componente es determinado por el cociente entre su varianza, el valor propio (asociado a su vector propio), y la suma de los valores propios de la matriz. Esto se ve expresado en la ecuación (A.8).

- Las covarianzas entre cada componente principal y las variables $\mathbf{X}$ están definidas por el producto de las coordenadas del vector propio y el valor propio:

$$cov(Y_i; x_1, ..., x_n) = (\lambda_i a_{i1}, ..., \lambda_i a_{ip})$$

- La correlación entre una componente principal y una variable $\mathbf{X}$ es proporcional al cociente de esa variable y el coeficiente de proporcionalidad, que es definido como el cociente entre la desviación típica del componente y la desviación típica de la variable, tal como se muestra en la ecuación (A.9).

- Las r componentes principales ($r < p$) proporcionan la predicción lineal óptima con las r variables del conjunto de variables $\mathbf{X}$.

Ejemplo A.1 *Sea La matriz de varianzas-covarianzas poblaciones:*

$$\Sigma = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

Calcularemos los componentes principales y la proporción de varianza total de cada componente.

Entonces, se determinará los autovalores de la matriz $\mathbf{\Sigma}$:

$$|\mathbf{\Sigma} - \lambda I| = 0$$

Entonces:

$$\begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 5-\lambda \end{vmatrix} = 0 = -\lambda^3 + 11\lambda^2 - 36\lambda + 36 = 0$$

De ello se tiene que los autovalores son: $\lambda_1 = 6$, $\lambda_2 = 3$ y $\lambda_3 = 2$ (se ha ordenado de mayor a menor). Ahora se procede a calcular los autovectores:
Para $\lambda_1 = 6$, entonces:

$$\begin{pmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

De ello se tiene que $x = y$ y $z = 2x$, de manera que su autovector será:

$$\begin{pmatrix} x \\ x \\ 2x \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Para $\lambda_2 = 3$, entonces:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

De ello se tiene que $y = -z$ y $x = -z$, de manera que su autovector será:

$$\begin{pmatrix} -z \\ -z \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

Finalmente, para $\lambda_3 = 2$, entonces:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

De ello se tiene que $z = 0$ y $x = -y$, de manera que su autovector será:

$$\begin{pmatrix} -y \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

Ahora, se normalizará los autovalores. Entonces, para el vector $a_1 = (1, 1, 2)$ se tiene que:

$$||a_1|| = \sqrt{1^2 + 1^2 + 2^2} = \sqrt{6}$$

De modo que el vector normalizado sería:

$$\hat{a_1} = \frac{1}{\sqrt{6}}(1, 1, 2)$$

Ahora para el vector $a_2 = (-1, -1, 1)$ se tiene que:

$$||a_1|| = \sqrt{-1^2 + -1^2 + 1^2} = \sqrt{3}$$

De modo que el vector normalizado sería:

$$\hat{a_1} = \frac{1}{\sqrt{3}}(-1, -1, 0)$$

Finalmente, para $a_3 = (-1, 1, 0)$ se tiene que:

$$||a_1|| = \sqrt{-1^2 + 1^2 + 0^2} = \sqrt{2}$$

De modo que el vector normalizado sería:

$$\hat{a_1} = \frac{1}{\sqrt{2}}(-1, 1, 0)$$

Las componentes principales ($\hat{y}_i$) serán:

$$\hat{y}_1 = \hat{a}_1 \cdot X = \frac{1}{\sqrt{6}}(x_1 + x_2 + 2x_3)$$

$$\hat{y}_2 = \hat{a}_3 \cdot X = \frac{1}{\sqrt{3}}(-x_1 - x_2 + x_3)$$

$$\hat{y}_3 = \hat{a}_3 \cdot X = \frac{1}{\sqrt{2}}(-x_1 + x_2)$$

Ahora, se calculará la varianza total será:

$$var(\Sigma) = tr(\Sigma) = 11$$

La proporción para cada componente:

$$\frac{var(\hat{y}_1)}{var(\Sigma)} = \frac{\lambda_1}{11} = \frac{6}{11} = 0{,}54 = 54\,\%$$

$$\frac{var(\hat{y}_2)}{var(\Sigma)} = \frac{\lambda_2}{11} = \frac{3}{11} = 0{,}27 = 27\,\%$$

$$\frac{var(\hat{y}_1)}{var(\Sigma)} = \frac{\lambda_1}{11} = \frac{2}{11} = 0{,}18 = 18\,\%$$

Apéndice B

Apéndice. Análisis factorial

El propósito fundamental del análisis factorial es explicar y describir, si es posible, las relaciones de covarianza entre un conjunto de variables observables en función a otras variables no observables o *latentes*, denominadas factores (Johnson et al., 2014; Peña, 2002). En ese sentido, una suposición importante del análisis factorial es que no es posible observar estos factores *latentes* directamente; las variables dependen de los factores pero también están sujetas a errores aleatorios (Kent et al., 1979).
El análisis factorial establece una estrecha relación con el análisis de componentes principales pero, a ello también, existen fuertes diferencias. Como lo destaca Peña (2002), "los componentes principales se construyen para explicar las varianzas, mientras que los factores se construyen para explicar las covarianzas o correlaciones". De esta manera, de acuerdo a Johnson et al. (2014), el modelo factorial está motivado por el siguiente argumento:

> "Basically, the factor model is motivated by the following argument: Suppose variables can be grouped by their correlations. That is, suppose all variables within a particular group are highly correlated among themselves, but have relatively small correlations with variables in a different group. Then it is conceivably be that each group of variables represents a single underlying construct, or factor, that is responsible for the observed correlations (Johnson et al., 2014, 481)".

En efecto, la correlación supone un aspecto central a explicar dentro del análisis factorial, con lo cual la fortaleza esta determinada por la alta correlación entre las variables

observables. Además, "los componentes principales constituyen una herramienta descriptiva, mientras que el análisis factorial presupone un modelo estadístico formal de generación de la muestra dada"(Peña, 2002, 355).
Siguiendo con Johnson et al. (2014):

> "Factor analysis can be considered an extension of principal component analysis, both can be viewed as attempts to approximate the covariance matrix Σ. However, the approximation based on the factor analysis model is more elaborate. The primary question in factor analysis is whether the data are consistent with a prescribed structure (Johnson et al., 2014, 482)".

Lo que en esencia estrecha más los lazos entre las dos, se encuentra en la forma de extracción de los factores latentes, pues allí se considera determinante el rol que los componentes principales desarrollado por Hotelling (1933).
Suponiendo un vector $\mathbf{X}$ de variables observadas, de dimensión $(p \times 1)$ y que son elementos de una población. El modelo factorial establece que el vector de datos observados se genera mediante la relación:

$$\mathbf{X} = \mu + \Lambda\mathbf{f} + \mathbf{u} \tag{B.1}$$

Donde $\mathbf{f}$ es un vector de variables no observables de dimensión $(m \times 1)$, a este vector también se le denomina vector de factores que constituyen los *factores comunes*. Se supone que este vector sigue una distribución $N_m(\mathbf{0}, \mathbf{I})$; es decir, los factores presentan media cero y son independientes entre sí y siguen una distribución normal. Λ es una matriz de constantes desconocidas de $(p \times m)$, donde $(m < p)$. Esta matriz se denomina matriz de *carga* (*loading* en inglés), o *saturaciones* y contiene los coeficientes que describen como los factores $\mathbf{f}$ afectan a las variables observadas $\mathbf{X}$. Por su parte,$\mathbf{u}$ es un vector de residuos o perturbaciones no observables de dimensión $(p \times 1)$, también llamado *factores específicos*; se supondrá que este vector sigue una distribución $N_p(\mathbf{0},\Psi)$, donde Ψ es una matriz diagonal y muestra que las perturbaciones están incorrelacionadas con los factores[1].
Considerando una muestra aleatoria simple de n individuos generado por el modelo

[1]Con estas tres hipótesis se deduce que μ es la media de las variables $\mathbf{X}$ ya que los factores y las perturbaciones tienen media cero. Además, $\mathbf{X}$ tiene distribución normal debido a que es la suma de variables normales. Ahora, se denomina Σ a la matriz de covarianzas $\mathbf{X} \backsim N_p(\mu, \Sigma)$

factorial expresado en (B.1), entonces cada dato $\mathbf{x}_{ij}$ puede escribirse como:

$$x_{ij} = \mu_j + \lambda_{j1} f_{1i} + + \lambda_{jm} f_{mi} + u_{ij} \quad i = 1, , n \;\; j = 1, \cdots, p$$

o también

$$\begin{aligned} x_1 &= \mu_1 + \lambda_1 1 f_1 + \lambda_1 2 f_2 + \cdots + \lambda_1 n f_n + u_1 \\ x_2 &= \mu_2 + \lambda_2 1 f_1 + \lambda_2 2 f_2 + \cdots + \lambda_2 n f_n + u_1 \\ &\vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ x_p &= \mu_p + \lambda_p 1 f_1 + \lambda_p 2 f_2 + \cdots + \lambda_p n f_n + u_p \end{aligned} \tag{B.2}$$

(B.3)

El subíndice i representa al individuo de la variable j. Los efectos de los factores sobre x_{ij} son el producto de los coeficientes $\lambda_{j1}, \cdots, \lambda_{jm}$, que depende de la relación entre cada factor y la variable j[2], por los valores de los m factores en el elemento muestral i, $f_{1i}, \cdots, f_{mi}$. Expresando en términos matriciales:

$$\mathbf{X} = \mathbf{1}\mu' + \mathbf{F}\Lambda' + \mathbf{U} \tag{B.4}$$

Aquí $\mathbf{1}$ es un vector de $(n \times 1)$ de unos, $\mathbf{F}$ es una matriz de $(n \times m)$ que contiene los m factores para los n elementos de la población. $\mathbf{\Lambda'}$ es la transpuesta de la matriz de las cargas $(m \times p)$ que presenta los coeficientes constantes que relacionan las variables con sus factores; finalmente, $\mathbf{U}$ es la matriz $(n \times p)$ de las perturbaciones.
se asume

$$E(\mathbf{F}) = \mathbf{0}(n \times 1), \qquad V(\mathbf{F}) = E(\mathbf{F}\mathbf{F}') = \mathbf{I}(n \times n)$$

$$E(\mathbf{U}) = \mathbf{0}(p \times 1), \qquad V(\mathbf{U}) = E(\mathbf{U}\mathbf{U}') = \mathbf{\Psi}(p \times p) = \begin{bmatrix} \psi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \psi_p \end{bmatrix}$$

y que $\mathbf{F}$ y $\mathbf{U}$ son independiente,

$$Cov(\mathbf{U}, \mathbf{F}) = E(\mathbf{U}\mathbf{F}') = \mathbf{0}(p \times n)$$

[2]que son los mismos para todos los elementos de la muestra

Estos supuestos en relación a la expresión (B.4) constituye el *modelo factorial ortogonal* (Johnson et al., 2014). Como todos los factores están incorrelacionados entre sí y, además, los factores comunes están estandarizados para tener una varianza de 1. A veces es conveniente suponer, como se evidenció anteriormente, que F y U, y por lo tanto X, tienen una distribución multinormal.

Notar que

$$x_i = \mu_i + \sum_{i=1}^{k} \lambda_{ii} f_i + u_i, \qquad i = 1, \cdots, p$$

Así que[3]:

$$\sigma_{ii} = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_{ii},$$

Entonces, las varianza de $\mathbf{X} = \sigma_{ii}$ puede ser dividido en dos partes.

$$\underbrace{\sigma_{ii}}_{Var(x_i)} = \underbrace{\lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{in}^2}_{comunalidad} + \underbrace{\psi_i}_{varianza\ especifica}$$

Primero,

$$h_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2 \tag{B.5}$$

que es la parte de la varianza de x_i que representada la parte común con las otras variables vía a los *factores comunes*, esto es lo denominado *comunalidad*. Como lo describe Kent et al. (1979), $\lambda_{ij}^2 = C(x_i, f, i)$ representa la extensión para la cual x_i depende del j-ésimo factor común.

[3]Del modelo factorial ortogonal (B.4) implica que la estructura de varianza de $\mathbf{X}$:

$$\begin{aligned}(\mathbf{X} - \mu)(\mathbf{X} - \mu)' &= (\mathbf{\Lambda F} + \mathbf{U})(\mathbf{\Lambda F} + \mathbf{U})' \\ &= (\mathbf{\Lambda F} + \mathbf{U})((\mathbf{\Lambda F})' + \mathbf{U}') \\ &= \mathbf{\Lambda F}(\mathbf{\Lambda F})' + \mathbf{U}(\mathbf{\Lambda F})' + \mathbf{\Lambda F U}' + \mathbf{U U}'\end{aligned}$$

Así qué

$$\begin{aligned}\mathbf{\Sigma} = Cov(\mathbf{X}) &= E(\mathbf{X} - \mu)(\mathbf{X} - \mu)' \\ &= \Lambda E(\mathbf{F F}')\Lambda' + E(\mathbf{U F}')\mathbf{\Lambda}' + \Lambda(\mathbf{F U}') + E(\mathbf{U U}') \\ &= \Lambda\Lambda' + \Psi\end{aligned}$$

Como $Cov(\mathbf{U}, \mathbf{F}) = E(\mathbf{U}, \mathbf{F}') = \mathbf{0}$. Además, de acuerdo a la expresión (B.1), se tiene:

$$(\mathbf{X} - \mu)\mathbf{F}' = (\mathbf{\Lambda F} + \mathbf{U})\mathbf{F}' = \Lambda \mathbf{F F}' + \mathbf{U F}'$$

entonces:

$$Cov(\mathbf{X}, \mathbf{F}) = E(\mathbf{X} - \mu)\mathbf{F}' = \Lambda E(\mathbf{F F}') + E(\mathbf{U F}') = \Lambda$$

Por otro lado, ψ_{ii} es la parte de la varianza determinada de manera única por el factor u_i, que explica la variabilidad en x_i que no son compartidas por las otras variables, esta es denominada *varianza especifica.*

La validez del modelo de factor k se puede expresar en términos de una condición simple en Σ. Después:

$$\Sigma = \Lambda\Lambda' + \mathbf{\Psi}$$

Notar que $\Sigma\Sigma'$ es una matriz simétrica de rango $m < p$. Esta matriz como se describió, contiene la parte común del conjunto de las variables y depende de las covarianzas entre las variables y los factores. Por otro lado, Ψ es una matriz diagonal que contiene la parte especifica de cada variable (Peña, 2002).

El modelo factorial asume que las $p+p(p-1)/2 = p(p+1)/2$ varianzas y covarianzas para $\mathbf{X}$ pueden ser reproducidas a partir de las pm cargas factoriales λ_{ij} y las p varianzas específicas ψ_i. Cuando $m = p$, cualquier matriz de covarianza Σ puede ser reproducida exactamente como $\Lambda\Lambda'$, así que Ψ puede ser una matriz de ceros. Sin embargo, cuando $(m < p)$ el análisis factorial es más útil. En este caso, el modelo provee $p(p+1)/2$ parámetros en Σ.

B.0.1. Unicidad del modelo

Como se ha descrito, la matriz de cargas Λ y los factores $\mathbf{f}$ son no observables. Esto implica un problema de indeterminación. De acuerdo a Peña (2002), dos representaciones $(\Lambda, \mathbf{f})$ y $(\Lambda^*, \mathbf{f}^*)$ serán equivalente si

$$\Lambda\mathbf{f} = \Lambda^*\mathbf{f}^*$$

La situación descrita conduce a dos tipos de indeterminación. La primera se da cuando en un conjunto de datos puede explicarse con la misma precisión con factores incorrelacionados o correlacionados. La segunda implica que los factores no quedan determinados de manera única (Peña, 2002). Desarrollando la primera, se tiene que si $\mathbf{H}$ es cualquier matriz no singular, entonces $\mathbf{X} = \mu + \mathbf{\Lambda f} + \mathbf{U}$ puede escribirse como

$$\mathbf{X} = \mu + \mathbf{\Lambda H H^{-1} f} + \mathbf{U} \tag{B.6}$$

Si $\Lambda^* = \Lambda\mathbf{H}$ es la nueva matriz de carga, y $\mathbf{f}^* = \mathbf{H^{-1}f}$ a los nuevos factores, entonces

$$\mathbf{X} = \mu + \mathbf{\Lambda}^* + \mathbf{f}^* + \mathbf{U}$$

En esta nueva expresión, $\mathbf{f}^*$ tiene una distribución $N(\mathbf{0}, \mathbf{H^{-1}f})$, lo que indica que dichos factores están correlacionados. Análogamente, si los factores correlaciones y $\mathbf{f} \backsim N(\mathbf{0}, \mathbf{V_f})$, entonces siempre se puede encontrar una expresión equivalemnte a las variables mediante el modelo factorial.

En efecto, si $\mathbf{A}$ es una matriz tal que $\mathbf{V_f} = \mathbf{AA'}$, entonces esta matriz siempre existirá si $\mathbf{V_f}$ es definida positiva. Por tanto, $\mathbf{A^{-1}V_f(A^{-1})'}$ se puede escribir como

$$\mathbf{X} = \mu + \mathbf{\Lambda(A)(A^{-1})f} + \mathbf{U}$$

Considerando $\Lambda^* = \Lambda A$ como la nueva matriz de coeficientes de los factores y $\mathbf{f}^* = \mathbf{A^{-1}f}$ como los nuevos factores, el modelo es equivalente es otro con factores incorrelacionados. Notar que esta indeterminación se resuelve bajo la hipótesis que los factores están incorrelacionados.

La otra indeterminación se presente cuando, dado $\mathbf{H}$ una ortogonal, entonces

$$\mathbf{X} = \mu + \mathbf{\Lambda f} + \mathbf{U}$$

Y

$$\mathbf{X} = \mu + \mathbf{(\Lambda H)'((\Lambda H)} + \mathbf{U}$$

son indestinguibles. En ambos los factores están incorrelacionados, con matriz de covarianzas igual a la identidad. En tal sentido, el modelo factorial está indeterminado ante rotaciones. De acuerdo a Peña (2002), esta indeterminación se resuelve imponiendo restricciones sobre los componentes de la matriz de carga.

B.0.2. Número máximo de factores comunes

Siguiendo a Cuadras (1996) se tiene el modelo $\mathbf{X} = \mathbf{\Lambda f} + \mathbf{U}$, entonces el número m de factores comunes está limitado por un valor máximo m_a que se puede determinar considerando que existen $p(p-1)/2$ correlaciones diferentes y $p \cdot m$ cargas o saturaciones. Pero si $\mathbf{\Lambda}$ es una matriz factorial con factores $\mathbf{f}$, también lo es $\mathbf{\Lambda T}$, con factores $\mathbf{F}^* = \mathbf{T'F}$, donde T es una matriz ortogonal. Debido a que $\mathbf{TT'} = \mathbf{I}$ se introducen $m(m-1)/2$ restricciones y el número de parámetros libres de $\mathbf{A}$ entonces será $m \cdot m - m(m-1)/2$. Ahora, el número de correlaciones menos el número de parámetros libres será

$$d = p(p-1)/2 - [p \cdot m - m(m-1)/2] = \frac{1}{2}[(p-m)^2 - p - m] \qquad \text{(B.7)}$$

Si $d = 0$ se obtiene una ecuación de segundo grado, cuya solución prueba que

$$m \leq m_a = \frac{1}{2}(2p + 1 - \sqrt{8p + 1})$$

Por ende, si $m > m_a$ el modelo factorial estará sobredimensionado, debido a que hay más cargas libres que correlaciones. En el caso que $m = m_1$ el modelo es determinado y se puede encontrar Λ a partir de $\mathbf{R}$. Ahora, un caso interesante es $m < m_a$, ya que en ese caso es posible plantear la estimación estadística de Λ, donde $d > 0$, es el número de grados de libertad del modelo. El número máximo m^* de factores comunes está en función de p.

B.0.3. Estimación de los parámetros

Como lo menciona Kent et al. (1979), en la práctica solo se observa la matriz $\mathbf{B}$, cuya información es resumida por la media muestral $\bar{x}$ y la matriz de covarianzas $\mathbf{S}$. El problema que surge es cómo estimar Λ y Ψ y, por consecuencia, $\Sigma = \Lambda\Lambda' + \Psi$ con $\mathbf{S}$. En definitiva, se encontrará los estimadores $\hat{\Lambda}$ y $\hat{\Psi}$ que satisfagan la que

$$\Sigma'\Psi^{-1}\Sigma \quad es\ diagonal \tag{B.8}$$

o en su defecto

$$\Sigma'\mathbf{B}^{-1}\Sigma \quad es\ diagonal \quad \mathbf{D} = diag(\sigma_{11}, \cdots, \sigma_{pp}) \tag{B.9}$$

Donde en cada caso, los elementos de la diagonal están escritos de forma descendente. Tanto (B.8), como (B.9) son invariantes de escala, excepto a posibles cambios en el signo de las columnas. La matriz Λ es entonces, completamente determinada por las restricciones. Notar que si el número de factores es igual a uno ($k = 1$), las restricciones son irrelevantes. Asimismo, si algún ψ_{ii} es igual a cero, entonces la restricción (B.8) no puede ser usada.

Entonces, la ecuación a estimar será:

$$\mathbf{S} = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi} \tag{B.10}$$

Dada una estimación $\hat{\Lambda}$, entonces es natural establecer:

$$\hat{\psi_{ii}} = s_{ii} - \sum_{j=1}^{k} \hat{\lambda_{ij}}^2, \quad i = 1, ..., p \tag{B.11}$$

de modo que las ecuaciones diagonales en (B.10) siempre se cumplen exactamente. Entonces, sólo se considera las estimaciones para las que se satisface (B.11) y $\hat{\psi_{ii}} \geq 0$. En base a ello, $\hat{\Sigma} = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$ se obtiene

$$\hat{\sigma}_{ii} = \sum_{j=1}^{k} \hat{\lambda}_{ij}^2 + \hat{\psi}_{ii}$$

Por lo que la expresión (B.11) es equivalente a la condición siguiente

$$\hat{\sigma}_{ii} = s_{ii}, \quad i = 1, ..., p \tag{B.12}$$

Entonces, pueden ocurrir tres casos entorno a $\mathbf{S} = \hat{\Sigma}\hat{\Sigma}' + \Psi$ considerando los valores de s

- Si $s < 0$, entonces $\mathbf{S} = \hat{\Sigma}\hat{\Sigma}' + \Psi$ tendrá más parámetros que ecuaciones. Por lo que, en general, se pueden encontrar infinitas soluciones por Λ y Ψ y, por tanto, el modelo factorial no está bien definido.

- Si $s = 0$, entonces $\mathbf{S} = \hat{\Sigma}\hat{\Sigma}' + \Psi$ puede ser resuelto para Λ y Ψ, sujeto, claro, bajo las restricciones (B.10) o (B.11). En ese sentido, el modelo contiene tantos parámetros como Σ, y por lo tanto no ofrece ninguna simplificación de la suposición original de que $V(x) = \Sigma$.

- Si $s > 0$, entonces $\mathbf{S} = \hat{\Sigma}\hat{\Sigma}' + \Psi$ tendrá más ecuaciones que parámetros. Entonces no es posible resolver exactamente el problema en términos de Λ y Ψ. Pero sí, a través de soluciones aproximadas.

Si Σ se desvía significativamente de una matriz diagonal, entonces se puede considerar un modelo factorial, y el problema inicial es estimar las cargas factoriales λ_{ij} y la varianza específica ψ_i. En ese sentido, existen dos métodos de estimación que gozan de gran popularidad, el de *componentes principales* y el *método de máxima verosimilitud*; estos dos métodos puedes ser rotados para la simplificación de los factores (Johnson et al., 2014).

B.0.4. Estimación de los factores

Siguiendo a Peña (2002), existen dos procedimientos para estimar los factores: (i) Bartlett, en el cual se supone que el vector de valores de los factores para cada

observación es un parámetro a estimar; (ii) aletorio, que supone que las variables siguen un nivel de aleatoriedad.

Bartlett: los factores como parámetros

Suponiendo que se tiene una vector $(p \times 1)$ de variables en el individuo i, $\mathbf{x}_i$, que tiene una distribución normal, con media $\Lambda \mathbf{f}_i$, donde $\mathbf{f}_i$ es un vector $(m \times 1)$ de factores para el elemento i en la muestra, además que la matriz de covarianzas es ψ

$$\mathbf{x}_i \backsim N_p(\Lambda \mathbf{f}_i, \psi)$$

Los parámetros $\mathbf{f}_i$ puede ser estimados a través de máxima verosimilitud. El estimador resultante es el de mínimo cuadrados generalizados, dado por

$$\hat{\mathbf{f}}_i = (\hat{\Lambda}' \hat{\psi}^{-1} \hat{\Lambda})^{-1} \hat{\Lambda}' \hat{\psi}^{-1} \mathbf{x}_i \tag{B.13}$$

Si se conoce Λ, el modelo factorial

$$\mathbf{x}_i = \Lambda \mathbf{f}_i + \mathbf{u}_i$$

el cual es un modelo de regresión con variable dependiente $\mathbf{x}_i$, las variables explicativas son las columnas de Λ y los parámetros $\mathbf{f}_i$. Como $\mathbf{u}_i$ se distribuye como $N(\mathbf{N}, \psi)$, se tiene que utilizar mínimos cuadrados generalizados, lo que deriva a la expresión (B.13).

Los factores como variables aleatorias

Ahora lo que se supone es que los factores son variables aleatorias, y el objetivo es encontrar un predictor lineal que minimice el error cuadrático medio de predicción. Si $\mathbf{f}_i$ son los valores de los factores en el individuo i y $\mathbf{x}_i$ es el vector de variables observadas, el vector $(\mathbf{f}_i, \mathbf{x}_i)$ tendrá distribución normal multivariante y el objetivo es encontrar: $E\ [\mathbf{f}_i | \mathbf{x}_i]$. Entonces, se obtiene

$$E\ [\mathbf{f}_i | \mathbf{x}_i] = E[\mathbf{f}_i] + Cov\ (f_i, x_i) Var(x_i)^{-1} (x_i - E\ (\mathbf{x}_i))$$

Como $E\ [\mathbf{f}_i] = 0$ y las covarianzas entre los factores y las variables son los términos de la matriz de carga, y suponiendo que las variables de media cero y los parámetros conocidos, es posible escribir:

$$\hat{\mathbf{f}}_i = E\ [\mathbf{f}_i | \mathbf{x}_i] = \Lambda' \mathbf{V}^{-1} \mathbf{x}_i$$

la expresión (B.0.4) es el predictor lineal de los factores sobre los datos. Por lo que, Λ' representa las covarianzas entre factores y variables y $\mathbf{V}$ es la covarianzas entre variables. Asimismo, la ecuación puede escribirse como

$$\hat{\mathbf{f}}_i = (\mathbf{I} + \Lambda'\psi^{-1}\lambda)^{-1}\Lambda'\psi^{-1}\mathbf{x}_i \tag{B.14}$$

Si en las ecuaciones (B.0.4) y (B.14) se sustituye los valores teóricos por los estimados, se obtiene un vector $\mathbf{f}_i$ que representa la estimación del valor de los m factores en el individuo i. Aplicando sucesivamente a los n datos muestrales, se obtiene los valores de los factores para esos n individuos $(\mathbf{f}_1, ..., \mathbf{f}_n)$, donde cada $\mathbf{f}_i$ es un vector $(m \times 1)$ (Peña, 2002).

B.0.5. Uso de la matriz de correlación en la estimación

Considerando estimaciones de $\Lambda = \Lambda_x$ e $\Psi = \Psi_x$ que son invariantes de escala. Para ello es conveniente considerar el escalamiento separadamente de las relaciones entre las variables. Si $\mathbf{Y} = \mathbf{HXD_s^{-1/2}}$, donde $\mathbf{D_s} = \mathbf{diag(s_{11}, ..., s_{pp})}$ denota las variables estandarizadas, por lo que

$$\sum_{r=1}^{n} y_{rj} = 0 \quad y \quad \frac{1}{n}\sum_{r=1}^{n} y_{rj}^{z} = 1 \quad j = 1, ..., p$$

Entonces, $\mathbf{Y}$ tendrá una matriz de carga factorial estimada $\hat{\Lambda}_y = \mathbf{D}_s^{-1/2}\hat{\Lambda}_x$ y la estimación de la varianza especifica $\hat{\Psi}_y = \mathbf{D}_s^{-1}\Psi_x$, y $\mathbf{S} = \hat{\Lambda}\hat{\Lambda}' + \Psi$ puede ser escrita en términos de la matriz de correlaciones de $\mathbf{X}$ como

$$\mathbf{R} = \hat{\Lambda}_y\hat{\Lambda}_y' + \hat{\Psi}_{y'} \tag{B.15}$$

Notar que $\hat{\psi}_{ii} = s_{ii} - \sum_{j=1}^{k}\hat{\lambda}_{ij}^2$ se convierte en

$$\hat{\psi}_{yii} = 1 - \sum_{j=1}^{k}\hat{\lambda}_{yij}^2 \qquad i = 1, ..., p$$

De este modo Ψ_x, ya no es un parámetro del modelo, sino una función de Ψ_y. Sin embargo, $\mathbf{R}$ contiene p menos parámetros libres que $\mathbf{S}$, por lo que s la diferencia entre el número de ecuaciones y el número de parámetros en $\mathbf{R} = \hat{\Lambda}_y\hat{\Lambda}_y' + \hat{\Psi}_{y'}$ todavía está dada por $s = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k)$. Las ecuaciones p para los estimados de los parámetros de escala están dadas por $\hat{\sigma}_{ii} = s_{ii}$.

Ya que en la práctica es la relación entre las variables lo que interesa más que su escalamiento. los datos a menudo se resumen por **R** en lugar de **S**. Las estimaciones de escala dadas en $\hat{\sigma}_{ii} = s_{ii}$ no se mencionan explícitamente, y las cargas factoriales estimadas y las varianzas específicas se presentan en términos de las variables estandarizadas.

B.0.6. Rotación de los factores

La obtención de la matriz factorial no es más que el primer paso del análisis factorial. La restricciones:

$$\Lambda' \Psi^{-1} \Lambda \quad es \quad diagonal$$

y

$$\Lambda' \mathbf{D}^{-1} \Lambda \quad es \quad diagonal \quad \mathbf{D} = diag(\sigma_{11}, ..., \sigma_{pp})$$

Sobre las cargas factoriales es una conveniencia matemática para hacer que las cargas factoriales sean únicas; sin embargo, puede complicar el problema de la interpretación. Generalmente, la matriz obtenida no define unos factores interpretables. La interpretación de las cargas factoriales es la más sencilla si cada variable tiene una gran *carga* en, como máximo, un factor, y si todas las cargas factoriales son grandes y positivas o cercanas a cero, con pocos valores intermedios; es decir, se trata de conseguir que unas cargas sean altas en desmedro de otras, que serán bajas; como consecuencia, se destaca la influencia de los factores sobre las variables observables. Esto implica que para esa interpretación es necesario *rotar* los factores hacia otros, sin afectar la validez del modelo (Cuadras, 1996; Johnson et al., 2014).

Rotación Varimax

El método varimax de rotación ortogonal fue propuesto por Kaiser (1958). Su razón es proporcionar ejes con algunas cargas grandes y tantas cargas cercanas a cero como sea posible. Esto se logra mediante una maximización iterativa de una función cuadrática de las cargas; es decir, maximizar la varianza de los coeficientes que defines los efectos de cada factor sobre las variables observadas.

Si δ_{ij} son los coeficientes de la matriz de carga asociados al factor j en las $i = 1, ..., p$ ecuaciones después de la rotación y δ_j al vector que es la columna j de la matriz de carga después de la rotación. Lo que se busca es que la variabilidad de los coeficientes al

cuadrado de este vector sea la máxima posible. De acuerdo a ello, se toma el cuadrado de los coeficientes, pues lo que importa es su valor absoluto. Si, $\bar{\delta}_{.j} = \sum \delta_{ij}^2/p$ es la media de los cuadrados de los componentes del vector δ_j, entonces la variabilidad para el factor j es:

$$\frac{1}{p}\sum_{i=1}^{p}(\delta_{ij}^2 - \bar{\delta}_{.j})^2 = \frac{1}{p}\sum_{i=1}^{p}\delta_{ij}^4 - (1/p)^2(\sum_{i=1}^{p}\delta_{ij}^2)^2 \tag{B.16}$$

y el criterio es maximizar la suma de las varianzas para todos los factores, que esta definido por:

$$q = (1/p)\sum_{j=1}^{m}\sum_{i=1}^{p}\delta_{ij}^4 - (1/p)^2\sum_{J=1}^{m}(\sum_{i=1}^{p}\delta_{ij}^2)^2 \tag{B.17}$$

Si Λ es la matriz de carga estimada inicialmente, entonces el problemas es encontrar una matriz ortogonal $\mathbf{M}$, tal que la matriz δ este dada por:

$$\delta = \Lambda\mathbf{M}$$

De acuerdo a esto, los coeficientes δ_{ij} vienen dados por

$$\delta_{ij} = \lambda_i'\mathbf{m}_j$$

Siendo λ_i' la fila i de la matriz Λ Y $\mathbf{m}_j$ la columna j de la matriz $\mathbf{M}$ que se busca. Ahora, los términos de la matriz $\mathbf{M}$ se obtienen de la expresión (B.0.6) respecto a cada uno de sus términos m_{ij}, teniendo en cuenta la restricción de ortogonalidad $\mathbf{m}_i'\mathbf{m}_i = 1$ y $\mathbf{m}_i'\mathbf{m}_j = 0$ $(i \neq j)$. Este resultado es la rotación *varimax* (Peña, 2002).

Apéndice C

Apéndice. Relación entre el análisis factorial y el análisis de componentes principales

Como lo remarca Kent et al. (1979), tanto el análisis factorial como el análisis de componentes principales es un intento por explicar un conjunto de datos en un número de dimensiones más pequeño. Aquí algunos aspectos resaltantes:

- El análisis de componentes principales es una transformación de los datos, donde no existe una suposición sobre la forma de la matriz de covarianza. Por su parte, el análisis factorial supone un modelo bien definido, donde se presentan un conjunto de supuestos de los valores subyacentes. En el caso de no cumplirse, el análisis factorial puede entregar resultados erróneos.

- El énfasis del análisis de componentes principales está en una transformación de las variables observadas a los componentes; mientras que en el análisis factorial el énfasis está en una transformación de los factores subyacentes a las variables observadas.

- Cuando las varianza especifica es asumida como cero, el análisis de componentes principales y el análisis factorial son equivalentes. Por lo tanto, si el modelo factorial se mantiene y si las varianzas específicas son *pequeñas*, entonces se espera que el análisis de componentes principales y el análisis factorial den resultados similares. No obstante, si las variaciones específicas son muy grandes, estas serán

absorbidas por todos los componentes principales, tanto retenidos como rechazados, mientras que el análisis factorial hace una provisión especial para ellos.

Apéndice D

Apéndice. Descomposición en valores singulares

Fallas-Monge et al. (2021) definen la descomposición en valores singulares (SVD, por sus siglas en inglés) como un tipo de factorización que, generaliza el concepto de valores propios para cualquier matriz rectangular, mediante una estrategia similar a la diagonalización de matrices. De hecho, la descomposición espectral de la que se derivan los valores propios y sus vectores propios se definen a partir de matrices simétricas, por tanto, cuadradas.

La ventaja que surge de este tipo de factorización es que no necesita de imponer restricciones sobre la dimensión de la matriz. Así, los vectores singulares forman una base del subespacio atravesado por las columnas o las filas, definiendo una matriz de proyección sobre él (Kanatani, 2021). De esta manera la SVD utiliza dos bases diferentes para diagonalizar una matriz, donde sus bases son ortonormales determinadas por los vectores singulares a derecha e izquierda (Buffo and Verdiell, 2004). Lo interesante de ello es que las transformaciones ortogonales preservan normas y ángulos; en particular, permiten preservan las longitudes de los vectores del error, totalmente inevitables en los cálculos numéricos.

La DVS surge alrededor de 1875 mediante los trabajos de Eugenio Beltrami y Camille Jordan (Fallas-Monge et al., 2021). Por su parte, Buffo and Verdiell (2004) destaca que su "generalización a espacios de dimensión infinita fue desarrolladas en el contexto de ecuaciones integrales por Smith y Weyl", en 1907 y 1912, respectivamente. No obstante, su uso se popularizó a partir de la década del 60, cuando Golub y otros autores

calcularon de manera numérica y empezaron a usarla como herramienta para muchos algoritmos (Trefethen and Bau III, 1997; Golub and Van Loan, 2013).

Se parte, pues, definiendo A como una matriz de dimensión $m \times n$, $(m \geq n)$, entonces, una descomposición en valores singulares (SVD) es una factorización del tipo $A = U\Sigma V'$ con U una matriz con columnas ortogonales de $m \times m$, V es una matriz ortogonal cuadrada de $n \times n$, estas matrices son unitarias y por ende invertibles (rango completo); mientras que Σ es una matriz diagonal de $m \times n$.

Para cualquier matriz A de $m \times n$, la matriz $A^T A$ de $n \times n$ es simétrica y, por tanto, es diagonalizada ortogonalmente, además de ser semidefinida positiva. Así, los autovalores de $A^T A$ son reales y no negativos (≥ 0). Entonces, para cualquier autovector u, se tiene:

$$A^T Au = \lambda u \tag{D.1}$$

operando de manera que:

$$u^T A^T Au = \lambda u^T u$$

Ello indica que $||Au||^2 = \lambda||u||^2$, por tanto, $\lambda \geq 0$. Luego, $\sqrt{\lambda_i}$, si λ_i, $i = 1, ..., n$, son los autovalores de $A^T A$. A su vez, las raices cuadradas de los autovalores de $A^T A$ son los valores singulares de A, denominados mediantes $\sigma_1, ..., \sigma_n$. A esto, es una convención acomodar los valores singulares de modo que $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n$. Además, los autovectores de $A^T A$ son los vectores singulares a derecha v_i de A.

Si el rango de A es $r < n$ (lo mismo para $A^T A$), entonces $\sigma_{r+1} = \sigma_{r+2} = ... = \sigma_n = 0$, lo que deriva a que Σ de $n \times n$ tendrá una forma en bloques: un bloque diagonal D de dimensión $r \times r$ incluso pudiendo ser $r = n$, a su derecha una matriz de ceros $\mathbf{0}$ de $r \times n - r$, abajo a la izquierda otra matriz de ceros $\mathbf{0}$ de $n - r \times r$ y, finalmente, el bloque diagonal de ceros $\mathbf{0}$ de $n - r - r$.

$$\mathbf{\Sigma} = \begin{pmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

Donde

$$D = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{pmatrix}$$

La matriz D es no singular, donde los elementos de la diagonal son no-negativos y estos pueden ser ordenados en forma no creciente. El número de elementos en la diagonal D es igual al rango de A.

Los autovalores no nulos de la matriz AA^T son los σ_i^2, y los autovectores asociados son los vectores singulares a izquierda u_i de A. Cualquier otro complemento ortogonal a $U_1, ..., U_n$ puede elegirse como autovectores asociados al autovalor 0 (Buffo and Verdiell, 2004).

Además, si $A = U\Sigma V^T$, entonces $A^T = V\Sigma^T U^T$, a esto se pre-multiplica A obteniendo $AA^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T$, entonces:

$$AA^T U = U\Sigma\Sigma^T \tag{D.2}$$

La igualdad expresada en (D.2) da una diagonalización de AA^T. Si se denota u_j a las columnas de U, entonces $AA^T u_j = \sigma_j u_j$, para $j = 1, ..., m$. Por ende, las columnas de U corresponden a los vectores propios de AA^T, asociados a los valores propios σ_j de AA^T, que a su vez corresponden a los valores propios de $A^T A$. Por lo que en resumen, los valores propios de las matrices AA^T y $A^T A$ coinciden, y corresponden a los valores singulares de la matriz A (Fallas-Monge et al., 2021). De acuerdo a lo anterior, para determinar los valores singulares de una matriz A es conveniente trabajar o con $A^T A$, o AA^T, según cuál sea la matriz de menor tamaño, con el objetivo de tomar el polinomio característico de menor grado.

Si se procede con la matriz AA^T, entonces las últimas $m - r$ columnas de U se toman del núcleo de A^T; ello implica determinar $m - r$ vectores ortonormales a partir del sistema $A^T u = \mathbf{0}$. De esta manera: $u_1, ..., u_r, u_{r+1}, ..., u_m$ forman una base ortonormal para $\mathbb{R}^m$. Estos son los denominados vectores singulares por la izquierda. Así que siguiendo a Fallas-Monge et al. (2021), se tiene en cuenta:

- Para construir la matriz V si ya se tiene la matriz U, considerando $A = U\Sigma V^T$ entonces:

  $$A^T U = V\Sigma$$

  Así, $A^T u_j = \sigma_j v_j$. Por lo que, para $\sigma_j \neq 0$, $1 \leq j \leq \text{mín}\{m, n\}$, se obtiene:

  $$v_j = \frac{1}{\sigma_j} A^T u_j \tag{D.3}$$

Los vectores v_j expresados en (D.3) son ortogonales debido a que los vectores u_j también son ortogonales.

- Si se tiene la matriz V, la matriz U es construida como:

$$u_j = \frac{1}{\sigma_j} A v_j$$

Esto derivado de $A = U\Sigma V^T$, lo que puede expresarse como $AV = U\Sigma$; luego, $Av_j = \sigma_j u_j$, donde $1 \leq j \leq \text{mín}\{m, n\}$.

Si existen $k = min(m, n)$ valores singulares de A y $r = r(A)$ (rango de A), entonces, como se mostró anteriormente, los valores singulares son positivos. Si $r < k$, por lo tanto los $(k - r)$ valores singulares son ceros; luego, los valores singulares son únicos, aunque no así los vectores singulares. Notar que si $m = n$ y los σ_j son distintos dos a dos, entonces la matriz Σ es única. En dicho caso, las columnas de U y las columnas de V están unívocamente determinados salvo un escalar complejo de módulo uno (Buffo and Verdiell, 2004; Soldevilla Olivares, 2004).
En muchas aplicaciones, como lo menciona Fallas-Monge et al. (2021), es común utilizar una versión simplificada de la SVD, de manera tal que:

$$A = U\Sigma V^T$$

Puede ser representada como

$$\begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \begin{bmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} V_r & V_{n-r} \end{bmatrix}^T \tag{D.4}$$

Operando

$$\begin{bmatrix} U_r\Sigma_r & \mathbf{0} \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r\Sigma_r V_r^T \tag{D.5}$$

Entonces, dado que A es una matriz de $m \times n$, la descomposición reducida en valores singulares de A está dada por $U_r\Sigma_r V_r^T$, donde $r = r(A)$, que corresponde a los valores singulares no nulos de A que, a su vez, son los valores propios de no nulos de AA^T y A^TA. Esta forma reducida pierde ciertas propiedades a favor de la utilidad de tener matrices más pequeñas; como por ejemplo, las matrices U y V en una descomposición completa son cuadradas, en su forma reducida no necesariamente lo son (Fallas-Monge et al., 2021).

Siguiendo con Fallas-Monge et al. (2021), U_r puede verse como una matriz de 1 bloque-fila y 1 bloque-columna, y Σ_r como una matriz de 1 bloque-fila y r bloques-columna, entonces:

$$U_r\Sigma_r = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1r} \\ u_{21} & u_{22} & \cdots & u_{2r} \\ & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{mr} \end{bmatrix}_{1\times 1} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}_{1\times r}$$

O lo que es lo mismo

$$U_r \otimes \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}_{1\times r}$$

Entonces, se llega a

$$\begin{bmatrix} \sigma_1 u_{11} & \sigma_2 u_{12} & \cdots & \sigma_r u_{1r} \\ \sigma_1 u_{21} & \sigma_2 u_{22} & \cdots & \sigma_r u_{2r} \\ & \vdots & \ddots & \vdots \\ \sigma_1 u_{n1} & \sigma_2 u_{n2} & \cdots & \sigma_r u_{mr} \end{bmatrix}_{1\times r} = \begin{bmatrix} \sigma_1 u_1| & \sigma_2 u_2| & \ldots & |\sigma_r u_r \end{bmatrix}_{1\times r}$$

Como $V_r = (v_1\ v_2\ \cdots\ v_r)$, entonces V^T puede ser vista como una matriz por bloques de r bloques-fila y 1 bloque-columna. Es decir:

$$V_r^T = \begin{bmatrix} v_1^T \\ v_3^T \\ \vdots \\ v_r^T \end{bmatrix}_{r\times 1}$$

Por lo tanto,

$$U_r\Sigma_r V_r^T = \begin{bmatrix} \sigma_1 u_1| & \sigma_2 u_2| & \ldots & |\sigma_r u_r \end{bmatrix}_{1\times r} \cdot \begin{bmatrix} v_1^T \\ v_3^T \\ \vdots \\ v_r^T \end{bmatrix}_{r\times 1}$$

Finalmente, se llega a la siguiente expresión, que es conocida como la SVD reducida:

$$U_r\Sigma_r V_r^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T \tag{D.6}$$

Ejemplo D.1 *Considerando la siguiente matriz se obtendrá la SVD*[1]*:*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

La factorización SVD de $\mathbf{A}$ será $\mathbf{A} = \mathbf{U\Sigma V^T}$, donde $\mathbf{A}$ es una matriz de 2×3, $\mathbf{U}$ es una matriz de 2×2, $\mathbf{V}$ una matriz 3×3 y, finalmente, $\mathbf{\Sigma}$ es una matriz de 2×3.

$$\mathbf{M} = \mathbf{A^T A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Aquí se observa que M es una matriz simétrica, por tanto es diagonalizable. Asimismo, los valores propios de M son todos mayores o iguales que cero.
Entonces, considerando $\mathbf{M} = \mathbf{A^T A}$, se tienen que los valores propios son $3, 1, 0$. De acuerdo a lo anterior, se tiene que la base o los sub espacios propios (cuya dimensión son uno) son:

$$V_3 : \{(1, 2, -1)\}, \text{ Base ortonormal de } V_3 : \{(\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{-\sqrt{6}}{6})\}$$

$$V_1 : \{(1, 0, 1)\}, \text{ Base ortonormal de } V_1 : \{(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2})\}$$

$$V_0 : \{(1, -1, 1)\}, \text{ Base ortonormal de } V_0 : \{(\frac{\sqrt{3}}{3}, \frac{-\sqrt{3}}{3}, \frac{-\sqrt{3}}{3})\}$$

Ahora, los valores singulares de A serán: $\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{1}$. De esto se obtiene la matriz $\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Con los elementos anteriores, se obtiene la matriz $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{-\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \end{pmatrix}$$

[1] *Ejemplo extraído de lasmatematicas.es, disponible en* https://www.youtube.com/watch?v=N8xQ3iPn3lA.

Considerando los vectores de la matriz anterior se calculan los vectores que componentes **U**:

$$\frac{1}{\sigma_1}\mathbf{A}\begin{pmatrix}\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ \frac{-\sqrt{6}}{6}\end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix}0 & 1 & -1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ \frac{-\sqrt{6}}{6}\end{pmatrix} = \begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}$$

$$\frac{1}{\sigma_2}\mathbf{A}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ 0\\ \frac{\sqrt{2}}{2}\end{pmatrix} = \frac{1}{\sqrt{1}}\begin{pmatrix}0 & 1 & -1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ 0\\ \frac{\sqrt{2}}{2}\end{pmatrix} = \begin{pmatrix}\frac{-\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}$$

De los cálculos anteriores se obtiene la matriz U:

$$\mathbf{U} = \begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}$$

Obteniendo la descomposición:

$$\mathbf{A} = \begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}\begin{pmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3}\\ \frac{\sqrt{6}}{3} & 0 & \frac{-\sqrt{3}}{3}\\ \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3}\end{pmatrix}$$

Ejemplo D.2 *Considerando otro ejemplo con una variación respecto al anterior. Entonces de la siguiente matriz se obtendrá la SVD:*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}$$

La factorización SVD de $\mathbf{A}$ será $\mathbf{A} = \mathbf{U\Sigma V^T}$, donde $\mathbf{A}$ es una matriz de 3×2, $\mathbf{U}$ es una matriz de 3×3, $\mathbf{V}$ una matriz 2×2, ambas son ortogonales y, finalmente, $\mathbf{\Sigma}$ es una matriz de 3×2 de la misma dimensión que $\mathbf{A}$.

$$\mathbf{M} = \mathbf{A^T A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Aquí se observa que M es una matriz simétrica, por tanto es diagonalizable. Asimismo, los valores propios de M son todos mayores o iguales que cero.
Entonces, considerando $\mathbf{M} = \mathbf{A^T A}$, se tienen que los valores propios son 3 y 1. De acuerdo a lo anterior, se tiene que la base o los sub espacios propios (cuya dimensión son uno) son:

$$V_3 : \{(1, -1)\}, \text{ Base ortonormal de } V_3 : \{(\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\}$$

$$V_1 : \{(1, 1)\}, \text{ Base ortonormal de } V_1 : \{(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})\}$$

Ahora, los valores singulares de A serán: $\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{1}$. De esto se obtiene la matriz $\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Con los elementos anteriores, se obtiene la matriz $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Considerando los vectores de la matriz anterior se calculan los vectores que componentes **U**:

El primer vector será:

$$\frac{1}{\sigma_1}\mathbf{A}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{-\sqrt{2}}{2}\end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix}0 & 1\\ 1 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{-\sqrt{2}}{2}\end{pmatrix} = \begin{pmatrix}\frac{-\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ \frac{\sqrt{6}}{6}\end{pmatrix}$$

El segundo vector:

$$\frac{1}{\sigma_2}\mathbf{A}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix} = \frac{1}{\sqrt{1}}\begin{pmatrix}0 & 1\\ 1 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix} = \begin{pmatrix}\frac{\sqrt{2}}{2}\\ 0\\ \frac{\sqrt{2}}{2}\end{pmatrix}$$

Se observa que se tiene dos columnas de **U** que es 3×3, luego faltaría una última columna; para obtenerlo se considera los dos vectores columnas correspondientes a las columnas de **U** y se obtendrá un tercer vector, de forma que dichos tres vectores constituyan una base ortonormal de $\mathbb{R}^3$.

Entonces:

$$u_1 = (-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}) \quad u_2 = (\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2})$$

lo que se busca es un vector u_3, tal que el conjunto formado por $\{u_1, u_2, u_3\}$ sea una base ortogonal. Para ello se puede considerar Gram-Schmidt, o se puede buscar un vector e_3 que sea ortogonal a u_1, u_2 de modo que $u_3 = \frac{e_3}{||e_3||}$.

Calculando el subespacio ortogonal de:

$$< u_1, u_2 >^{\perp}$$

Entonces considerando el subespacio generado por u_1 y u_2, lo cual sería:

$$u_1 = (-\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}) = \sqrt{6}(\frac{-1}{6}, \frac{1}{3}.\frac{1}{6}) = \frac{\sqrt{6}}{6}(-1, 2, 1)$$

Entonces, para u_1 podemos tomar un proporción, el cual sería: $(-1, 2, 1)$. Para u_2 sería:

$$u_2 = (\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}) = \frac{\sqrt{2}}{2}(1, 0, 1)$$

Por tanto, $< u_1, u_2 >$ sería:

$$< u_1, u_2 >=< (-1, 2, 1), (1, 0, 1) >$$

Luego, considerando lo anterior

$$< u_1, u_2 >^{\perp}=< (-1, 2, 1), (1, 0, 1) >^{\perp}=< (-1, -1, 1) >$$

Por lo que considerando e_3 como:

$$e_3 =< (-1, -1, 1) >$$

el cual es un vector ortogonal a u_1 y u_2. Ahora calculando u_3 como:

$$u_3 = \frac{(-1, -1, 1)}{\sqrt{3}} = (\frac{-1}{3}, \frac{-1}{3}, \frac{1}{3}) = (\frac{-\sqrt{3}}{3}, \frac{-1\sqrt{3}}{3}, \frac{\sqrt{3}}{3})$$

Por lo que este vector u_3 es de modulo 1, de forma el conjunto formado $\{u_1, u_2, u_3\}$ es una base ortonormal de $\mathbb{R}^3$.
Completando la matriz U:

$$\mathbf{U} = \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \end{pmatrix}$$

Con ello tenemos la SVD:

$$\mathbf{A} = \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{-\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Apéndice E

Apéndice. La Descomposición en Valores singulares (SVD), el Análisis de Componentes principales (ACP) y el Análisis factorial

Existe una estrecha relación entre la Descomposición en Valores Singulares y el Análisis de Componentes Principales. Y es que, la SVD es usada para generalizar el ACP, tal como se verá más adelante.

La relación existente entre SVD y el ACP ha sido utilizada en psicología durante muchos años, desde Tucker (1958) con su artículo *an inter-battery method of factor analysis*; en el cual se postula el siguiente modelo:

$$\mathbf{X}_{p1} = \mu_1 + \Lambda_1 \mathbf{Z} + \Gamma_1 \mathbf{y}_1 + \mathbf{e}_1$$
$$\mathbf{X}_{p2} = \mu_2 + \Lambda_2 \mathbf{Z} + \Gamma_2 \mathbf{y}_2 + \mathbf{e}_2$$

Donde μ_1, μ_2 son los vectores de medias, $\Gamma_1, \Gamma_2, \Lambda_1, \Lambda_2$ son matrices de coeficientes, $\mathbf{Z}$ es un vector de variables latentes comunes a $\mathbf{x}_{p1}$ y $\mathbf{x}_{p2}$. Por su parte, $\mathbf{y}_1$ e $\mathbf{y}_2$ son vectores de variables latentes específicos a $\mathbf{x}_{p1}$ y $\mathbf{x}_{p2}$, y finalmente, $\mathbf{e}_1$ y $\mathbf{e}_2$ son los errores. De esta manera, Tucker (1958) ajusta el modelo utilizando la SVD de la matriz de correlaciones, y señala que su procedimiento es equivalente a encontrar combinaciones

lineales de dos baterías que tienen covarianza máxima.

De la misma manera, de acuerdo a Jolliffe (2002), Bretherton et al. (1992) y Wallace et al. (1992) popularizaron este análisis bajo el nombre de análisis de descomposición en valores singulares (SVD), en el cual calcularon el análisis de componentes a través de una SVD de la matriz de covarianzas.

Como se vio anteriormente, la SVD usa las mismas transformaciones unitarias que PCA realizadas en las filas o columnas de X (usando las columnas o filas como espacios de características). De esa manera, los valores singulares de la SVD son las raíces cuadradas de los autovalores del PCA. Partiendo de la expresión:

$$X = U\Sigma V^T \tag{E.1}$$

Se observa que

$$XX^T = U\Sigma^2 U^T$$

Además que

$$X^T X = V\Sigma^2 V^T$$

Ahora, de acuerdo a Kroonenberg (2008) la matriz de datos $X = x_{ij}$ puede ser descompuesta en los scores (o puntuaciones) para los individuos y los loading (cargas) para las variables, es decir,

$$x_{ij} = \hat{x_{ij}} + e_{ij} = \sum_{s=1}^{S} u_{is} f_{js} + e_{ij} \tag{E.2}$$

o matricialmente

$$X = \hat{X} + E = UF^T + E \tag{E.3}$$

Aquí s representa un componente, mientras que S corresponde al total de componentes extraídos en una solución, $\hat{X}$ es denominada la base estructural de la matriz de datos $U = (u_{is})$ es una matriz $I \times S$ normalizada de los coeficientes de los individuos sobre los S componentes, $F = (f_{is})$ es una matriz $J \times S$ de los coeficiente de las variables sobre las S componentes, y E es la matriz $I \times J$ de los residuos. Además, las columnas tanto de A como de F refieren a la misma componente.

Ahora, f_{is}, que es los coeficientes de las variables, puede ser descompuesto como el producto de los coeficientes normalizados de las variables, v_{js}, y la variabilidad de los

componente, g_{ss}, de manera que (E.2) puede expresarse como:

$$x_{ij} = \sum_{s=1}^{S} u_{is} f_{js} + e_{ij} = \sum_{s=1}^{S} u_{is}[g_{ss} v_{js}] + e_{ij}$$

o en notación matricial:

$$X = UGV^T + E \tag{E.4}$$

Esta en efecto, es la SVD; donde $V = (v_{js})$ es una matriz normalizada $J \times S$ que contiene los puntajes de las variables sobre los S componentes, y $G = (g_{ss})$ es una matriz $S \times S$ diagonal, es decir que los elementos $g_{ss'} = 0$ si $s \neq s'$, que contiene los valores singulares de la matriz X. Los componentes de U son los vectores singulares izquierdos, mientras que los componentes de V son los vectores singulares derechos.
Tal como se vio en la expresión (D.6), la SVD puede ser escrita como la suma de matrices de rango uno, de la forma:

$$X = \sum_{s=1}^{S} g_{ss}(u_s \otimes v_s) + E \tag{E.5}$$

Donde g_{ss} corresponde al s-ésimo valor singular, u_{ss} es el vector singular izquierdo y v_{ss} es el vector singular derecho. Ahora, cada producto $u_s \otimes v_s$ genera una matriz del mismo tamaño que la matriz de datos $(n \times p)$, pero de rango uno. Ahora, para obtener la representación completa de la matriz de datos original, X, se requiere S aproximaciones, cada una ponderada por su respectivo valor singular, g_{ss}. Como G es una matriz diagonal, donde $g_{ss'} = 0$ si $s \neq s'$; entonces, la combinación $u_{is}v_{js'}$ tienen peso igual a cero, o lo que es lo mismo, la combinación de los componentes no contribuyen a la imagen estructural, $\hat{X}$.
Expandiendo la expresión (E.5):

$$X = g_{11}(u_1 \otimes v_1) + g_{22}(u_2 \otimes v_2) + \cdots + g_{ss}(u_s \otimes v_s) + E \tag{E.6}$$

En la SVD, u_s, es la s-ésima columna de A, que está exclusivamente vinculado a b_s, la s-ésima columna B, como el producto $(u_s \otimes v_s)$, siendo cero cuando $s \neq s'$. Esto básicamente debido a que u_s y v_s son referidos al mismo componente y la misma interpretación. Además, el número de columnas de A y B son necesariamente iguales, ya que $g_{ss'} = 0$.

Apéndice F

Programación del modelo de Tucker 3 en R

```
    install.packages("openxlsx")

#Cambiando directorio
setwd("XXXXXX")
library(readxl)
library(openxlsx)

ruta_archivo <- "XXXXX/BD1.xlsx"
nombres_hojas <- excel_sheets(ruta_archivo)
lista <- lapply(nombres_hojas, function(hoja) {
  read_excel(ruta_archivo, sheet = hoja)
})
lista <- setNames(lapply(nombres_hojas, function(hoja) {
  read_excel(ruta_archivo, sheet = hoja)
}), nombres_hojas)

N=length(lista)
Xa=as.matrix(lista[[1]])

for (i in 2:N){
```

```
21    Xa=cbind(Xa,lista[[i]])
22  }
23
24  matricesT=lapply(lista,t)
25  Xb=matricesT[[1]]
26
27  for (i in 2:N){
28    Xb=cbind(Xb,matricesT[[i]])
29  }
30
31  Xc=as.vector(as.matrix(lista[[1]]))
32
33  for (i in 2:N){
34    Xc=rbind(Xc,as.vector(as.matrix(lista[[i]])))
35  }
36
37  Xa=as.data.frame(Xa)
38  Xb=as.data.frame(Xb)
39
40  row.names(Xc)=NULL
41  Xc=as.data.frame(Xc)
42
43  matriciaciones=list(Xa,Xb,Xc)
44
45  names(matriciaciones)=c("Xa","Xb","Xc")
46  k=length(names(lista))
47  j=ncol(lista[[1]])
48  indcond=paste(row.names(lista[[1]]),names(lista)[1],sep="-")
49  varcond=paste(names(lista[[1]]),names(lista)[1],sep="-")
50  indvar=paste(row.names(lista[[1]]),names(lista[[1]])[1],sep="-")
51
52  for(i in 2:j){
```

```
  indvar=c(indvar,paste(row.names(lista[[1]]),
  names(lista[[1]])[i],sep="-"))
}

for(i in 2:k){
  indcond=c(indcond,paste(row.names(lista[[1]]),
  names(lista)[i],sep="-"))
  varcond=c(varcond,paste(names(lista[[1]]),
  names(lista)[i],sep="-"))
}

names(matriciaciones$Xa)=varcond
names(matriciaciones$Xb)=indcond
names(matriciaciones$Xc)=indvar
row.names(matriciaciones$Xc)=names(lista)

#write.xlsx(matriciaciones,"matriciaciones.xlsx")

#Preprocesamiento
X = matriciaciones
ind=row.names(X$Xa)
#variables
var=row.names(X$Xb)
#ocasiones
cond=row.names(X$Xc)
#Extrayendo los nombres de los encabezados
indcond=names(X$Xb)
varcond=names(X$Xa)
indvar=names(X$Xc)
#Creando las matrices para el preprocesamiento
Xa=as.matrix(X$Xa)
Xb=as.matrix(X$Xb)
```

```
Xc=as.matrix(X$Xc)
#Ubicando el tamaño de las matrices Xa, Xb y Xc
na=nrow(Xa)
nb=nrow(Xb)
nc=nrow(Xc)
#Cantidad de columnas de Xa
Xacol=ncol(Xa)
#Considerando la cantidad de filas de la matriz Xa
I=na
# Port es una matriz que sirve para normalizar
Port=diag(c(rep(1,I)))-(1/I)*c(rep(1,I))%*%t(c(rep(1,I)))
#Aquí Xac ya está normalizada
Xac=Port%*%Xa
#Transponiendo los valores de Xac (ya que las siguientes formas
#son solo matriciaciones de la primera normalizacion)
Xbac=t(Xac[,1:nb])

for (i in seq((nb+1),nb*nc,nb)){
  Xbac=cbind(Xbac,t(Xac[,i:(i+(nb-1))]))
}

for(j in 0:(nb-1)){
  D=Xac[,(j+1)]
    for(i in (j + seq((nb+1),nb*nc,nb))){
      D=cbind(D,Xac[,i])
    }
    if(j==0) {
      Xcac=t(D)
      next
    }
    Xcac=cbind(Xcac,t(D))
}
```

```
#Ahora wj y wk son las desviaciones para la normalización final
#(esto es tal cual la formula)

Wj=diag(1/sqrt(diag(Xbac%*%t(Xbac))/(na*nc)))
Wk=diag(1/sqrt(diag(Xcac%*%t(Xcac))/(na*nb)))
Ya=Xac%*%kronecker(Wk,Wj)
Yb=t(Ya[,1:(1+(nb-1))])
for (i in seq((nb+1),nb*nc,nb)){
  Yb=cbind(Yb,t(Ya[,i:(i+(nb-1))]))
  for(j in 0:(nb-1)){
    D=Ya[,(j+1)]
    for(i in (j + seq((nb+1),nb*nc,nb))){
      D=cbind(D,Ya[,i])
    }
    if(j==0) {
      Yc=t(D)
    next
    }
  }
  Yc=cbind(Yc,t(D))
}
#Ahora creando un nuevo data frame con los valores ya normalizados
Xac=as.data.frame(Xac)
Xbac=as.data.frame(Xbac)
#Renombrando
row.names(Xcac)=NULL
Xcac=as.data.frame(Xcac)
row.names(Xac)=ind
row.names(Xbac)=var
row.names(Xcac)=cond
names(Xac)=varcond
names(Xbac)=indcond
```

```
149 names(Xcac)=indvar
150 # generando una nueva lista con los resultados finales normalizados
151 X1=list(Xac,Xbac,Xcac)
152 names(X1)=c("Xac","Xbac","Xcac")
153 #write.xlsx(X1,"Preprocesamiento.xlsx")
154
155 class(X1)="matriciacion"
156 lista.original = lista
157
158 ################################################################
159 #Estimación
160 N=100
161 #Como se usa nuevamente el nombre lista, chancaré con los datos X1
162 lista = X1
163 #Extrayendo los nombres de las listas
164 varcond=names(lista[[1]])
165 indcond=names(lista[[2]])
166 indvar=names(lista[[3]])
167 cond=row.names(lista[[3]])
168 var=row.names(lista[[2]])
169 ind=row.names(lista[[1]])
170 #Convirtiendo a matrices
171 X1=as.matrix(lista[[1]])
172 X2=as.matrix(lista[[2]])
173 X3=as.matrix(lista[[3]])
174 #Calculando la cantidad de filas de las matrices -1
175 #Individuos
176 P=nrow(X1)-1
177 #Variables
178 Q=nrow(X2)-1
179 #Ocasiones
180 R=nrow(X3)-1
```

```
#Creando los valores de las tres dimensiones del cubo
c1=c(1:R)
c2=c(1:Q)
c3=c(1:P)
#Se crea un vector concatenando repetidamente Q-veces
#los elementos del vector R
a1=c(rep(R,Q))

#Se crea un vector concatenando repetidamente P-veces
#los elementos del vector Q*R
a2=c(rep(Q*R,P))

#Se crea un vector concatenando repetidamente a1-veces
#los elementos del vector c2
s1=c(rep(c2,a1))

w1=c(rep(c1,P*Q))
w2=c(rep(s1,P))
w3=c(rep(c3,a2))
#Creando el core
Z=cbind(w3,w2,w1)

pq=w3*w2
qr=w2*w1
pr=w3*w1

#"L" Es una matriz boleana que contiene contiene TRUE
#en las posiciones donde se cumple que pq es mayor o igual a w1,
#qr es mayor o igual a w3 y pr es mayor o igual a w2.

L=pq>=w1 & qr>=w3 & pr>=w2

```

```
213 #Z es una matriz condicionada a los valores que cumplen L
214 Z=Z[L,]
215
216 SCE1=c(rep(0,N))
217 SSX1=sum(X1^2)
218
219 A=list()
220 B=list()
221 C=list()
222 G1=list()
223 G2=list()
224 G3=list()
225
226 X1est=c()
227 G1t=c()
228 G2t=c()
229 G3t=c()
230 At=c()
231 Bt=c()
232 Ct=c()
233
234 #row(Z) crea un vector que indica la fila de cada elemento en Z.
235 #tapply() aplicar una función (en este caso sum()) a los
236 #subconjuntos de un vector o matriz, determinados por un
237 #factor (en este caso row(Z))
238
239 Ncomp=as.vector(tapply(Z,row(Z),sum)) #numero de componentes
240 #Combinando Ncomo a Z
241 Z=cbind(Z,Ncomp)
242 #Buscando el min
243 maxi=min((P+1),((Q+1)*(R+1)))+min((Q+1),_
244 ((P+1)*(R+1)))+min((R+1),((P+1)*(Q+1)))-3
```

```
245
246 P+Q+R
247
248 if (maxi==P+Q+R){
249   #Esta línea filtra los datos en la matriz Z utilizando
250   #una condición La condición Z[, 4] != max(Z[, 4])
251   #selecciona las filas de la matriz Z donde el
252   #valor en la cuarta columna no es igual al valor
253   #máximo de la cuarta columna en otras palabras,
254   #elimina las filas que contienen el valor máximo en
255   #la cuarta columna.
256   #El resultado se asigna nuevamente a la variable Z
257   Z=Z[Z[,4]!=max(Z[4,]),]
258 } else {
259   #En esta línea, se filtran las filas de la matriz Z donde el valor
260   #en la cuarta columna es menor que maxi.
261   #Las filas que cumplen esta condición se mantienen en la matriz Z,
262   #mientras que las que no la cumplen se eliminan.
263   Z=Z[Z[,4]<maxi,]
264 }
265 #Cantidad de ocasiones
266 nZ=nrow(Z)
267 #Creando el vector que repite nZ-veces el valor 0
268 SCE=c(rep(0,nZ))
269 #devuelve el índice de las filas de Z si estuvieran ordenadas
270 #de manera creciente según los valores de la cuarta columna.
271 o=order(Z[,4])
272
273 #XCCC = Z[,4]
274 #write.xlsx(XCCC,"XCCC.xlsx")
275 #write.xlsx(o,"O.xlsx")
276
```

```
277 tiempo_total <- system.time({
278
279   for(i in 1:nZ){
280     Bt=svd(X2)
281     Ct=svd(X3)
282     for(j in 2:(N+1)){
283       Aant=At
284       Bant=Bt
285       Cant=Ct
286
287       G1ant=G1t
288       G2ant=G2t
289       G3ant=G3t
290
291       SSG1ant=sum(G1ant^2)
292
293       X1estant=X1est
294
295       At=svd(X1%*%kronecker(Ct$u[,1:Z[i,2:3][2]],Bt$u[,1:Z[i,2:3][1]]))
296       Bt=svd(X2%*%kronecker(Ct$u[,1:Z[i,2:3][2]],At$u[,1:Z[i,1]]))
297       Ct=svd(X3%*%kronecker(Bt$u[,1:Z[i,2:3][1]],At$u[,1:Z[i,1]]))
298
299       G1t=t(At$u[,1:Z[i,1]])%*%X1%*%
300       kronecker(Ct$u[,1:Z[i,2:3][2]],Bt$u[,1:Z[i,2:3][1]])
301
302       G2t=t(Bt$u[,1:Z[i,2:3][1]])%*%X2%*%
303       kronecker(Ct$u[,1:Z[i,2:3][2]],At$u[,1:Z[i,1]])
304
305       G3t=t(Ct$u[,1:Z[i,2:3][2]])%*%X3%*%
306       kronecker(Bt$u[,1:Z[i,2:3][1]],At$u[,1:Z[i,1]])
307
308       X1est<-(At$u[,1:Z[i,1]]%*%G1t%*%
```

```
      kronecker(t(Ct$u[,1:Z[i,2:3][2]]),t(Bt$u[,1:Z[i,2:3][1]])))

      E1=X1-X1est

      SCE1[j]=sum(E1^2)

      dif=abs(SCE1[j]- SCE1[(j-1)])

      if(dif<=1e-10) {
        A[[i]]=Aant$u[,1:Z[i,1]]
        B[[i]]=Bant$u[,1:Z[i,2]]
        C[[i]]=Cant$u[,1:Z[i,3]]

        G1[[i]]=G1ant;G2[[i]]=G2ant;G3[[i]]=G3ant;SCE[i]=SCE1[j-1]
        break
      }
      if(j==(N+1)) {
        SSG1=sum(G1t^2)

        A[[i]]=At$u[,1:Z[i,1]]
        B[[i]]=Bt$u[,1:Z[i,2]]
        C[[i]]=Ct$u[,1:Z[i,3]]

        G1[[i]]=G1t
        G2[[i]]=G2t
        G3[[i]]=G3t

        SCE[i]=SCE1[j]
        break
      }
    }
  }
```

```
})

# Tiempo total de ejecución en segundos
tiempo_total <- tiempo_total[["elapsed"]]
print(paste("Tiempo de ejecución total:", tiempo_total, "segundos"))

#write.xlsx(C,"matrizC.xlsx")#Nota: tarda mucho

SCE=round(SCE,2)
SPM=cbind(Z,SCE)
SPM=SPM[o,]# se reordena la matriz SPM utilizando los
#índices almacenados previamente en la variable o

id=1:nrow(SPM)

SPM=as.data.frame(SPM)
names(SPM)=c("P","Q","R","No de comp.","SCE")
SCEmin=tapply(SPM[,5], SPM[,4],min)
ncomp=as.numeric(names(SCEmin))

plot(ncomp,SCEmin,type="b",xlab="Suma del_
número de componentes S=(P+Q+R)",ylab="SCE",cex=.5)
print(SPM)
print(SCEmin) #este es el que se tiene que
#usar (filtra los valores minimos entre los S)

#write.xlsx(SPM,"SMPmatrix.xlsx")

length(SCEmin)->n #la cantidad de datos de SCEmin

SCE2=SCEmin[2:n]
SCE1=SCEmin[1:(n-1)]
```

```
373
374 diffS=SCE1-SCE2 #se pierde un grado de libertad
375
376 b=c()
377 nd=length(diffS)
378
379 for(i in 1:(nd-1)){
380   if(diffS[(i+1)]!=0)
381     b[i]=diffS[i]/max(c(diffS[(i+1):nd]))
382
383   if(diffS[(i+1)]==0)
384     b[i]=0
385 }
386
387 nb=names(diffS)[2:(n-1)]
388 mod=nb[b==max(b)]
389
390 id=row.names(SPM[SPM[,4]==mod & SPM[,5]==SCEmin[names(SCEmin)==mod],])
391
392 #Esta línea nos dice que fila de SPM tomar (por ejemplo la fila 42)
393 cat("Selected model with Dif-fit criterion is:",id,"\n")
394 #cat("Select the combination of components
395 #(number between 1 and", nZ,")","\n")
396
397 #j=as.integer(readLines(n = 1))
398
399 j = as.integer(4)#osea se elije la fila 42
400 jo=o[j] #tomamos el elemento de la fila 42 el cual es 65
401
402 def=as.integer(SPM[4,1:4]) #es un vector con el
403 #numero de componentes en los tres modos
404
```

```
405 p=def[1]
406 q=def[2]
407 r=def[3]
408
409 A=A[[jo]]
410 B=B[[jo]]
411 C=C[[jo]]
412
413 G1=G1[[jo]]
414 G2=G2[[jo]]
415 G3=G3[[jo]]
416
417 res=list(A,B,C,G1,G2,G3,X1,varcond,indcond,indvar,ind,var,cond)
418 names(res)=c("A","B","C","Ga","Gb","Gc","Xa","varcond",_
419 "indcond","indvar","ind","var","cond")
420
421 res
422
423 #Cálculo de medidas de ajuste del modelo de Tucker 3
424
425 AJtotal=(sum(res$Ga^2))/(sum(res$Xa^2))*100
426 AJtotal
427 contrib.compsA=(diag(as.matrix(res$Ga)%*%t(as.matrix(res$Ga)))/sum(res$Xa^2))*100
428 contrib.compsA
429 contrib.compsB=(diag(as.matrix(res$Gb)%*%t(as.matrix(res$Gb)))/sum(res$Xa^2))*100
430 contrib.compsB
431 contrib.compsC=(diag(as.matrix(res$Gc)%*%t(as.matrix(res$Gc)))/sum(res$Xa^2))*100
432 contrib.compsC
433 contrib.triadacomps=(as.matrix(res$Ga)*as.matrix(res$Ga))*(100/sum(res$Xa^2))
434 contrib.triadacomps
435
436 res=list(AJtotal,contrib.compsA,contrib.compsB,
```

```
contrib.compsC,contrib.triadacomps)

names(res)=c("AJtotal","contrib.compsA",
"contrib.compsB","contrib.compsC","contrib.triadacomps")
res

#Representaciones de las entidades de cada modo
install.packages("ade4")
install.packages("MASS")
library(ade4)
library(MASS)

#conf.mode=function(lista,x=1,y=2,s=1,mode="A",xlim=NULL,ylim=NULL)

x = 1
y = 2
s = 1
xlim=NULL
ylim=NULL

mode="A"

res=list(A,B,C,G1,G2,G3,X1,varcond,indcond,indvar,ind,var,cond)
names(res)=c("A","B","C","Ga","Gb","Gc","Xa",_
"varcond","indcond","indvar","ind","var","cond")

res

A=res$A
A
B=res$B
B
```

```
469 C=res$C
470 C
471
472 nind=nrow(A) #filas (individuos)
473 nvar=nrow(B) #columnas(variable)
474 ncond=nrow(C) #ocasiones
475
476 ncond=length(C)
477
478 Ga=res$Ga
479 Gb=res$Gb
480 Gc=res$Gc
481
482 ind=res$ind
483 var=res$var
484 cond=res$cond
485
486 varcond=res$varcond
487 indcond=res$indcond
488 indvar=res$indvar
489
490 Fa=kronecker(C,B)%*%t(Ga)
491 Fb=kronecker(C,A)%*%t(Gb)
492 Fc=kronecker(B,A)%*%t(Gc)
493
494 Ta=diag(svd(Fa)$d)
495 Tb=diag(svd(Fb)$d)
496 Tc=diag(svd(Fc)$d) #aquí se genera un problema
497 #porque la dimensión es 1, por lo que la solución no es necesario tomarlo
498
499 Amod=A%*%Ta
500 Bmod=B%*%Tb
```

```
501 Cmod=C%*%Tc
502
503 Famod=Fa%*%ginv(Ta)
504 Fbmod=Fb%*%ginv(Tb)
505 Fcmod=Fc%*%ginv(Tc)
506
507 sub="Configuracion de las entidades del modo"
508 sub=paste(sub,mode)
509 ejes="(ejes "
510 ejes=paste(ejes,"x,y,)")
511 sub=paste(sub,ejes)
512
513 if(mode=="A"){
514   Ad=A[,c(x,y)]
515   Amodd=Amod[,c(x,y)]
516   Fad=Fa[,c(x,y)]
517   Famodd=Famod[,c(x,y)]
518   if(ncol(A)==1)
519     stop("Selected mode must have for than one column")
520   s.label(Amodd,clabel=0.5,label=ind,xlim=xlim,ylim=ylim,sub=sub,csub=1.05)
521   s.arrow(s*Famodd,xax=1,yax=2,label=varcond,clabel=0.5,boxes=FALSE)
522 }
523
524 mode="B"
525
526 if(mode=="B"){
527   Bd=B[,c(x,y)]
528   Bmodd=Bmod[,c(x,y)]
529   Fbd=Fb[,c(x,y)]
530   Fbmodd=Fbmod[,c(x,y)]
531   if(ncol(B)==1) stop("Selected mode must have for than one column")
532   s.label(Bmodd,clabel=0.5,label=var,xlim=xlim,ylim=ylim,sub=sub,csub=1.0)
```

```
533   s.arrow(s*Fbmodd,xax=1,yax=2,label=indcond,clabel=0.5,boxes=TRUE,add.plot=FALSE)
534 }
535
536 mode="C"
537
538 if(mode=="C"){
539   Cd=C[,c(x,y)]
540   Cmodd=Cmod[,c(x,y)]
541   Fcmodd=Fcmod[,c(x,y)]
542   Fcd=Fc[,c(x,y)]
543   if(ncol(C)==1) stop("Selected mode must have for than one column")
544   s.label(Cmodd,clabel=0.5,label=cond,xlim=xlim,ylim=ylim,sub=sub,csub=1.25)
545   s.arrow(s*Fcmodd,xax=1,yax=2,label=indvar,clabel=0.5,boxes=FALSE,add.plot=TRUE)
546 }
547
548 #Representaciones de las trayectorias de dos modos
549 install.packages("ade4")
550 library(ade4)
551 #traj.modes=function(lista,comb="AC",x=1,y=2,s=1,xlim=NULL,ylim=NULL){
552
553   x=1
554   y=2
555   s=1
556
557   xlim=NULL
558   ylim=NULL
559
560   A=res$A
561   B=res$B
562   C=res$C
563
564   comb="AC" #Arrancamos con esto
```

```
565
566   nind=nrow(A)
567   nvar=nrow(B)
568   ncond=15#nrow(C)
569
570   Ga=res$Ga
571   Gb=res$Gb
572   Gc=res$Gc
573
574   ind=res$ind
575   var=res$var
576   cond=res$cond
577
578   varcond=res$varcond
579   indcond=res$indcond
580   indvar=res$indvar
581
582   Fa=kronecker(C,B)%*%t(Ga)
583   Fb=kronecker(C,A)%*%t(Gb)
584   Fc=kronecker(B,A)%*%t(Gc)
585
586   fac1=factor(rep(1:nind,ncond))
587
588   sub="Configuracion de la combinacion"
589   sub=paste(sub,comb)
590   ejes="(ejes ";ejes=paste(ejes,x,y,")")
591   sub=paste(sub,ejes)
592
593   comb="AC"
594   if(comb=="AC"){
595     Bd=B[,c(x,y)]
596     Fbd=Fb[,c(x,y)]
```

```
597     if(ncol(B)==1) stop("Selected modes combination cannot be computed")
598     s.traject(Fbd,fac=fac1,ord = (1:length(fac1)), xax = 1,
599     yax = 2,label = NULL, clabel = 0.8, cpoint = 1, pch = 1, grid = TRUE,
600               addaxes = TRUE,edge = FALSE, origin = c(0,0),
601               include.origin = TRUE, sub = sub, csub = 1.25,
602               possub = "bottomleft",
603               cgrid = 0.5, pixmap = NULL, contour = NULL,
604               xlim = xlim, ylim = ylim, area = NULL)
605     s.arrow(s*Bd,xax=1,yax=2,label=var,clabel=0.5,boxes=FALSE)
606     text(Fbd[,1],Fbd[,2],indcond,pos=1,cex=0.5)
607   }
608
609   fac2=factor(rep(1:nvar,ncond))
610   comb="BC"
611   if(comb=="BC"){
612     Ad=A[,c(x,y)]
613     Fad=Fa[,c(x,y)]
614     if(ncol(A)==1) stop("Selected modes combination cannot be computed")
615     s.traject(Fad,fac=fac2,ord = (1:length(fac2)), xax = 1, yax = 2,
616               label = NULL, clabel = 0.5, cpoint = 1, pch = 1,
617               xlim = xlim, ylim = ylim,
618               grid = TRUE, addaxes = TRUE, edge = FALSE,
619               origin = c(0,0), include.origin = TRUE,
620               sub = sub, csub = 1.25, possub = "bottomleft",
621               cgrid = 1, pixmap = NULL, contour = NULL, area = NULL)
622     s.arrow(s*Ad,xax=1,yax=2,label=ind,clabel=0.5,boxes=TRUE,add.plot=FALSE)
623     text(Fad[,1],Fad[,2],varcond,pos=1,cex=0.5)
624   }
625
626   fac3=factor(rep(1:nind,nvar))
627   comb="AB"
628   if(comb=="AB"){
```

```
629     Cd=C[,c(x,y)]
630     Fcd=Fc[,c(x,y)]
631     if(ncol(C)==1) stop("Selected modes combination cannot be computed")
632     s.traject(Fcd,fac=fac3,ord = (1:length(fac3)), xax = 1, yax = 2,
633               label = NULL, clabel = 0.8, cpoint = 1,
634               pch = 1,xlim = xlim, ylim = ylim,
635               grid = TRUE, addaxes = TRUE, edge = FALSE, origin = c(0,0),
636               include.origin = TRUE, sub = sub, csub = 1.25,
637               possub = "bottomleft",cgrid = 1, pixmap = NULL,
638               contour = NULL, area = NULL)
639     s.arrow(s*Cd,xax=1,yax=2,label=cond,clabel=0.5,boxes=TRUE,add.plot=FALSE)
640     text(Fcd[,1],Fcd[,2],indvar,pos=1,cex=0.5)
641   }
642
643 #Representación de los elementos del arreglo central (función: repcore)
644 #function(lista,compA,compB){
645
646     compA = 1 # iniciaré con 1
647     compB = 1
648
649     ini=res$Ga[compA,]
650     n=ncol(res$Ga)
651
652     nc=nrow(res$Gc)
653     nb=nrow(res$Gb)
654
655     ind=seq(compB,n-nc+compB,by=nc)
656     ind2=1:nc
657     texto1=""
658     texto2="C"
659     texto1=rep(texto1,nc)
660     texto2=rep(texto2,nc)
```

```
661     t=paste(texto1,ind2)
662     t=paste(t,texto2)
663     sub=paste("Comp",compA,"A, en comp",compB,"B")
664     plot(ini[ind],sub=sub,xlim=c(0.65,nc),xlab="",ylab="",las=1)
665     lines(ini[ind])
666     text(ini[ind],labels=t,pos=2)
667
668 ##################FIN#######################
```