

03061

1
29



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

**UNIDAD ACADÉMICA DE LOS CICLOS PROFESIONALES Y
DE POSGRADO DEL COLEGIO DE CIENCIAS Y
HUMANIDADES**

**INSTITUTO DE INVESTIGACIONES EN MATEMATICAS
APLICADAS Y EN SISTEMAS**

**METODO STATIS, UNA PROPUESTA PARA EL
ANALISIS DE CUBOS DE DATOS**

T E S I S
PARA OBTENER EL GRADO DE:
MAESTRA EN ESTADISTICA E
INVESTIGACION DE OPERACIONES
P R E S E N T A :
L A M A T E M A T I C A
MARTHA MARIA TELLEZ ROJO SOLIS

BAJO DE SUPERVISION DEL
DR. RUBEN HERNANDEZ CID

MEXICO, D. F.

JUNIO DE 1996

**TESIS CON
FALLA DE ORIGEN**

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

En primera instancia quisiera agradecer a la Universidad Nacional Autónoma de México, institución en la que he cursado todos mis estudios y de la que he recibido un apoyo fundamental para poder concluir esta maestría.

Al Dr. Rubén Hernández Cid, por su dedicación y paciencia puesta en este trabajo, pero sobre todo, por su permanente impulso.

A mis Profesores del Departamento de Estadística del IIMAS, Silvia, Raúl, Federico, Lety, Belén y Nacho, por haberme dado tanto, que no podría limitarlo a las palabras. Mil gracias por todo!

A los Profesores Silvia Ruiz Velasco, Raúl Rueda, Jorge Olgún e Ignacio Méndez, sinodales del examen de grado, por su cuidadosa lectura y valiosos comentarios.

A Salvador, mi apoyo y motivación fundamental.

A Carlos y Lourdes.

**A mis cuatro hermanos Carlos, Loanna,
Cristina y José Antonio.**

A mis dos grandes amigos, Andrés y Salvador.

Contenido

Introducción	3
1 Otro enfoque del Análisis de Componentes Principales.	7
1.1 Introducción	7
1.2 Espacios vectoriales y dualidad.	8
1.2.1 Algunas métricas de interés.	13
1.3 Inercia de una nube de puntos.	17
1.4 Operadores adjuntos y autoadjuntos.	20
1.5 Construcción de la nube de inercia máxima.	25
1.5.1 Descripción de individuos.	29
1.5.2 Descripción de variables.	31
1.6 Objetos suplementarios.	33
1.6.1 Individuos suplementarios.	33
1.6.2 Variables suplementarias.	35
2 El Método Statis.	37
2.1 Introducción	37
2.2 El concepto de "Estudio".	38
2.3 Comparación entre objetos.	41
2.3.1 Estructura común de los individuos a través de los estudios.	41
2.3.2 Estructura común de las variables a través de los estudios.	48
2.3.3 Comparación entre las matrices de datos.	50
2.4 Imagen euclidiana de los estudios.	50
2.5 Estudios Suplementarios.	54
2.6 Construcción del <i>Compromiso</i> .	55
2.6.1 Interpretación del <i>compromiso</i> .	59
2.7 Trayectorias de los individuos.	63
2.7.1 Información de las variables en las trayectorias de los individuos.	64
2.8 Representación de las variables en <i>estudios</i> (X_i, m, d_i) .	65

3 Aplicación del Método Statis a la dinámica demográfica en México (1930-1990).	67
3.1 Introducción.	67
3.2 El Problema.	67
3.3 Selección de métricas y coeficientes para el uso del <i>Método Statis</i> .	68
3.4 Evolución demográfica de los estados de la República.	69
3.4.1 Posición <i>compromiso</i> de los Estados.	70
3.4.2 Trayectorias de cada Estado.	72
3.5 Evolución de las variables.	74
3.6 Conclusiones.	76
Gráficas	77
Rutinas en S-Plus	97
Bibliografía	113

Introducción.

Retomando los orígenes de la Estadística, nos encontramos con que esta disciplina, considerada por muchos parte de las Matemáticas, surgió con los objetivos fundamentales de "describir" y "analizar" información, lo que generó dos grandes ramas de la Estadística, la Estadística Descriptiva y la Inferencia Estadística. Ambos enfoques fueron desarrollándose exitosamente pero con el devenir del tiempo, el "análisis" fue cobrando mayor relevancia y la "descripción" pasó a un segundo término, no obstante de ser un paso fundamental como primer acercamiento a un conjunto de datos. Conocer y entender la información disponible, permitirá encaminar el análisis de una manera más adecuada, permitiéndonos además, reconocer la factibilidad del cumplimiento de los supuestos que las técnicas analíticas demandan.

Cuando se pretende realizar un estudio sobre un conjunto de individuos a los cuales se les midieron ciertas características, las técnicas descriptivas proporcionan este primer acercamiento, las cuales interpretadas adecuadamente y con la ayuda de un profundo entendimiento de los objetivos que se persiguen, darán las pautas para iniciar un buen análisis del banco de datos en cuestión. Si se tuviera un problema en el que sólo se tuviera una variable de interés, la tarea descriptiva quedaría completamente resuelta utilizando instrumentos gráficos y medidas descriptivas básicas de la Estadística Univariada; no obstante, al estar ante un problema de la vida real, es muy frecuente que el fenómeno bajo estudio demande la medición de más de una característica para lograr un mejor entendimiento del mismo, teniendo que recurrir entonces a la Estadística Multivariada. Ambas posibilidades se han trabajado ampliamente y el problema descriptivo se encuentra prácticamente resuelto; como ejemplo pensemos en el caso multivariado en el que las variables bajo estudio son cuantitativas; para poder describir esta situación se tienen un gran número de técnicas útiles, entre las que destaca el Análisis de Componentes Principales (ACP), que permite describir

el comportamiento de un conjunto de individuos a los que se les midieron un cierto grupo de variables, de tal suerte que la información recabada puede resumirse en una matriz de tantos renglones como individuos se tengan y tantas columnas como variables se hayan medido. El atractivo fundamental de esta técnica, radica en la posibilidad representar a los objetos observados en un espacio de dimensión menor al original con la mínima pérdida de información posible, y cuando adicionalmente resulta que esta dimensión puede ser dos o máximo tres, la herramienta gráfica se vuelve sumamente poderosa y atractiva por su interpretabilidad.

Para el caso en el que los individuos fueron observados además, en diferentes momentos del tiempo, la necesidad de describir su comportamiento cobra aún mayor importancia dada la complejidad de la información, pues no basta realizar un análisis descriptivo en cada período de observación, ya que se perdería el carácter evolutivo del fenómeno. En este caso, el conjunto de datos puede resumirse en tantas matrices como ocasiones se hayan observado los sujetos, dando lugar a lo que se conoce como "Cubos de Datos".

El *Método Statis* propuesto inicialmente por Y. Escoufier y trabajado posteriormente por Christine Lavit, propone una técnica que describe este tipo de problemas, logrando generalizaciones aún mayores, como admitir la posibilidad de que el conjunto de individuos (o el de variables) cambie de período a período, que permite incluir situaciones prácticas en las que algunos individuos no puedan finalizar el estudio en su totalidad o bien, que algunas variables cobren o pierdan importancia dentro del contexto del problema y que el investigador se vea en la necesidad de modificar el conjunto de características a medir antes de terminar el período de observación. Desgraciadamente esta propuesta no ha sido muy explotada y se encuentra muy poca literatura al respecto, no obstante pensamos que vale la pena difundirla por su gran utilidad además de presentar la teoría que la sustenta de una manera mucho más accesible de cómo se propuso originalmente; asimismo creemos que implementar el método en rutinas computacionales puede contribuir también en esta tarea.

La idea fundamental del *Método Statis* radica en el uso de la estructura algebraica subyacente en los datos, en contraposición a la filosofía de imponerles modelos que los ajusten. Esto permite que pueda aplicarse a cualquier cubo de datos cuantitativos

que se disponga, pero con la consecuente limitación, que es importante no perder de vista, de que *Statis* es una técnica descriptiva que permite conocer con bastante profundidad a los individuos y características observadas a través de instrumentos gráficos, utilizando la información disponible de manera óptima, pero no puede ser extrapolada a la población que fue muestreada por no ser una técnica de tipo inferencial. Por otro lado se debe de reconocer, que si el tipo de fenómeno bajo estudio es un censo, lo que se necesita precisamente es una técnica descriptiva; además, dada la enorme complejidad de los cubos de datos, los supuestos distribucionales que demandan las escasas técnicas inferenciales que atacan este tipo de problemas, son tan restrictivos que podrán aplicarse adecuadamente sólo en muy raras ocasiones, cobrando así, mayor importancia una herramienta descriptiva.

Como se mencionó anteriormente, el *Método Statis* constituye una alternativa al análisis de cubos de datos y puede considerarse como una generalización del ACP, por lo que se diseñó la estructura de esta tesis de la manera siguiente. En el Capítulo 1 se expone detalladamente la parte técnica del método cuando el cubo bajo estudio está formado por una única matriz, con la idea de ver la forma cómo *Statis* recupera el ACP. Como este método utiliza conceptos matemáticos avanzados, se buscó hacer esta exposición desde la presentación de los mismos, no obstante la lectura de este trabajo demanda un manejo fluido del Algebra Lineal. En el Capítulo 2, utilizando los conceptos presentados en el Capítulo 1, se ataca el problema más general en el que se tienen varias matrices de datos que resumen las mediciones en cada período de observación y en el que alguno de los conjuntos, individuos o variables, puede sufrir modificaciones durante el estudio.

Finalmente, tratando de presentar un visión global de *Statis*, en el Capítulo 3 se presenta una aplicación de este método. El problema que se atacó se resume brevemente a continuación. El objetivo era entender la dinámica demográfica en México a nivel estatal y nacional que sufrió el país en el período comprendido entre 1930 y 1990, a través de seis variables de interés: número de nacimientos de hombres y de mujeres, número de defunciones generales y de niños menores de un año y número de matrimonios y divorcios. La información disponible se obtuvo de los censos nacionales presentados en el Anuario Estadístico de los Estados Unidos Mexicanos

[10] publicado por el Instituto Nacional de Geografía e Informática (INEGI) y el análisis se llevó a cabo utilizando las rutinas desarrolladas como parte de este trabajo en el lenguaje S-PLUS con licencia del Instituto de Investigación en Matemáticas Aplicadas y Sistemas (IIMAS) y que se anexan en el apéndice al final de esta tesis.

Capítulo 1

Otro enfoque del Análisis de Componentes Principales.

1.1 Introducción

El Análisis de Componentes Principales (ACP) es una técnica que nos permite estudiar el comportamiento de varias características cuantitativas medidas sobre un grupo de individuos, es decir, observaciones multivariadas que pueden resumirse en una matriz de datos con tantos renglones como sujetos se hayan observado y con un número de columnas igual a la cantidad de variables bajo estudio. Esta técnica nos brinda la posibilidad de describir relaciones entre individuos, entre variables y sobre todo, nos permite llevar a cabo esta descripción con un número menor de variables perdiendo la mínima información posible.

Sin embargo, el Análisis de Componentes Principales, se ve restringido ante aplicaciones donde surge la necesidad de estudiar simultáneamente varias matrices de datos; cuando no basta el análisis interno de cada una de ellas, sino que además se requiere describir sus relaciones. Consideremos las tres situaciones siguientes:

- *Situación 1.* Un grupo de n individuos se somete a un estudio que consiste en la medición de p variables cuantitativas en k diferentes ocasiones, con el objeto de estudiar la posible existencia de algún patrón evolutivo de estas ca-

racterísticas en los individuos bajo estudio. El resumen de esta información se puede concentrar entonces, en k matrices de dimensión $n \times p$ cada una.

- *Situación 2.* Se quiere describir el comportamiento de p variables en k diferentes grupos de individuos, donde cada uno de ellos cuenta con n_i sujetos, dando lugar a k matrices de dimensión $n_i \times p$.
- *Situación 3.* Un grupo de n individuos se somete a la medición de características agrupadas por algún criterio de seriación en k grupos con p_i variables cada uno de ellos, lo que da lugar a k matrices de dimensión $n \times p_i$.

La idea de este capítulo es presentar el Análisis de Componentes Principales de una matriz desde una perspectiva basada en su estructura algebraica, que nos permita abordar este tipo de situaciones sin requerir de supuestos distribucionales.

1.2 Espacios vectoriales y dualidad.

Cuando realizamos un Análisis de Componentes Principales clásico, disponemos de n individuos a los que les hemos medido p variables cuantitativas, de tal manera que la información obtenida de cada uno de los sujetos puede guardarse en un vector de dimensión p , teniendo un total de n de ellos. Así, estos vectores pueden interpretarse como elementos de un espacio vectorial con la forma siguiente

$$\begin{aligned} \mathbf{x}'_1 &= (x_1^1, \dots, x_p^1), \\ \mathbf{x}'_2 &= (x_1^2, \dots, x_p^2), \\ &\vdots \\ \mathbf{x}'_n &= (x_1^n, \dots, x_p^n). \end{aligned}$$

Análogamente, desde el punto de vista de las variables, esta información se puede interpretar como un conjunto de p vectores columna de dimensión n , que son elementos

de un espacio formado por las variables bajo estudio. Esto se resume como

$$\begin{aligned} \underline{x}^{1'} &= (x_1^1, \dots, x_n^1), \\ \underline{x}^{2'} &= (x_1^2, \dots, x_n^2), \\ &\vdots \\ \underline{x}^{p'} &= (x_1^p, \dots, x_n^p), \end{aligned}$$

de tal manera que el subíndice denota al número de individuo y el supraíndice se refiere a la variable correspondiente.

Así, la información observada nos lleva a pensar en dos espacios vectoriales que serán nuestro punto de partida en el análisis,

F : espacio de individuos. Inducido por n elementos de dimensión p de la forma $\underline{x}_i^j = (x_1^i, \dots, x_p^i)$ con $i = 1, \dots, n$ cuya base canónica será denotada por $\{f_1, \dots, f_p\}$.

E : espacio de variables. Generado por p elementos de dimensión n de la forma $\underline{x}^{j'} = (x_1^j, \dots, x_n^j)$ con $j = 1, \dots, p$ cuya base canónica se denotará por $\{e_1, \dots, e_n\}$.

Esto da lugar a la matriz de datos observada X de dimensión $n \times p$,

$$X = \begin{pmatrix} x_1^1 & x_1^2 & \dots & x_1^p \\ x_2^1 & x_2^2 & \dots & x_2^p \\ \vdots & \vdots & & \vdots \\ x_n^1 & x_n^2 & \dots & x_n^p \end{pmatrix}$$

que puede expresarse en términos de las variables

$$X = (\underline{x}^1, \dots, \underline{x}^p) \text{ con } \underline{x}^i \in E,$$

o bien en términos de los individuos

$$X' = (\underline{x}_1, \dots, \underline{x}_n) \text{ con } \underline{x}_j \in F.$$

Puesto que en las aplicaciones de la vida cotidiana, las características cuantitativas de interés se miden con números reales, el espacio de individuos F comúnmente será \mathbf{R}^p y el espacio de variables E será \mathbf{R}^n , lo que en general simplifica la teoría pero que desde el punto de vista técnico no representa una condición indispensable.

Así, el elemento x_j^i de la matriz observada, se puede ver como la i -ésima entrada del j -ésimo vector de E , o bien como la j -ésima entrada del i -ésimo vector de F ,

es decir, este elemento puede obtenerse a través de una función h_E que consista en extraer la i -ésima entrada de un vector en E o haciendo lo correspondiente en F

<i>variables</i>	<i>individuos</i>
$h_E : E \rightarrow \mathbf{R}$	$h_F : F \rightarrow \mathbf{R}$
$h_E(\mathbf{x}^j) = x_i^j$	$h_F(\mathbf{x}_i) = x_i^j$

Esto nos lleva a estudiar el espacio formado por las aplicaciones lineales de un espacio vectorial cualquiera a los números reales.

Definición 1.2.1 *El dual G^* de un espacio vectorial G es el espacio vectorial formado por el conjunto de aplicaciones lineales de $G \rightarrow \mathbf{R}$, denotado comúnmente por $\mathcal{L}(G, \mathbf{R})$.*

El espacio vectorial dual tiene algunas propiedades interesantes que lo hacen de gran utilidad para nosotros. Entre otras cosas, el dual comparte la estructura del espacio original; esto puede verse a través del siguiente resultado que garantiza la existencia de un isomorfismo¹ entre ellos.

Teorema 1.2.1 *Sea (G, q) un espacio vectorial de dimensión finita n , con producto escalar q , denotado por $(\underline{a}, \underline{b})_q$ con $\underline{a}, \underline{b} \in G$, y sea $\mathbf{q} : G \rightarrow \mathbf{R}$ una transformación lineal. Entonces existe un único vector $\underline{y} \in G$ tal que $\mathbf{q}(\underline{x}) = (\underline{x}, \underline{y})_q$, por esta razón $\mathbf{q}(\underline{x})$ se denotará por $\mathbf{q}_y(\underline{x})$.*

Demostración. Sea $\beta = \{\underline{q}_1, \dots, \underline{q}_n\}$ una base ortonormal de G , a través de la cual podemos expresar a un vector fijo $\underline{y} \in G$ como

$$\underline{y} = \sum_{i=1}^n (\underline{q}_i, \underline{y})_q \underline{q}_i.$$

Si definimos $h : G \rightarrow \mathbf{R}$ como $h(\underline{x}) = (\underline{x}, \underline{y})_q$, entonces h es claramente lineal (herencia de la linealidad del producto escalar q); al aplicar esta transformación a cualquier

¹Un isomorfismo entre dos espacios vectoriales G y F , es una función lineal biyectiva entre ellos.

vector $q_j \in \beta$, tenemos

$$\begin{aligned} h(q_j) &= (q_j, y)_q = (q_j, \sum_{i=1}^n (q_i, y)_q q_i)_q = (\sum_{i=1}^n (q_i, y)_q q_i, q_j)_q \\ &= \sum_{i=1}^n ((q_i, y)_q q_i, q_j)_q = \sum_{i=1}^n (q_i, y)_q (q_i, q_j)_q \\ &= \sum_{i=1}^n (q_i, y)_q \delta_{ij} = (q_j, y)_q = \alpha_y(q_j) \end{aligned}$$

como h y α_y coinciden en los elementos de la base β , podemos concluir que $h = \alpha_y$.

La unicidad de y , se deduce directamente de las propiedades elementales del producto escalar, cuando se supone la existencia de otro vector y' que cumpla que

$$\alpha_{y'}(x) = (x, y')_q = (x, y)_q, \text{ para toda } x \in G,$$

de donde se implica que $y = y'$. \square

Construyamos entonces, conjuntos $\{e_1^*, \dots, e_n^*\}$, y $\{f_1^*, \dots, f_p^*\}$ que formen una base para sus respectivos espacios duales E^* y F^* , de tal manera que los vectores observados puedan expresarse como combinaciones lineales de ellos

$$\begin{array}{ll} \text{variables} & \text{individuos} \\ e_i^* : E \rightarrow \mathbf{R} & f_j^* : F \rightarrow \mathbf{R} \\ e_i^*(x^j) = x_i^j & f_j^*(x_i) = x_i^j. \end{array}$$

Estas expresiones se resumen a través de las transformaciones a y a' que relacionan los espacios involucrados y cuyas matrices asociadas son claramente X' y X respectivamente

$$\begin{array}{ll} a : E^* \rightarrow F & a' : F^* \rightarrow E \\ a(e_i^*) = x_i & a'(f_j^*) = x^j. \end{array}$$

El siguiente diagrama permite visualizar esta información en conjunto

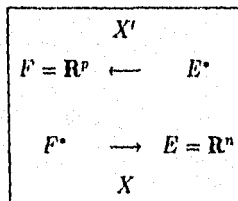


diagrama 1

Para entender las relaciones que existen entre las características y los individuos observados, necesitamos definir algunas medidas de "proximidad" o "asociación" entre objetos de la misma clase. Esto podrá hacerse a través de las métricas heredadas del producto escalar que elegimos en cada uno de los espacios, de tal manera que si queremos cambiar de medida basta cambiar de producto interior y con ello a su matriz asociada.

Consideremos el producto escalar clásico m entre individuos definido sobre F de la manera siguiente

$$m : F \times F \rightarrow \mathbf{R}$$

$$(\underline{x}_i, \underline{x}_j)_m = \underline{x}_i^* * \underline{x}_j$$

donde "*" denota el producto punto entre dos vectores, de donde se deduce la norma de un vector $\underline{x} \in F$, $\|\underline{x}\|_m^2 = (\underline{x}, \underline{x})_m$ y la distancia entre dos vectores $\underline{x}_i, \underline{x}_j$: $d_{\underline{x}_i, \underline{x}_j} = \|\underline{x}_i - \underline{x}_j\|_m$. Como se mencionó anteriormente, a este producto escalar clásico le corresponde una aplicación lineal $\mathbf{m}_y \in F^*$ que define un isomorfismo entre F y F^* cuya matriz asociada, en términos de las bases mencionadas, será llamada $M = I_p$.

Asimismo, podemos pensar en el correspondiente producto escalar euclidiano d para el espacio de variables E

$$d : E \times E \rightarrow \mathbf{R}$$

$$(\underline{x}^i, \underline{x}^j)_d = \underline{x}^{i'} * \underline{x}^j$$

con su aplicación lineal asociada $\mathbf{d}_y \in E^*$ y con matriz $D = I_n$.

Finalmente, sólo nos resta definir las métricas sobre los espacios duales F^* y E^* , de tal manera que se preserven las normas de los vectores originales; si atendemos a la asignación hecha anteriormente ($\underline{e}_i^* \mapsto \underline{x}_i$ y $f_j^* \mapsto \underline{x}^j$) y llamamos V a la matriz asociada a una métrica dada sobre F^* y W , a aquella sobre E^* , queremos entonces que:

- el i -ésimo individuo $\underline{x}_i \in F$: $\|\underline{x}_i\|_m = \|\underline{e}_i^*\|_w$ para $i = 1, \dots, n$ y
- la j -ésima variable $\underline{x}^j \in E$: $\|\underline{x}^j\|_d = \|f_j^*\|_v$ para $j = 1, \dots, p$.

las matrices quedan definidas por

$$\begin{aligned} V &= X'DX \\ W &= XMX' \end{aligned} \quad (1.1)$$

pudiendo resumirse toda esta información en el siguiente *Diagrama de Dualidad*, que no es más que la complementación del diagrama 1

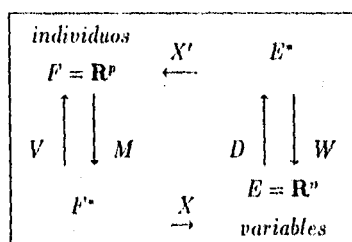


diagrama 2

1.2.1 Algunas métricas de interés.

Con el objeto de resaltar las proximidades y las diferencias entre las características bajo estudio, así como lograr mayor interpretabilidad, se propone efectuar una translación a cada vector del espacio de variables por su media aritmética ponderada por la "importancia" de cada individuo, lo que produce un corrimiento de la nube de datos que la lleva a tener por centro de gravedad, el origen. La razón por la cual hablamos de "media ponderada" evadiendo su sentido probabilístico, se debe a la filosofía del método, que busca no involucrar supuestos distribucionales. Precisemos estos conceptos.

Definición 1.2.2 La Media Ponderada de la j -ésima característica se define como

$$\bar{x}^j = \sum_{i=1}^n p_i x_i^j,$$

donde p_i refleja la "importancia" asociada al i -ésimo individuo y $\sum_{i=1}^n p_i = 1$. De donde se obtiene la Media Aritmética para el caso particular en que $p_i = 1/n \forall i = 1, \dots, n$.

Definición 1.2.3 El conjunto \mathcal{I} de individuos observados: $\{\underline{x}_1, \dots, \underline{x}_n\}$, elementos de \mathbb{R}^p , es llamado Nube de Individuos y su Centro de Gravedad se define como el vector $\underline{q} \in F$, que tiene en cada una de sus entradas, la media de la variable correspondiente, es decir

$$\underline{q} = \begin{pmatrix} \bar{x}^1 \\ \vdots \\ \bar{x}^p \end{pmatrix}.$$

Definición 1.2.4 Una matriz observada X cuya nube de individuos tiene por centro de gravedad al vector $\underline{q} \in \mathbb{R}^p$ se conoce como Matriz Centrada o Corregida por su Media.

Definición 1.2.5 La Varianza de la j -ésima característica observada se define como

$$s_j^2 = \sum_{i=1}^n p_i (x_i^j - \bar{x}^j)^2.$$

Definición 1.2.6 La Covarianza muestral entre la j -ésima y la k -ésima variable, se define por

$$s_{jk} = \sum_{i=1}^n p_i (x_i^j - \bar{x}^j)(x_i^k - \bar{x}^k).$$

Esta ponderación de los individuos se debe incluir en la métrica asignada al espacio de variables, de tal manera que la "distancia" entre ellas refleje las diferentes aportaciones hechas por cada uno de los individuos involucrados. Definamos entonces el siguiente producto escalar d_p , sobre la base $\{\underline{e}_1, \dots, \underline{e}_n\}$ de E

$$(\underline{e}_i, \underline{e}_j)_{d_p} = \begin{cases} p_i & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

cuya matriz asociada D_p es diagonal con elementos $\{p_1, \dots, p_n\}$, donde

$$(\underline{u}, \underline{v})_{d_p} = \underline{u}' D_p \underline{v}.$$

Proposición 1.2.1 d_p define un producto escalar sobre E .

Demostración. Debemos probar tres propiedades:

- Simetría. Sean $\underline{u}, \underline{v} \in E$

$$(\underline{u}, \underline{v})_{d_p} = \underline{u}' D_p \underline{v} = (\underline{u}' D_p \underline{v})' = \underline{v}' D_p \underline{u} = (\underline{v}, \underline{u})_{d_p}.$$

- Bilinealidad. Sean $\underline{u}, \underline{v}, \underline{z} \in E$

$$(\underline{u} + \underline{v}, \underline{z})_{d_p} = (\underline{u} + \underline{v})' D_p \underline{z} = (\underline{u}' + \underline{v}') D_p \underline{z} = \underline{u}' D_p \underline{z} + \underline{v}' D_p \underline{z} = (\underline{u}, \underline{z})_{d_p} + (\underline{v}, \underline{z})_{d_p},$$

$$(\lambda \underline{u}, \underline{v})_{d_p} = (\lambda \underline{u})' D_p \underline{v} = \lambda (\underline{u}' D_p \underline{v}) = \lambda (\underline{u}, \underline{v})_{d_p}.$$

- Por último probaremos que si $\underline{u} \in E$ entonces $(\underline{u}, \underline{u})_{d_p} \geq 0$. Como $\underline{u} \in E$ puede expresarse como $\underline{u} = \sum_i (\underline{\varepsilon}_i, \underline{u}) \underline{\varepsilon}_i$, usando la bilinealidad de d_p vemos que

$$\begin{aligned} (\underline{u}, \underline{u})_{d_p} &= (\sum_i (\underline{\varepsilon}_i, \underline{u}) \underline{\varepsilon}_i, \sum_i (\underline{\varepsilon}_i, \underline{u}) \underline{\varepsilon}_i)_{d_p} = \sum_i (\underline{\varepsilon}_i, \underline{u}) (\underline{\varepsilon}_i, \sum_i (\underline{\varepsilon}_i, \underline{u}) \underline{\varepsilon}_i)_{d_p} = \\ &= [\sum_i (\underline{\varepsilon}_i, \underline{u})]^2 (\underline{\varepsilon}_i, \underline{\varepsilon}_i)_{d_p} = [\sum_i (\underline{\varepsilon}_i, \underline{u})]^2 p_i, \end{aligned}$$

que es positivo y se anula únicamente cuando $(\underline{\varepsilon}_i, \underline{u}) = 0 \forall i$, es decir, cuando $\underline{u} = 0$. \square

Retomando el *Diagrama de Dualidad* (diagrama 2), vemos que la matriz $V = X'DX$ (ver 1.1) definida de esta manera para el producto escalar d , se expresa entonces como $V = X' D_p X$ para este nuevo producto escalar d_p , lo que reconstruye la matriz de varianzas y covarianzas tradicional cuando los pesos asignados a los individuos son iguales y la matriz X es centrada. Basta observar la entrada jk de la matriz V que está dada por

$$[V]_{jk} = \sum_{i=1}^n p_i x_i^j x_i^k,$$

además, la norma que define este producto escalar es tal que reconstruye la varianza muestral

$$\|\underline{x}^j\|_{d_p}^2 = (\underline{x}^j, \underline{x}^j)_{d_p} = \sum_{i=1}^n p_i (x_i^j)^2 = s_j^2,$$

entonces, si llamamos α al ángulo formado entre las variables \underline{x}^i y \underline{x}^j centradas, encontramos

$$\cos \alpha = \frac{(\underline{x}^i, \underline{x}^j)_{d_p}}{\|\underline{x}^i\|_{d_p} \|\underline{x}^j\|_{d_p}} = \frac{s_{ij}}{s_i s_j},$$

que es precisamente su correlación. Es así como se interpreta la relación entre las variables: si éstas, pensadas como vectores, están muy cercanas a la colinealidad ($\alpha = 0$ ó $\alpha = \pi$) tendremos evidencia de correlación máxima, y si se encuentran cercanas a la ortogonalidad ($\alpha = \pi/2$) habrá indicios en favor de una correlación nula.

Asimismo, al situarnos en el espacio de individuos, no es raro encontrarse ante una aplicación en la que las variables medidas se refieran a características completamente diferentes, cuyo rango de valores y variabilidades discrepen de manera importante. Por ejemplo, imaginemos que estamos midiendo el salario que recibe un trabajador y el número de hijos que tiene, si este individuo recibiera un aumento de \$1000 en el salario no representa, ni remotamente, una situación comparable a sufrir el mismo aumento en la {cantidad de hijos}, por lo que sus varianzas deben interpretarse bajo criterios diferentes. Este efecto ocasionado por la diferencia de dispersiones así como de unidades, podría enturbiar el análisis, de tal manera que es recomendable disponer de observaciones ya divididas entre su desviación estándar, donde esta distorsión se haya eliminado previamente. Estamos hablando de aplicar la siguiente transformación (lineal) a los datos centrados

$$y_i^j = \frac{x_i^j}{s_j}$$

y que equivale a asignar un producto escalar particular al espacio vectorial F . Concretamente nos referimos al producto escalar $m_{1/s}$ que se define sobre la base $\{\underline{f}_1, \dots, \underline{f}_p\}$ de la manera siguiente

$$(\underline{f}_i, \underline{f}_j)_{m_{1/s}} = \begin{cases} \frac{1}{s_j} & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

cuya matriz asociada es una diagonal con entradas $M_{1/s} = \left\{ \frac{1}{s_1}, \dots, \frac{1}{s_p} \right\}$ y da lugar a una métrica en E^* con matriz asociada $W = X M_{1/s} X'$ (ver 1.1). La demostración de que $m_{1/s}$ define un producto escalar, es completamente análoga a la Proposición 1.2.1.

Cabe señalar que el efecto de esta transformación sobre una matriz centrada es simplemente su reducción, por lo que la distancia entre dos variables nos dará como

resultado precisamente su correlación.

1.3 Inercia de una nube de puntos.

Uno de los objetivos del Análisis de Componentes Principales, es la búsqueda de proyecciones ortogonales de la nube de puntos original, que nos permitan obtener una representación gráfica (*imagen euclidiana*) de los individuos o variables bajo estudio que describa de la mejor manera posible, su comportamiento y relaciones en el sentido de pérdida mínima de información; esto nos sugiere buscar un espacio de menor dimensión a la original (*subespacio afín*), donde podamos lograrlo. Para presentar formalmente esta construcción, necesitamos introducir algunos conceptos que nos serán de utilidad.

Definición 1.3.1 *Un espacio vectorial (G, q) es llamado Euclidiano si es real, de dimensión finita y con un producto escalar q .*

Definición 1.3.2 *Un conjunto no vacío H es un Espacio Afín Euclidiano asociado a un espacio vectorial euclidiano G , si existe una aplicación a tal que*

$$\begin{aligned} a : H \times G &\rightarrow H \\ a(M, \underline{x}) &= M + \underline{x} \end{aligned}$$

que cumpla:

1. $\forall M \in H, \forall \underline{x}, \forall \underline{y} \in G$

$$\begin{aligned} M + \underline{0} &= M \\ M + (\underline{x} + \underline{y}) &= (M + \underline{x}) + \underline{y}. \end{aligned}$$

2. *Para cualesquiera dos elementos $M_i, M_j \in H$, existe un único vector $\underline{x} \in G$ tal que $M_i + \underline{x} = M_j$. Por esta razón, llamaremos al vector \underline{x} , $\underline{M_i M_j}$ y a los elementos del conjunto H los llamaremos puntos.*

Algunas observaciones que nos ayudarán son las siguientes:

- La dimensión del espacio afín es la misma que la del espacio vectorial original.
- El par formado por un punto $O \in H$ llamado origen y una base del espacio vectorial G se conoce como pareja de referencia, de tal manera que las coordenadas del punto $M \in H$ son las componentes del vector $\underline{OM} \in G$ en la base dada.

Como se comentó anteriormente, buscamos la representación de los objetos en un subespacio afín euclidiano, que capte la máxima variabilidad posible de los datos originales. La pregunta que surge entonces es, ¿cómo podemos generalizar la noción de dispersión en un espacio multidimensional y cómo llevar a cabo esta transformación? Esto será posible bajo el concepto de *inercia*, que resume la idea de proximidad de una nube de puntos con respecto a otro particular, con la ventaja adicional de que incluye la idea de la ponderación de los individuos, a través de la métrica que se haya decidido usar.

Definición 1.3.3 *El momento de inercia de una nube de individuos $\{\underline{x}_i, i = 1, \dots, n\}$ en el espacio vectorial F con métrica m respecto a un punto $\underline{a} \in F$, se define como*

$$I_{\underline{a}} = \sum_i p_i \|\underline{x}_i - \underline{a}\|_m^2.$$

Esta medida de inercia respecto a cualquier punto, puede recuperarse a través de la inercia con respecto al centro de gravedad de la nube. Este resultado se conoce como el *Teorema de Huyghens* y será de gran utilidad en este desarrollo.

Teorema 1.3.1 *La inercia de una nube de individuos $\{\underline{x}_i, i = 1, \dots, n\}$ en el espacio vectorial F con métrica m , con respecto a un punto \underline{a} puede descomponerse como*

$$I_{\underline{a}} = I_{\underline{g}} + \|\underline{g} - \underline{a}\|_m^2.$$

Demostración.

$$\begin{aligned} I_{\underline{a}} &= \sum_i p_i \|\underline{x}_i - \underline{a}\|_m^2 = \sum_i p_i (\underline{x}_i - \underline{a}, \underline{x}_i - \underline{a})_m \\ &\quad \sum_i p_i (\underline{x}_i - \underline{g} + \underline{g} - \underline{a}, \underline{x}_i - \underline{g} + \underline{g} - \underline{a})_m \end{aligned}$$

$$\begin{aligned}
I_{\underline{a}} &= \sum p_i(\underline{x}_i - \underline{q}, \underline{x}_i - \underline{q} + \underline{q} - \underline{a})_m + \sum p_i(\underline{q} - \underline{a}, \underline{x}_i - \underline{q} + \underline{q} - \underline{a})_m \\
&= \sum p_i(\underline{x}_i - \underline{q}, \underline{x}_i - \underline{q})_m + \sum p_i(\underline{x}_i - \underline{q}, \underline{q} - \underline{a})_m + \sum p_i(\underline{q} - \underline{a}, \underline{x}_i - \underline{q})_m \\
&\quad + \sum p_i(\underline{q} - \underline{a}, \underline{q} - \underline{a})_m \\
&= \sum p_i \|\underline{x}_i - \underline{q}\|_m^2 + 2 \sum p_i(\underline{x}_i - \underline{q}, \underline{q} - \underline{a})_m + \|\underline{q} - \underline{a}\|_m^2 \sum p_i \\
&= I_{\underline{q}} + 2(\sum p_i \underline{x}_i - \underline{q} \sum p_i, \underline{q} - \underline{a})_m + \|\underline{q} - \underline{a}\|_m^2 = I_{\underline{q}} + \|\underline{q} - \underline{a}\|_m^2. \quad \square
\end{aligned}$$

Queda claro entonces, que el punto más cercano a la nube utilizando el criterio de inercia mínimo, es precisamente su centro de gravedad, que para el caso de nubes centradas, como éste coincide con el origen, se puede reducir la definición 1.3.3 a

$$\psi = \sum_i p_i \|\underline{x}_i\|_m^2.$$

Este concepto se relaciona con el esquema de dualidad que habíamos discutido a través del siguiente resultado.

Teorema 1.3.2 *La inercia de una nube de individuos centrada, puede expresarse como*

$$\psi = \text{traza}(VM).$$

Demostración.

$$\psi = \sum_i p_i \|\underline{x}_i\|_m^2 = \sum_i p_i(\underline{x}_i, \underline{x}_i)_m = \sum_i p_i(\underline{x}_i' M \underline{x}_i),$$

$\underline{x}_i' M \underline{x}_i \in \mathbf{R}$ entonces $\underline{x}_i' M \underline{x}_i = \text{traza}(\underline{x}_i' M \underline{x}_i)$ por lo que

$$\begin{aligned}
\psi &= \sum_i p_i \text{traza}(\underline{x}_i' M \underline{x}_i) = \sum_i p_i \text{traza}(\underline{x}_i \underline{x}_i' M) \\
&= \sum_i \text{traza}(p_i \underline{x}_i \underline{x}_i' M) = \text{traza} \sum_i (p_i \underline{x}_i \underline{x}_i' M) = \text{traza}(VM).
\end{aligned}$$

La última igualdad se justifica por una expresión alternativa de V , $\sum p_i \underline{x}_i \underline{x}_i'$ cuya prueba es inmediata. \square

Corolario 1.3.1

$$\psi = \text{traza}(VM) = \text{traza}(WD).$$

Demostración. La demostración se deduce de las propiedades del operador *traza*.

$$\psi = \text{traza}(VM) = \text{traza}(X'DXM) = \text{traza}(XMX'D) = \text{traza}(WD). \quad \square$$

A continuación se presentan algunas observaciones interesantes

- Si $M = D_{1/s^2}$ entonces

$$\begin{aligned} \psi &= \text{traza}(VD_{1/s^2}) = \text{traza}(V(D_{1/s^2})^{1/2}(D_{1/s^2})^{1/2}) \\ &= \text{traza}((D_{1/s^2})^{1/2}V(D_{1/s^2})^{1/2}) = \text{traza}(C) = p \end{aligned}$$

donde C es la matriz de correlación.

- Si $M = I_p$, la matriz identidad deducida a partir de la base canónica, la *Inercia* recupera la traza de la matriz de varianzas y covarianzas $\psi = \text{traza}(VI) = \text{traza}(V)$.

Pero ¿cómo se relaciona este idea de la inercia de una nube, con las matrices definidas en el *Esquema de Dualidad*? La respuesta no es sencilla, pero se tratará de aclarar en la siguiente sección.

1.4 Operadores adjuntos y autoadjuntos.

A partir del siguiente teorema, se desprenden los conceptos de "operador adjunto" y "autoadjunto", que serán indispensable para el desarrollo de esta sección.

Teorema 1.4.1 Sea (G, q) un espacio vectorial de dimensión finita con producto escalar q , y sea T un operador lineal de G , es decir $T : G \rightarrow G$. Entonces existe un único operador lineal T^* en G , tal que

$$(T(\underline{x}), \underline{y})_q = (\underline{x}, T^*(\underline{y}))_q \text{ para todo } \underline{x}, \underline{y} \in G.$$

Demostración. Sea $y \in G$ y definamos $q: G \rightarrow R$ como $q(x) = (T(x), y)_q$ para todo $x \in G$. Primero probaremos que q es lineal. Sean $x_1, x_2 \in G$ y $c \in R$

$$\begin{aligned} q(cx_1 + x_2) &= (T(cx_1 + x_2), y)_q = (cT(x_1) + T(x_2), y)_q \\ &= c(T(x_1), y)_q + (T(x_2), y)_q = cq(x_1) + q(x_2). \end{aligned}$$

Por el teorema 1.2.1, sabemos que existe un único vector $y_0 \in G$ tal que $q_{y_0}(x) = (x, y_0)_q$ y para que sea igual a $(T(x), y)_q$ tenemos que definir $T^*: G \rightarrow G$ como $T^*(y) = y_0$, así

$$(T(x), y)_q = (x, T^*(y))_q,$$

de donde es fácil probar que T^* es lineal. Sean $y_1, y_2 \in G$ y $c \in R$,

$$\begin{aligned} (x, T^*(cy_1 + y_2))_q &= (T(x), cy_1 + y_2)_q = c(T(x), y_1)_q + (T(x), y_2)_q \\ &= c(x, T^*(y_1))_q + (x, T^*(y_2))_q \\ &= (x, cT^*(y_1))_q + (x, T^*(y_2))_q, \end{aligned}$$

como x es arbitrario, se tiene que $T^*(cy_1 + y_2) = cT^*(y_1) + T^*(y_2)$.

Sólo nos falta probar que T^* es única. Supongamos $U: G \rightarrow G$ lineal y que satisface que $(T(x), y)_q = (x, U(y))_q$ para toda $x \in G$, entonces

$$(T(x), y)_q = (x, T^*(y))_q = (x, U(y))_q \text{ para toda } x, y \in G,$$

lo que implica directamente que $T^* = U$. \square

Definición 1.4.1 El operador T^* descrito en el teorema anterior se conoce como el *Adjunto del operador T* .

Nótese que también

$$(x, T(y))_q = (T(y), x)_q = (y, T^*(x))_q = (T^*(x), y)_q,$$

lo que informalmente podría resumirse como la adición de un “*” cuando se cambia la posición del operador dentro del producto escalar.

El siguiente teorema nos explica la manera cómo se relacionan las matrices asociadas a estos operadores.

Teorema 1.4.2 Sea (G, q) un espacio vectorial de dimensión finita n , con producto interior q , y sea β una base ortonormal para G . Si T es un operador lineal en G cuya matriz asociada a la base β se denota por $A = [T]_\beta$ y $B = [T^*]_\beta$ la correspondiente para el operador T^* , entonces:

$$[T^*]_\beta = [T]_\beta^*$$

donde $[T]_\beta^*$ representa la matriz adjunta de $[T]_\beta$.

Demostración. Sea $\beta = \{q_1, \dots, q_n\}$ entonces la entrada ij de la matriz B está dada por

$$\begin{aligned} B_{ij} &= (T^*(q_j), q_i) = (q_i, T^*(q_j)) = (T(q_i), q_j) \\ &= A_{ji} = (A^*)_{ij}, \end{aligned}$$

que es precisamente la definición de la adjunta de una matriz para el caso en que todas sus entradas sean reales, de donde se concluye que $B = A^*$. \square

Observemos que

- La matriz asociada A^* al operador adjunto T^* de T , no es más que la transpuesta de aquella asociada a T , es decir $A^* = A'$.
- Estos resultados pueden extenderse al caso en el que el campo escalar del espacio vectorial G sea el de los números complejos, sin olvidar que la simetría en el producto escalar tiene una ligera variante: $(x, y)_q = \overline{(y, x)_q}$ donde la barra significa conjugación compleja; esto da como resultado que la matriz adjunta A^* , sea la transpuesta conjugada de A , es decir $(A^*)_{ij} = \overline{A_{ji}}$.

Dentro de la clase de operadores adjuntos existe un grupo de particular interés, constituido por aquellos operadores cuyo adjunto coincide con él mismo.

Definición 1.4.2 Sea (G, q) un espacio vectorial con producto interior q y sea T un operador lineal en G . T se denomina operador autoadjunto (o hermitiano) si $T = T^*$. Asimismo, una matriz cuadrada A se conocerá como autoadjunta (o hermitiana) si $A = A^*$.

Así, para el caso de matrices con entradas reales, ser autoadjunta equivale a ser simétrica. Una propiedad importante que gozan estos operadores, es que todos sus valores propios son reales.

Teorema 1.4.3 *Sea T un operador lineal autoadjunto en un espacio vectorial (G, q) . Si λ es un valor propio de T , entonces $\lambda \in \mathbf{R}$.*

Demostración. La demostración se hará en tres etapas:

Primero se probará que $\|T(\underline{x})\|_q = \|T^*(\underline{x})\|_q \quad \forall \underline{x} \in G$.

$$\begin{aligned} \|T(\underline{x})\|_q^2 &= (T(\underline{x}), T(\underline{x}))_q = (T^*T(\underline{x}), \underline{x})_q = (TT^*(\underline{x}), \underline{x})_q = (T^*(\underline{x}), T^*(\underline{x}))_q \\ &= \|T^*(\underline{x})\|_q^2. \end{aligned}$$

La segunda parte relaciona los valores propios de los dos operadores; si λ es valor propio de T entonces $\bar{\lambda}$ es valor propio de T^* .

Si λ es el valor propio de T asociado al vector propio \underline{x} , es decir $T(\underline{x}) = \lambda\underline{x}$, y definimos el operador U como el polinomio característico asociado a T , $U = T - \lambda I$; se tiene que si $U(\underline{x}) = \underline{0}$, entonces

$$\begin{aligned} \underline{0} = \|U(\underline{x})\|_q &= \|U^*(\underline{x})\|_q = \|(T - \lambda I)^*(\underline{x})\|_q = \|(T^* - \bar{\lambda}I)(\underline{x})\|_q \\ &= \|T^*(\underline{x}) - \bar{\lambda}I(\underline{x})\|_q \Rightarrow T^*(\underline{x}) = \bar{\lambda}\underline{x}, \end{aligned}$$

lo que nos dice que $\bar{\lambda}$ es un valor propio de T^* . Con estos dos resultados probados, la tesis del teorema se demuestra inmediatamente,

$$\lambda\underline{x} = T(\underline{x}) = T^*(\underline{x}) = \bar{\lambda}\underline{x} \text{ como } \underline{x} \neq \underline{0}, \text{ entonces } \lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbf{R}. \quad \square$$

Una propiedad que caracteriza a este tipo de operadores es aquella que afirma que un operador es autoadjunto si y sólo sí, el espacio vectorial sobre el cual se define, tiene una base ortonormal formada por los valores propios de dicho operador. Su demostración no es difícil, pero se prefirió omitir con el objeto de no hacer de este trabajo un escrito excesivamente técnico. Para mayores detalles se sugiere consultar [7].

Finalmente estamos en la posibilidad de presentar los conceptos y resultados que nos permitirán relacionar todo lo presentado en esta sección, con el material expuesto previamente.

Definición 1.4.3 Sea a un operador autoadjunto sobre un espacio vectorial (G, q) con matriz asociada A . El cociente de Rayleigh para un vector $\underline{x} \neq \underline{0}$, elemento del espacio vectorial G , se define por el escalar

$$R_a(\underline{x}) = \frac{(\underline{a}(\underline{x}), \underline{x})_q}{\|\underline{x}\|_q^2},$$

equivalentemente

$$R_A(\underline{x}) = \frac{(A\underline{x}, \underline{x})_q}{\|\underline{x}\|_q^2}.$$

Teorema 1.4.4 Para una matriz autoadjunta A con valores propios ordenados decrecientemente $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, se tiene que

$$\max_{\underline{x} \neq \underline{0}} R_A(\underline{x}) = \lambda_1,$$

$$\min_{\underline{x} \neq \underline{0}} R_A(\underline{x}) = \lambda_n.$$

Demostración. Como se mencionó anteriormente, podemos encontrar una base ortonormal $\{\underline{x}_1, \dots, \underline{x}_n\}$ de vectores propios de A , tales que $A\underline{x}_i = \lambda_i \underline{x}_i$, $1 \leq i \leq n$, donde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ son los valores propios asociados al correspondiente \underline{x}_i . Por el teorema 1.4.3, sabemos que todos estos $\lambda_i \in \mathbf{R}$; cada vector \underline{x} en el espacio vectorial (G, q) , puede expresarse como combinación lineal de esta base de la siguiente manera

$$\underline{x} = \sum_{i=1}^n (\underline{x}, \underline{x}_i)_q \underline{x}_i,$$

por lo tanto

$$\begin{aligned} R_A(\underline{x}) &= \frac{(A\underline{x}, \underline{x})_q}{\|\underline{x}\|_q^2} = \frac{(A \sum_{i=1}^n (\underline{x}, \underline{x}_i)_q \underline{x}_i, \sum_{j=1}^n (\underline{x}, \underline{x}_j)_q \underline{x}_j)_q}{\|\underline{x}\|_q^2} \\ &= \frac{(\sum_{i=1}^n (\underline{x}, \underline{x}_i)_q A\underline{x}_i, \sum_{j=1}^n (\underline{x}, \underline{x}_j)_q \underline{x}_j)_q}{\|\underline{x}\|_q^2} = \frac{(\sum_{i=1}^n (\underline{x}, \underline{x}_i)_q \lambda_i \underline{x}_i, \sum_{j=1}^n (\underline{x}, \underline{x}_j)_q \underline{x}_j)_q}{\|\underline{x}\|_q^2} \\ &= \frac{\sum \sum \lambda_i (\underline{x}, \underline{x}_i)_q (\underline{x}, \underline{x}_j)_q (\underline{x}_i, \underline{x}_j)_q}{\|\underline{x}\|_q^2} = \frac{\sum \sum \lambda_i (\underline{x}, \underline{x}_i)_q (\underline{x}, \underline{x}_j)_q \delta_{ij}}{\|\underline{x}\|_q^2} = \frac{\sum_{i=1}^n \lambda_i ((\underline{x}, \underline{x}_i)_q)^2}{\|\underline{x}\|_q^2} \\ &\leq \frac{\lambda_1 \sum_{i=1}^n ((\underline{x}, \underline{x}_i)_q)^2}{\|\underline{x}\|_q^2} = \lambda_1. \end{aligned}$$

La otra parte del teorema se demuestra de una manera semejante. \square

La igualdad se alcanza cuando el cociente de Rayleigh se evalúa precisamente en el vector propio asociado al valor característico máximo (mínimo) \underline{x}_1 (\underline{x}_n)

$$R_A(\underline{x}_1) = \frac{(A\underline{x}_1, \underline{x}_1)_q}{\|\underline{x}_1\|_q^2} = \frac{(\lambda_1 \underline{x}_1, \underline{x}_1)_q}{\|\underline{x}_1\|_q^2} = \frac{\lambda_1 (\underline{x}_1, \underline{x}_1)_q}{\|\underline{x}_1\|_q^2} = \lambda_1.$$

1.5 Construcción de la nube de inercia máxima.

Esta sección busca ligar las dos previas y explicar de dónde surge la construcción propuesta.

Al retomar lo que se discutió sobre la inercia de una nube de puntos (ver definición 1.3.3), recordemos que para el caso de nubes centradas se pudo reducir a

$$\psi = \sum_i p_i \|\underline{x}_i\|_m^2,$$

donde m representa la métrica usada en el espacio de individuos (F, m) . Como sabemos, cada uno de los vectores \underline{x}_i que conforman esta nube, pueden expresarse de manera única como combinación lineal de alguna base ortonormada de (F, m) , digamos $\{\underline{f}_1, \dots, \underline{f}_p\}$, concretamente $\underline{x}_i = \sum_{j=1}^p (\underline{x}_i, \underline{f}_j)_m \underline{f}_j$, de tal suerte que podemos encontrar una expresión equivalente para ψ a través de esta base

$$\begin{aligned} \sum_{i=1}^n p_i \|\underline{x}_i\|_m^2 &= \sum_{i=1}^n p_i (\underline{x}_i, \underline{x}_i)_m = \sum_{i=1}^n p_i \left(\sum_{j=1}^p (\underline{x}_i, \underline{f}_j)_m \underline{f}_j, \sum_{k=1}^p (\underline{x}_i, \underline{f}_k)_m \underline{f}_k \right)_m \\ &= \sum_{i=1}^n p_i \sum_{j=1}^p (\underline{x}_i, \underline{f}_j)_m \sum_{k=1}^p (\underline{x}_i, \underline{f}_k)_m (\underline{f}_j, \underline{f}_k)_m \\ &= \sum_{i=1}^n p_i \sum_{j=1}^p (\underline{x}_i, \underline{f}_j)_m \sum_{k=1}^p (\underline{x}_i, \underline{f}_k)_m \delta_{jk} \\ &= \sum_{i=1}^n p_i \sum_{j=1}^p (\underline{x}_i, \underline{f}_j)_m^2 = \sum_{i=1}^n \sum_{j=1}^p p_i (\underline{x}_i, \underline{f}_j)_m^2. \end{aligned} \tag{1.2}$$

Si retomamos el *Esquema de Dualidad* discutido en la sección 1.2 y la forma como individuos y variables se consideraron elementos de espacios vectoriales con sus respectivas métricas y relaciones duales; podemos encontrar, usando las técnicas introducidas previamente, que las matrices que resumen esta información (ver la expresión 1.1) resultan ser autoadjuntas, lo que nos permite introducir el Cociente de Rayleigh (ver definición 1.4.3) como vínculo para encontrar la nube de inercia máxima. Veamos con más detalle.

En el teorema 1.3.2 y el corolario siguiente, probamos que $\psi = \text{traza}(VM) = \text{traza}(WD)$. Si analizamos con cuidado estas matrices y utilizamos el hecho de que M y D son matrices diagonales, vemos que pueden expresarse como el producto de dos matrices simétricas con entradas reales, es decir, autoadjuntas

$$\begin{aligned} VM &= X'(DXM) = AA^* \\ WD &= X(MX'D) = A'(A')^*. \end{aligned}$$

La razón por la que se han nombrado precisamente AA^* y $A'(A')^*$, se explica en el siguiente teorema, que lo enunciaremos para el caso de la matriz $VM = AA^*$, pero puede referirse análogamente para $WD = A'(A')^*$.

Teorema 1.5.1 *La inercia de una nube de puntos centrada, es la suma de los valores del cociente de Rayleigh asociados a la matriz AA^* , sobre la base ortonormada de vectores $\{f_1^*, \dots, f_p^*\}$.*

Demostración. Como se concluyó con anterioridad, la inercia se puede escribir alternativamente en términos de la base ortonormal $\{f_1, \dots, f_p\}$ (ver 1.2), a través de la expresión

$$\psi = \sum_{i=1}^n p_i \|\mathbf{x}_i\|_m^2 = \sum_{i=1}^n \sum_{j=1}^p p_i (\mathbf{x}_i, \mathbf{f}_j)_m^2.$$

Asimismo se había definido el producto escalar d_p y la aplicación a (ver pág. 11) sobre los elementos de la base $\{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$ que nos permitía pasar de un espacio vectorial a otro de la siguiente manera

$$\begin{aligned} a : (E^*, w) &\rightarrow (F, m) \\ a(\mathbf{e}_i^*) &= \mathbf{x}_i, \end{aligned}$$

cuya matriz asociada a la base canónica de F es X' . Análogamente

$$\begin{aligned} a' : (F^*, v) &\rightarrow (E, d) \\ a'(\mathbf{f}_j^*) &= \mathbf{x}^j, \end{aligned}$$

con matriz asociada X . Así, el operador adjunto a^* se define por

$$a^* : (F, m) \rightarrow (E^*, w),$$

de tal manera que

$$(a(\mathbf{e}_i^*), \mathbf{f}_j)_m = (\mathbf{e}_i^*, a^*(\mathbf{f}_j))_w,$$

de donde es fácil probar que $a^* = w^{-1}a'm = da'm$ por lo que la matriz asociada al operador aa^* es $X'DXM = VM$ y que habíamos llamado precisamente AA^* , (así como $WD = X'(DXM)$ que habíamos llamado $A'(A')^*$), entonces

$$(\mathbf{x}_i, \mathbf{f}_j)_m = (a(\mathbf{e}_i^*), \mathbf{f}_j)_m = (\mathbf{e}_i^*, a^*(\mathbf{f}_j))_w,$$

$$\begin{aligned}
 R_{AA^*}(\underline{f}_j) &= (AA^*\underline{f}_j, \underline{f}_j)_m = (A^*\underline{f}_j, A^*\underline{f}_j)_w \\
 &= \left(\sum_i (A^*\underline{f}_j, \underline{e}_i^*)_{w_i} \underline{e}_i^*, \sum_i (A^*\underline{f}_j, \underline{e}_i^*)_{w_i} \underline{e}_i^* \right)_w \\
 &= \sum_i (A^*\underline{f}_j, \underline{e}_i^*)_{w_i}^2 (\underline{e}_i^*, \underline{e}_i^*)_w = \sum_i (A^*\underline{f}_j, \underline{e}_i^*)_{w_i}^2 p_i \\
 &= \sum_i (\underline{x}_i, \underline{f}_j)_m^2 p_i.
 \end{aligned}$$

por lo que se tiene

$$\psi = \sum_{j=1}^p R_{AA^*}(\underline{f}_j). \quad \square$$

Al referirnos una vez más al teorema 1.3.2, se concluye que

$$\psi = \sum_{j=1}^p R_{AA^*}(\underline{f}_j) = \text{traza}(VM).$$

Ahora bien, el resultado anterior nos dice que la inercia total puede recuperarse a través del cambio a una base ortonormal, sin embargo, no debemos perder de vista nuestro objetivo principal que consiste en describir lo mejor posible, el comportamiento de individuos y variables en un espacio de menor dimensión. La pregunta que surge ahora es: ¿cómo debemos elegir dicha base, de tal suerte que al proyectar sobre un espacio afín de dimensión menor a la original se recupere la mayor inercia posible o, equivalentemente, perder la mínima cantidad de información disponible?

Si pensamos en los teoremas 1.4.4 y 1.5.1, así como en los comentarios que hicimos sobre ellos, sabemos que una posible base ortonormal que podemos elegir y que reconstruiría la inercia total a través de los cocientes de Rayleigh, es aquella formada por los vectores propios de la matriz VM . Esta idea resulta muy interesante, por varias razones.

- Sabemos que la inercia es la suma de estos cocientes (teorema 1.5.1) sobre cualquier base ortonormal, en particular por aquella formada por vectores propios, pero además sabemos que el máximo de estos sumandos (teorema 1.4.4) se alcanza cuando el cociente se evalúa en el vector propio asociado al máximo valor característico λ_1 , de tal manera, que ésta será la dirección que capte la mayor proporción de inercia.
- Si eliminamos la dirección del vector propio asociado a λ_1 , nos queda un subespacio afín ortogonal a esta dirección (recordemos que forman un conjunto de

vectores ortogonales), con la inercia restante igual al total original menos la que fue “explicada” por la dirección del vector excluido. Si aplicamos nuevamente el teorema 1.4.4, podemos concluir que el siguiente vector que “explica” la mayor cantidad de la inercia sobrante, es aquél que corresponde al valor propio máximo λ_2 y así sucesivamente.

- Cabe recordar también, que los cocientes de Rayleigh a los que nos hemos referido, se aplican sobre las matrices VM o WD , que como ya se dijo, son autoadjuntas y podemos garantizar que sus valores propios son no negativos. El número de estos valores que sean estrictamente positivos, coincide con el rango de la matriz correspondiente y cuando este número es inferior a la dimensión original, se refleja una situación en la que en realidad, todos los puntos se encontraban en un subespacio propio, contenido estrictamente en el original.

En una aplicación de la vida real, lo que comúnmente sucede es que necesitamos una representación en un espacio de dos, o a lo más tres dimensiones, de tal manera que busquemos la proyección ortogonal en un subespacio de, digamos, dimensión h . Por lo que se discutió anteriormente, se propone cambiar la base de nuestros datos observados, a aquella definida por los vectores propios, y una vez hecha esta transformación, proyectar ortogonalmente sobre las h primeras direcciones, es decir, sobre los vectores asociados a los h valores propios más grandes. Como no se incluyeron todos los vectores de la base, es claro que no se recuperará el total de la inercia, de donde surge la siguiente definición.

Definición 1.5.1 *Al proyectar sobre un espacio de dimensión h , definido por los primeros h vectores propios asociados a los máximos valores característicos $\lambda_1 \geq \dots \geq \lambda_h$, la cantidad*

$$\frac{\sum_{i=1}^h \lambda_i}{\text{traza}(VM)},$$

se conoce como “la proporción de inercia explicada” en el subespacio de dimensión h generado y es una medida de la calidad global de la representación.

Definición 1.5.2 El vector propio normalizado \underline{u}_i del espacio de individuos $F = \mathbf{R}^p$, asociado al valor propio λ_i de la matriz VM ($i = 1, \dots, p$), genera lo que se conoce como el i -ésimo Eje Principal.

Siguiendo con la notación empleada hasta ahora $VM\underline{u}_i = \lambda_i\underline{u}_i$, al premultiplicar por la matriz M tenemos que $MVM\underline{u}_i = \lambda_i M\underline{u}_i$, y si llamamos $\underline{v}_i = M\underline{u}_i$ podemos concluir que $MV\underline{v}_i = \lambda_i\underline{v}_i$, lo que nos dice que si \underline{u}_i es vector propio de VM asociado a λ_i , entonces $\underline{v}_i = M\underline{u}_i$ es el vector propio asociado también a λ_i , pero de la matriz MV .

Definición 1.5.3 El vector $\underline{e}^i = X\underline{v}_i = XM\underline{u}_i \in E = \mathbf{R}^n$ se conoce como la i -ésima Componente Principal ($i = 1, \dots, p$).

Si recordamos quiénes eran las matrices V y W y pensamos concretamente en las métricas $D_p = D$ y $D_{1j_s} = M$, (ver la expresión 1.1) podemos notar que XM no es más que la matriz de datos reducida, además

$$\begin{aligned}\lambda_i\underline{u}_i &= MV\underline{v}_i = MX'D_pX\underline{v}_i = MX'D_p\underline{e}^i, \\ &\Rightarrow XM X'D_p\underline{e}^i = \lambda_i X\underline{v}_i = \lambda_i\underline{e}^i, \\ &\Rightarrow WD_p\underline{e}^i = \lambda_i\underline{e}^i,\end{aligned}$$

lo que nos dice que \underline{e}^i es vector propio de WD_p .

1.5.1 Descripción de individuos.

Hasta ahora, hemos expuesto la técnica de Componentes Principales utilizando un enfoque diferente que nos ha permitido librarnos del concepto de probabilidad que admite diversas interpretaciones, así como de supuestos distribucionales que difícilmente se cumplen en problemas prácticos. No obstante estas ventajas teóricas, si nos restringimos al uso de las métricas más comunes, los resultados presentados hasta este momento pueden obtenerse mediante un paquete estadístico convencional, sin embargo este enfoque permite además de generalizar a otras métricas, hacer extensiones que nos dan la posibilidad de evaluar la calidad de representación de cada individuo,

lo cual, aunque no se encuentre en la paquetería estándar se recomienda incluir en el análisis.

La interpretación geométrica de todos estos conceptos es sumamente interesante. Al inicio, teníamos una nube de n puntos en \mathbf{R}^p , donde cada dimensión estaba asociada a una de las características medidas; cuando matemáticamente se propuso cambiar de base a aquella formada por los vectores propios ortonormados de la matriz VM $\{\underline{v}_i\}$, lo que en realidad hicimos fue un cambio de coordenadas, una rotación sobre el espacio \mathbf{R}^p donde las direcciones adecuadas fueron aquellas asociadas a estos vectores propios, ordenados decrecientemente de acuerdo a la magnitud de su valor característico correspondiente.

Estos vectores, al premultiplicarse por la matriz M , dan origen a los $\{\underline{v}_i\}$ (conservando el orden heredado de los $\{\underline{u}_i\}$) y son ellos los que nos permitirán conocer las nuevas coordenadas de cada punto después del cambio de ejes. La i -ésima coordenada de los n individuos se encuentra en el vector llamado *componente principal*, que no es más que una combinación lineal de las características que se midieron inicialmente, con coeficientes dados precisamente por los $\{\underline{v}_i\}$, es decir, $\underline{c}^i = X \underline{v}_i$ ($\underline{c}^i \in \mathbf{R}^n$).

La construcción de la nueva nube de puntos, resulta entonces simplemente de postmultiplicar la matriz de datos observados X , por aquella cuyas columnas son los vectores $\{\underline{v}_i\}$; la proyección sobre un subespacio de dimensión menor h , consistirá en incluir únicamente los h primeros de ellos, con la garantía de haber captado la mayor inercia posible, si se comparara contra cualquier otra combinación lineal que se quisiera intentar. Además mencionamos que la proporción de inercia explicada puede considerarse una medida de la calidad de representación de la nube en su conjunto, sin embargo, carecemos de información relevante de la calidad marginal de un individuo particular.

Intuitivamente, un individuo \underline{x}_i se considera bien representado, si el ángulo que forma con su proyección es pequeño, o equivalentemente, si su norma y la correspondiente a su representación son similares. Llamando α al ángulo formado entre el individuo \underline{x}_i y su proyección $\underline{\pi}_i$, se tiene que

$$\cos \alpha = \frac{\|\underline{\pi}_i\|_m}{\|\underline{x}_i\|_m}$$

De tal manera que un punto se considerará bien representado si $\cos \alpha \rightarrow 1$ y conforme $\cos \alpha \rightarrow 0$ tendremos evidencia de menor calidad. Esta es la razón por la cual $\cos \alpha$ es llamado el **coeficiente de calidad de representación del i -ésimo individuo**.

Cuando se tiene una buena calidad de representación, los ejes en el nuevo sistema de referencia, adquirirán una interpretación natural acorde al problema real, de acuerdo con la magnitud de los coeficientes de las combinaciones lineales que los determinan, de tal suerte que cada eje podrá pensarse como un resumen de algunas de las características originales, y la posición de los puntos bajo este nuevo sistema coordenado, se interpretará con respecto a estos resúmenes, lo que puede ser de gran utilidad al ubicarse en el contexto del problema.

1.5.2 Descripción de variables.

Lo que presentamos en esta sección, no se incluye en la mayoría de los paquetes estándar, no obstante consideramos que representa una aportación valiosa de este enfoque y debe mencionarse con detalle.

Como se comentó anteriormente, sabemos que tendremos p valores propios no negativos de la matriz autoadjunta VM , de los cuales ρ de ellos serán estrictamente positivos ($\rho \leq p$), lo que indica que un individuo pudo haber aportado la misma información pero a través de ρ variables en lugar de p , es decir, "medimos de más". Este número ρ será referido en lo sucesivo como el rango de VM puesto que coincide con él.

De los p componentes principales, sólo ρ de ellos están asociados a valores propios estrictamente positivos que serán los que realmente nos brinden información valiosa, a este subconjunto lo llamaremos $\{\xi^1, \dots, \xi^\rho\}$ y de acuerdo a la teoría resumida en el *Esquema de Dualidad* de la sección 1.2, deberá medirse con la métrica d_ρ . Lo que puede verse fácilmente es que forman una base ortogonal (no normado) que genera un subespacio ρ -dimensional en \mathbf{R}^n .

Proposición 1.5.1 *El conjunto de ρ componentes principales $\{\xi^1, \dots, \xi^\rho\}$, (ver definición 1.5.3), forma una base d_ρ -ortonormada.*

Demostración. Recordemos algunos detalles que necesitaremos,

$\{\underline{u}_i\}$ forman un conjunto d_p ortonormado de vectores propios de VM , $VM\underline{u}_i = \lambda_i \underline{u}_i$.

También habíamos deducido que el vector $\underline{v}_i = M\underline{u}_i$ es un vector característico de la matriz MV , $MV\underline{v}_i = \lambda_i \underline{v}_i$.

Asimismo dijimos que $V = X'D_p X$ y que M es una matriz diagonal.

$$\begin{aligned} (\underline{e}^i, \underline{e}^j)_{d_p} &= (X\underline{v}_i, X\underline{v}_j)_{d_p} = (XM\underline{u}_i, XM\underline{u}_j)_{d_p} = \underline{u}'_i M X' D_p X M \underline{u}_j \\ &= \underline{u}'_i M (X' D_p X) M \underline{u}_j = \underline{u}'_i M V M \underline{u}_j = \lambda_j \underline{u}'_i M \underline{u}_j \\ &= \lambda_j \delta_{ij}. \quad \square \end{aligned}$$

Esta proposición nos permite afirmar que bajo el planteamiento original, aquellos vectores que representaban la medición de una característica bajo estudio sobre toda la muestra ($\underline{x}^i \in \mathbf{R}^n$ $i = 1, \dots, p$), pueden expresarse en términos de las componentes principales y la métrica d_p . Lo que no debemos olvidar es que no conforman un conjunto normado pues

$$\|\underline{e}^j\|_{d_p} = \sqrt{\lambda_j}$$

de tal manera que $\left\{ \frac{1}{\sqrt{\lambda_j}} \underline{e}^j \right\}_{j=1, \dots, p}$ constituye una base ortonormada del subespacio mencionado, sobre la cual, la variable \underline{x}^i tiene por j -ésima coordenada a

$$\frac{1}{\sqrt{\lambda_j}} (\underline{x}^i, \underline{e}^j)_{d_p},$$

de tal suerte que

$$\|\underline{x}^i\|_{d_p}^2 = \sum_{j=1}^p \frac{1}{\lambda_j} \left\{ (\underline{x}^i, \underline{e}^j)_{d_p} \right\}^2.$$

Es importante recordar que si las variables son centradas, cada producto-escalar reconstruye la covarianza entre la variable y su correspondiente componente principal (ver definición 1.2.6), por lo que cada coordenada será igual a esta covarianza dividida entre $\sqrt{\lambda_j}$

$$\frac{1}{\sqrt{\lambda_j}} (\underline{x}^i, \underline{e}^j)_{d_p} = \frac{\text{cov}(\underline{x}^i, \underline{e}^j)}{\sqrt{\lambda_j}},$$

y para el caso de variables centradas y reducidas, tendremos que lo que se reconstruye es la correlación

$$\frac{1}{\sqrt{\lambda_j}} (\underline{x}^i, \underline{e}^j)_{d_p} = \text{corr} (\underline{x}^i, \underline{e}^j).$$

Así, la proyección de un vector-variable \underline{x}^i sobre un subespacio de dimensión h (comúnmente 2), estará dada por

$$\underline{\pi}^i = \left(\frac{1}{\sqrt{\lambda_1}} (\underline{x}^i, \underline{e}^1)_{d_p}, \dots, \frac{1}{\sqrt{\lambda_h}} (\underline{x}^i, \underline{e}^h)_{d_p} \right),$$

y si \underline{x}^i es centrada y reducida, tendremos que las coordenadas de esta proyección pueden expresarse por

$$\underline{\pi}^i = \left(\text{corr} (\underline{x}^i, \underline{e}^1), \dots, \text{corr} (\underline{x}^i, \underline{e}^h) \right),$$

por lo que los puntos asociados a las variables estarán contenidos en una hipersfera centrada de dimensión h y radio 1. El caso más frecuente, consiste en que esta representación se construye sobre un plano y está formada por una nube de puntos, cada uno asociado a una variable, todos ellos contenidos en un círculo de radio 1 y con centro en el origen. La primera coordenada nos da el valor de la correlación entre la variable y el primer componente principal, de tal suerte que abscisas cercanas a 1 ó -1, serán evidencia de que la variable representada por el punto correspondiente, queda muy bien explicada en términos de la primera componente, teniendo la interpretación análoga para las ordenadas. Asimismo, valores de coordenadas cercanas a 0 sugieren una mala correspondencia entre la variable y la componente respectiva. Esta gráfica se conoce como el *Círculo de Correlaciones*.

1.6 Objetos suplementarios.

1.6.1 Individuos suplementarios.

En muchas situaciones prácticas, necesitamos comparar el comportamiento de los individuos que estamos analizando con algún otro grupo que se tome como referencia sin que éste se incluya como parte del análisis, sino únicamente en la representación

final, lo que permitiría ubicarlos gráficamente y compararlos con el conjunto de interés. Aplicaciones de este estilo pueden encontrarse cuando se estudia una población y se necesita relacionarla con otra con fines comparativos, que por no formar parte del grupo bajo estudio, no deberá incluirse en la búsqueda del nuevo sistema de coordenadas.

Definición 1.6.1 *El conjunto de individuos cuyos pesos asociados son idénticamente cero, se conocen como Individuos Suplementarios, asimismo los elementos del grupo cuyos pesos son diferentes de cero se identifican como Individuos Activos.*

Supongamos que se tienen de k individuos suplementarios, lo que da un total de $n+k$ sujetos disponibles. Esta información se guardará en k vectores que se añadirán a la matriz X , después de los individuos bajo estudio, lo que da lugar a la matriz aumentada X_{n+k}

$$X_{n+k} = \begin{pmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_n \\ \underline{x}_{n+1} \\ \vdots \\ \underline{x}_{n+k} \end{pmatrix} \quad \text{con } \underline{x}_i \in \mathbf{R}^p.$$

El espacio vectorial E , formado por los vectores asociados a las variables y su dual E^* , serán entonces de dimensión $n+k$; esto implica que la matriz diagonal asociada a su métrica D'_p , será de dimensión $(n+k) \times (n+k)$ cuyos últimos k renglones tendrán 0 en todas sus entradas. Esto conduce al hecho de que el rango de la matriz VM no será alterado, sus valores y vectores propios serán los mismos que aquéllos que se hubieran obtenido utilizando únicamente los objetos activos. Los conjuntos $\{\underline{u}_i\}$ y $\{\underline{v}_i\}$ asociados a los valores característicos diferentes de cero, también permanecerán sin cambios.

La construcción de la nube de puntos que represente a los individuos en un espacio de menor dimensión, se hará siguiendo la misma idea de antes, calculamos $\underline{z}_i = X_{n+k} \underline{v}_i$, que nos dará como resultado un vector de $(n+k) \times 1$ con la i -ésima coordenada en el nuevo sistema de referencia para todos los individuos disponibles,

tanto activos como suplementarios. Esto lo hacemos para $i = 1$ y 2 si lo que queremos finalmente, es una representación en el plano euclidiano.

Por último vale la pena recalcar un par de detalles. Así como se sugirió utilizar una medida de calidad de representación para los individuos activos, se propone ahora para los individuos suplementarios. La definición es exactamente la misma. Asimismo se recomienda ampliamente, identificar los puntos asociados a los individuos suplementarios con algún símbolo que los distinga de los activos, lo que aumentará la interpretabilidad de la gráfica al hacer la comparación.

1.6.2 Variables suplementarias.

Análogo al concepto de individuos suplementarios, se puede hablar de variables suplementarias. Esto aparece en situaciones en las que se quieran comparar las variables observadas con otro conjunto que se haya medido sobre los mismos individuos pero que no se quiera introducir en el análisis excepto con fines comparativos. La metodología correspondiente es similar a la relativa a individuos suplementarios salvo por pequeños ajustes necesarios.

Partiendo de las p variables de interés centradas y reducidas previamente, podemos pensar en una ponderación de las mismas por pesos p^1, \dots, p^p que hayan sido asignados a cada una, atendiendo a su importancia relativa en el estudio. Estos pesos, al igual que aquéllos para los individuos, se introducirán a través del producto escalar del espacio de individuos F

$$(\underline{x}_i, \underline{x}_j)_{m_p} = \begin{cases} p^i & \text{si } i = j \\ 0 & \text{si } i \neq j, \end{cases}$$

con matriz asociada $M_p = \text{diag}\{p^1, \dots, p^p\}$, de donde el correspondiente producto escalar entre dos individuos se expresa como

$$(\underline{x}_i, \underline{x}_j)_{m_p} = \underline{x}_i' M_p \underline{x}_j.$$

La introducción de l variables suplementarias se lleva a cabo agregándolas como vectores columna en la matriz observada X , que para distinguirla de la original se

denotará por: X^{p+l} , de tal manera que el espacio vectorial de individuos será ahora de dimensión $p+l$ y la matriz X^{p+l} de $n \times (p+l)$ con la forma siguiente

$$X^{p+l} = (\underline{x}^1, \dots, \underline{x}^p : \underline{x}^{p+1}, \dots, \underline{x}^{p+l}).$$

Para que estas variables no se incluyan en el análisis pero sí en la representación final, se debe modificar adecuadamente el producto escalar m_p mencionado recientemente

$$\forall i, \forall j \in \{1, \dots, p+l\} : (\underline{f}_i, \underline{f}_j)_{m_p} = \begin{cases} p^i & \text{si } i = j \text{ y además } i \leq p, j \leq p \\ 0 & \text{si } i \neq j \text{ o bien } i > p \text{ ó } j > p, \end{cases}$$

de tal forma que la matriz asociada $M_p' = \text{diag}\{p^1, \dots, p^p, 0, \dots, 0\}$ es "esencialmente" igual a M_p excepto por los últimos l renglones idénticamente cero.

Así, el rango de la matriz VM_p' sigue siendo p (igual que la matriz VM sobre la cual se expuso la teoría previa) y sus vectores y valores propios permanecen invariantes, de tal suerte que lo mismo sucede para los conjuntos $\{\underline{u}_i\}$ y $\{\underline{v}_i\}$.

A partir de este momento, se sigue la misma idea presentada para el caso de tener únicamente variables activas, se construye el círculo de correlaciones con cada una de la $p+l$ variables utilizando tantas componentes principales como la dimensión del subespacio donde se proyecte. Así, si éste se escoge de dimensión h , las coordenadas de la proyección de la i -ésima variable \underline{x}^i estarán dadas por

$$\underline{\pi}^i = \left(\frac{1}{\sqrt{\lambda_1}} \text{corr}(\underline{x}^i, \underline{e}^1), \dots, \frac{1}{\sqrt{\lambda_h}} \text{corr}(\underline{x}^i, \underline{e}^h) \right) \text{ con } i = \{1, \dots, p+l\}.$$

De aquí que la interpretación sea exactamente la misma que para las variables activas, coordenadas cercanas a 1 ó -1 serán indicios de variables bien representadas por la componente correspondiente. Tanto para no confundir las variables activas con las suplementarias, así como para que el gráfico final cumpla con su misión comparativa, se sugiere asignar símbolos diferentes dependiendo del tipo de variable que se esté representando.

Capítulo 2

El Método Statis.

2.1 Introducción

El método Statis nos permite estudiar varias matrices de datos cuantitativos simultáneamente. El interés por un método de este estilo, surge en aplicaciones como las descritas en el capítulo anterior.

1. Cuando se registraron algunas variables en un grupo de individuos en diferentes momentos del tiempo y se trata de describir el comportamiento que siguieron a lo largo del período de observación;
2. cuando se registró un cierto grupo fijo de variables en diferentes conjuntos de individuos y se pretende describir el comportamiento de dichas variables al cambiar de sujetos, y por último
3. cuando un grupo de individuos se somete a la medición de algunas características agrupadas por algún criterio de seriación y se busca describir su evolución al pasar de un grupo a otro.

La primera situación es la que se presenta con mayor frecuencia en la vida real y es útil tanto para conocer la evolución de los individuos a través del tiempo, así como el comportamiento de las variables bajo estudio durante ese período. Asimismo, las dos últimas situaciones coinciden en que buscan describir el “desarrollo” de alguno

de los dos grupos de objetos que permanece fijo (variables o individuos), al cambiar el otro.

Al retomar las ideas expuestas en el capítulo anterior, en las que tanto los individuos como las variables se interpretaron como elementos de espacios vectoriales, se puede notar que la primera situación corresponde al análisis de la evolución conjunta de dos espacios y las dos últimas pretenden describir el comportamiento de los elementos de un espacio vectorial al cambiar el que define el otro grupo, sin olvidar que ambos se encuentran ligados a través del *Diagrama de Dualidad*. A cada uno de estos escenarios, corresponde una estrategia de análisis diferente, por lo que será fundamental hacer la elección adecuada al problema bajo estudio.

2.2 El concepto de "Estudio".

Recordemos las ideas y notación introducida previamente.

(F, m) , espacio vectorial de dimensión p , generado por las mediciones obtenidas de los individuos observados y cuyo producto escalar se denotó por m , con matriz asociada M que depende del producto escalar particular al que nos refiramos. Veamos algunos ejemplos, habíamos mencionamos cuatro posibles métricas de interés,

- para el producto euclidiano en el que $(x_i, x_j) = x_i' * x_j$, M será igual a I_p .
- la métrica que estandariza las variables, que tiene como matriz asociada $D_{1/s} = \text{diag} \left\{ \frac{1}{s_1}, \dots, \frac{1}{s_p} \right\}$, y a partir de la cual al aplicar la transformación $Y = XD_{1/s}$ obtenemos la llamada matriz estándar;
- la métrica resumida en la matriz $D_{1/s^2} = \text{diag} \left\{ \frac{1}{s_1^2}, \dots, \frac{1}{s_p^2} \right\}$, que al descomponerse como $D_{1/s} D_{1/s}$ reconstruye la matriz de correlaciones por medio de la transformación $C = D_{1/s} V D_{1/s}$;
- aquella que pondera las variables por pesos p^1, \dots, p^p cuya matriz asociada es $D^p = \text{diag}\{p^1, \dots, p^p\}$. Comúnmente se aplica sobre variables previamente centradas y reducidas.

(E, d) , espacio vectorial de dimensión n inducido por las variables medidas y cuyo producto escalar se denotó por d con matriz asociada D . Mencionamos dos métricas de particular interés,

- la métrica euclidiana cuya matriz asociada es $D = I_n$, su definición es análoga a la expuesta para el espacio de individuos; en el caso de disponer de variables centradas la distancia inducida reconstruye la covarianza y para características que además están centradas y reducidas, se recupera la correlación muestral;
- la métrica que pondera la importancia de cada uno de los n individuos, a través de la asignación de pesos p_i con $i = 1, \dots, n$, cuya matriz asociada es $D_p = \text{diag}\{p_1, \dots, p_n\}$ y que incluye, como caso particular, la igualdad de pesos cuando $p_i = 1/n$. Análogamente al caso anterior, esta métrica reconstruye la varianza muestral entre variables cuando éstas son centradas, pero ahora, ponderada por los pesos de los individuos. Para el caso de variables centradas y reducidas reproduce la correlación muestral ponderada.

Finalmente, hablamos de los espacios duales (F^*, v) y (E^*, w) que al incluirlos en el análisis, nos permitieron relacionar los espacios originales; sus productos escalares asociados se representaron a través de las matrices $V = X'DX$ y $W = XMX'$, donde D y M representan las matrices asociadas a los productos escalares d y m , cualesquiera que estos sean, de los espacios E y F respectivamente.

Cuando se piensa realizar un aplicación con esta técnica, es importante empezar por distinguir claramente el objetivo que se persigue, identificar los objetos de interés que motivan el análisis, ya sea la descripción de los individuos, de las variables o ambos, ya que esto definirá la estrategia a usar. Asimismo debemos escoger las métricas que satisfagan lo mejor posible nuestros requerimientos, asignar pesos a los individuos y decidir si se trabajará con una matriz centrada y reducida o en su forma original. Todas estas elecciones se resumen en el concepto de "estudio", que se presenta a continuación.

Definición 2.2.1 *Un Estudio es una terna (X, m, d) , en la que se establecen los productos escalares escogidos para realizar un análisis de la matriz X , m es un producto*

escalar sobre F que induce una distancia entre individuos y d es un producto escalar sobre E que induce una distancia entre variables.

La relación entre individuos y variables, se estableció a través de las aplicaciones lineales

$$\begin{aligned} a : (E^*, w) &\rightarrow (F, m) & a' : (F^*, v) &\rightarrow (E, d) \\ a(\underline{e}_i^*) &= \underline{x}_i & a'(\underline{f}_j^*) &= \underline{x}_j^j, \end{aligned} \quad (2.1)$$

con matrices asociadas N' y X respectivamente, lo que nos permitió presentar diferentes expresiones de la inercia de una nube de puntos centrada, ψ^1 ,

$$\begin{aligned} \psi &= \text{traza}(vm) = \text{traza}(wd) = \text{traza}(aa^*) = \text{traza}(a'(a')^*) \\ &= \lambda_1 + \lambda_2 + \dots + \lambda_p, \end{aligned}$$

donde p es el rango de las matrices VM y WD , cuyos valores propios no nulos ordenados en forma decreciente son $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$.

Cuando disponemos de más de un estudio y se busca describir la evolución de individuos y/o variables a través de ellos, nos encontramos ante una situación donde se puede aplicar el *Método Statis* propuesto originalmente por Y. Escoufier y trabajado posteriormente por Christine Lavit en su libro "*Analyse Conjointe de Tableaux Quantitatifs*" [5].

La idea central del método radica en la búsqueda de una estructura común de los elementos de interés al pasar de un estudio a otro; esta estructura que tiene por objetivo resumir las relaciones que se presenten en cada tabla, se conoce como *Intraestructura* y su construcción dependerá del elemento de interés, es decir, para el caso en que nuestro interés recaiga en la descripción de individuos, resumirá las relaciones entre sujetos que se presentaron en cada tabla de datos, asimismo resumirá las relaciones entre variables cuando éstas sean el motivo de nuestra investigación.

¹La expresión obtenida en el capítulo anterior, se presentó para las matrices asociadas a estos productos escalares, lo cual no constituye diferencia alguna.

2.3 Comparación entre objetos.

2.3.1 Estructura común de los individuos a través de los estudios.

Recordando algunos de los esquemas generales descritos en la introducción de este capítulo, distinguimos aquél en el que un grupo fijo de n individuos se somete a un análisis que consiste en la medición de p variables cuantitativas en k diferentes momentos y se quiere describir la evolución que sufre cada uno de ellos a lo largo del período de observación. Una situación más general sería aquella en la que se permitiera cambiar el conjunto de variables en cada intervalo de tiempo, ya sea en número o en elementos. Ambas situaciones coinciden en que el conjunto de interés está formado por el grupo de individuos y éste se conserva a lo largo del análisis, a diferencia del formado por las variables observadas, que podría ir cambiando.

Desde el punto de vista algebraico, disponemos de k estudios en los cuales, el espacio vectorial formado por las variables (\mathbf{R}^n) permanece fijo, a diferencia de aquél formado por los individuos. La información disponible se resume entonces, en k ternas, de la forma

$$(X_1, m_1, d), \dots, (X_k, m_k, d),$$

donde el espacio vectorial que permanece fijo es $(E, d) = \mathbf{R}^n$ y cada uno de los tiempos de observación induce un espacio cuya dimensión puede cambiar, atendiendo al número p_j de variables al tiempo j , $(E_j, m_j) = \mathbf{R}^{p_j}$ ($j = 1, \dots, k$)². A continuación se presenta el esquema de dualidad que describe la estructura de esta situación restringida al caso en que se comparen dos estudios (X_1, m_1, d) y (X_2, m_2, d) .

²Cabe recordar que en el capítulo 1 p_i se usó para referir el peso asociado al individuo i , no obstante aquí surge de manera natural como el número de variables en el i -ésimo estudio por lo que se recomienda tener en cuenta esta duplicidad. En caso de confusión, se hará la aclaración pertinente.

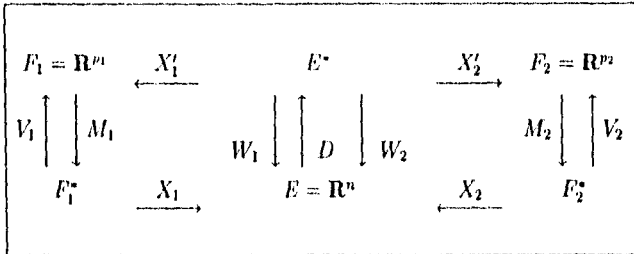


diagrama 3

donde

$$V_1 = X_1' D X_1 \quad W_1 = X_1 M_1 X_1'$$

$$V_2 = X_2' D X_2 \quad W_2 = X_2 M_2 X_2'$$

Si al cambiar las condiciones de observación, llámese variables o tiempos calendario, los individuos conservan esencialmente las distancias entre ellos, podremos afirmar que existe una estructura común que se mantiene durante este período; de tal manera que deberemos centrar nuestro esfuerzo en comparar las matrices de productos escalares entre individuos W_j , lo que equivale a comparar las formas bilineales $w_j \in \mathcal{L}(E^*, E)$ contra una "estructura media" que podamos definir. Puesto que necesitamos establecer un criterio de comparación entre elementos de un espacio vectorial (aquel formado por las matrices cuadradas de dimensión n si pensamos en W_j o bien el espacio de funciones $\mathcal{L}(E^*, E)$ si trabajamos con w_j), definiremos un producto escalar que nos permita hacerlo.

Teorema 2.3.1 Sean $a, b \in \mathcal{L}(E, F)$. La aplicación que a cada pareja (a, b) le asigna la traza de a^*b , denotada por

$$\text{traza}(a^*b) = (a, b)_{HS},$$

define un producto escalar en $\mathcal{L}(E, F)$ llamado comúnmente, de Hilbert-Schmidt.

Demostración. Sean $a, b, c \in \mathcal{L}(E, F)$ y $\lambda \in \mathbf{R}$,

- para probar simetría, utilizaremos lo que se vió en el teorema 1.5.1, donde se dedujo que $a^* = da'm$, por lo que

$$\begin{aligned}(a, b)_{HS} &= \text{traza}(a^*b) = \text{traza}(da'mb) = \text{traza}(da'mb)' = \text{traza}(b'm'ad') \\ &= \text{traza}(b'mad) = \text{traza}(db'ma) = \text{traza}(b^*a) = (b, a)_{HS}.\end{aligned}$$

•

$$(\lambda a, b)_{HS} = \text{traza}((\lambda a)^*b) = \text{traza}(\lambda a^*b) = \lambda \text{traza}(a^*b) = \lambda(a, b)_{HS}.$$

•

$$\begin{aligned}(a + b, c)_{HS} &= \text{traza}((a + b)^*c) = \text{traza}((a^* + b^*)c) = \text{traza}(a^*c + b^*c) \\ &= \text{traza}(a^*c) + \text{traza}(b^*c) = (a, c)_{HS} + (b, c)_{HS}.\end{aligned}$$

- Como se mencionó en la sección 1.5, el operador a^*a es autoadjunto con todos sus valores propios $\lambda_1, \dots, \lambda_p$ diferentes de cero por lo que

$$(a, a)_{HS} = \text{traza}(a^*a) = \sum_{i=1}^p \lambda_i \geq 0,$$

la igualdad se da si $\lambda_i = 0 \forall i$ lo cual sucede si y sólo si $a \equiv 0$. \square

Así, para comparar dos matrices W_1 y W_2 que contienen los productos escalares entre todos los individuos ponderados por sus respectivos pesos, se propone el coeficiente denominado *Covarianza Vectorial* (*covv*) que no es más que su producto escalar Hilbert-Schmidt

$$\text{covv}(W_1, W_2) = (W_1, W_2)_{HS} = \text{traza}(W_1^*W_2).$$

Si este coeficiente se calcula sobre las matrices (o aplicaciones) normalizadas, se conoce como *Correlación Vectorial* o coeficiente *RV*

$$RV(W_1, W_2) = \left(\frac{W_1}{\|W_1\|} \cdot \frac{W_2}{\|W_2\|} \right)_{HS},$$

a partir del cual la distancia entre dos matrices se simplifica a

$$d_{RV}(W_1, W_2) = \left\| \frac{W_1}{\|W_1\|} - \frac{W_2}{\|W_2\|} \right\|_{HS} = \sqrt{2(1 - RV(W_1, W_2))},$$

de donde podemos concluir que la condición $RV(W_1, W_2) = 1$ implica distancia nula entre las matrices normalizadas y que $\frac{W_1}{\|W_1\|_{HS}} = \frac{W_2}{\|W_2\|_{HS}}$.

A través de la métrica que define este producto escalar, podemos observar que la matriz de productos escalares entre individuos tiene la misma norma Hilbert-Schmidt que aquélla que contiene los correspondiente para las variables (matriz de varianzas y covarianzas)

Proposición 2.3.1

$$\|W\|_{HS}^2 = \|V\|_{HS}^2.$$

Demostración. Recordemos que

$$\begin{aligned} w &: (E^*, w) \rightarrow (E, d) & v &: (F^*, v) \rightarrow (F, m) \\ w^* &: (E, d) \rightarrow (E^*, w) & v^* &: (F, m) \rightarrow (F^*, v) \\ W^* &= DWD & V^* &= MVM \end{aligned}$$

con $W = XMX'$ y $V = X'DX$ (ver 1.1), entonces,

$$\begin{aligned} \|W\|_{HS}^2 &= \text{traza}(W^*W) = \text{traza}(DWDW) = \text{traza}(DXMX'DXMX') \\ &= \text{traza}(MX'DXMX'DX) = \text{traza}(MVMV) = \text{traza}(V^*V) \\ &= \|V\|_{HS}^2. \quad \square \end{aligned}$$

La notación que introduciremos nos permitirá distinguir los vectores asociados a individuos y variables en cada uno de los estudios, conservando las ideas correspondientes al caso de una sola matriz observada. Para esta situación más sencilla, habíamos llamado \underline{x}_i al i -ésimo individuo de la tabla y \underline{x}^j a la j -ésima variable, por lo que el subíndice denotaba el número de individuo y el supraíndice la variable, convención que seguiremos usando en esta generalización,

$(\underline{x}_l)_i$, vector asociado al i -ésimo individuo del estudio l , $(x_l)_i \in F_l$.

$(\underline{x}_l)^j$, vector asociado a la j -ésima variable del estudio l , $(x_l)^j \in E$.

W_l , matriz de productos escalares entre individuos de la tabla l , asociada a la transformación lineal $w_l : (E^*, w_l) \rightarrow (E, d)$.

La introducción de una métrica entre *estudios* se hizo con el objeto de establecer una distancia entre las matrices de productos escalares W_l , cuyo cálculo será posible a través del siguiente resultado.

Teorema 2.3.2 *La distancia entre dos matrices de productos escalares entre individuos W_1 y W_2 inducida por la métrica de Hilbert-Schmidt, puede expresarse como*

$$\begin{aligned} d_{HS}^2(W_1, W_2) &= \sum_{i=1}^n \sum_{j=1}^n p_i p_j (W_1 - W_2)_{ij}^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n p_i p_j \left(((x_1)_i, (x_1)_j)_{m_1} - ((x_2)_i, (x_2)_j)_{m_2} \right)^2. \end{aligned}$$

Demostración.

$$d_{HS}^2(W_1, W_2) = \|W_1 - W_2\|_{HS}^2 = \text{traza}((W_1 - W_2)^*(W_1 - W_2)).$$

La traza de una aplicación lineal a sobre un espacio vectorial G con base $\{g_1, \dots, g_n\}$ se define formalmente como

$$\text{traza}(a) = \sum_{i=1}^n (a(g_i), g_i),$$

por lo que debemos especificar sobre qué espacio vectorial estamos trabajando

$$\begin{aligned} (\mathbf{w}_1 - \mathbf{w}_2) : (E^*, w) &\rightarrow (E, d) \\ (\mathbf{w}_1 - \mathbf{w}_2)^* : (E, d) &\rightarrow (E^*, w) \\ (\mathbf{w}_1 - \mathbf{w}_2)^*(\mathbf{w}_1 - \mathbf{w}_2) : (E^*, w) &\rightarrow (E^*, w), \end{aligned}$$

así, utilizando además la definición de operador adjunto tenemos (ver teorema 1.4.1):

$$\begin{aligned} d_{HS}^2(W_1, W_2) &= \sum_{j=1}^n \left((\mathbf{w}_1 - \mathbf{w}_2)^*(\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*), \underline{e}_j^* \right)_w \\ &= \sum_{j=1}^n \left((\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*), (\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*) \right)_d \\ &= \sum_{j=1}^n p_j \left((\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*), (\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*) \right)_{L_n} \\ &= \sum_{j=1}^n p_j \left\| (\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*) \right\|^2, \end{aligned}$$

pero como $(\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*) \in (E, d)$, podemos expresar este vector como combinación lineal de la base $\{\underline{e}_1, \dots, \underline{e}_n\} : (\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*) = \sum_{i=1}^n \left((\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*), \underline{e}_i \right)_d \underline{e}_i$, de donde lo anterior se reduce a

$$\begin{aligned} d_{HS}^2(W_1, W_2) &= \sum_{j=1}^n p_j \sum_{i=1}^n \left((\mathbf{w}_1 - \mathbf{w}_2)(\underline{e}_j^*), \underline{e}_i \right)_d^2 (\underline{e}_i, \underline{e}_i)_d \\ &= \sum_{j=1}^n \sum_{i=1}^n p_j p_i \left((\mathbf{w}_1(\underline{e}_j^*), \underline{e}_i) - (\mathbf{w}_2(\underline{e}_j^*), \underline{e}_i) \right)^2 \\ &= \sum_{j=1}^n \sum_{i=1}^n p_j p_i (W_1 - W_2)_{ij}^2, \end{aligned}$$

y como W_i es la matriz de productos escalares entre individuos de la tabla i

$$(W_1)_{ij} = ((x_1)_i, (x_1)_j)_{m_1},$$

$$(W_2)_{ij} = ((x_2)_i, (x_2)_j)_{m_2},$$

$$\|W_1 - W_2\|_{HS}^2 = \sum_{i=1}^n \sum_{j=1}^n p_i p_j \left(((x_1)_i, (x_1)_j)_{m_1} - ((x_2)_i, (x_2)_j)_{m_2} \right)^2. \quad \square$$

Además, se cuenta con un expresión alternativa para el producto escalar Hilbert-Schmidt entre matrices de este estilo, que permite expresarlo en función de las variables

Teorema 2.3.3

$$(W_1, W_2)_{HS} = \sum_{i=1}^{p_1} \sum_{j=1}^{p_2} \left((x_1)_i', (x_2)_j' \right)_d^2,$$

donde p_i es el número de variables en la tabla i .

Demostración.

$$\begin{aligned} (W_1, W_2)_{HS} &= \text{traza}(W_1^* W_2) = \text{traza}((D W_1 D) W_2) \\ &= \text{traza}(D(X_1 M_1 X_1') D(X_2 M_2 X_2')) \\ &= \text{traza}(M_1 X_1' D X_2 M_2 X_2' D X_1) = \text{traza}(M_1 V_{X_1, X_2} M_2 V_{X_1, X_2}'), \end{aligned}$$

utilizando la definición formal de la traza de una matriz vemos que $M_1 V_{X_1, X_2} M_2 V_{X_1, X_2}'$ es una matriz asociada a una aplicación lineal cuyo dominio y contradominio es (F_1^*, v_1) por lo que esto último puede expresarse como

$$(W_1, W_2)_{HS} = \sum_{i=1}^{p_1} (M_1 V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i, (f_1^*)_i)_{v_1};$$

analizando cada sumando vemos que

$$\begin{aligned} (M_1 V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i, (f_1^*)_i)_{v_1} &= (M_1 V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i)' V_1 (f_1^*)_i = \\ (V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i)' M_1 V_1 (f_1^*)_i &= (V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i, (f_1^*)_i)_I, \end{aligned}$$

por lo que

$$\begin{aligned} (W_1, W_2)_{HS} &= \sum_{i=1}^{p_1} (V_{X_1, X_2} M_2 V_{X_1, X_2}' (f_1^*)_i, (f_1^*)_i)_I \\ &= \sum_{i=1}^{p_1} (M_2 V_{X_1, X_2}' (f_1^*)_i, V_{X_1, X_2} (f_1^*)_i)_I \end{aligned}$$

$$(W_1, W_2)_{HS} = \sum_{i=1}^{p_1} (V'_{X_1, X_2}(\underline{f}_1^*)_i, V'_{X_1, X_2}(\underline{f}_1^*)_i)_{m_2} = \sum_{i=1}^{p_1} \|V'_{X_1, X_2}(\underline{f}_1^*)_i\|_{m_2}^2,$$

pero como $V'_{X_1, X_2}(\underline{f}_1^*)_i \in F_2$, puede expresarse como combinación lineal de la base $\{(\underline{f}_2^*)_j\}_{j=1}^{p_2}$, y que por ser ortogonal se puede aplicar el Teorema de Pitágoras

$$\begin{aligned} (W_1, W_2)_{HS} &= \sum_{j=1}^{p_2} \sum_{i=1}^{p_1} (V'_{X_1, X_2}(\underline{f}_1^*)_i, (\underline{f}_2^*)_j)_{m_2}^2 = \sum_{j=1}^{p_2} \sum_{i=1}^{p_1} ((\underline{f}_1^*)_i, V_{X_1, X_2}(\underline{f}_2^*)_j)^2 \\ &= \sum_{j=1}^{p_2} \sum_{i=1}^{p_1} ((\underline{f}_1^*)_i, X_1' D X_2 (\underline{f}_2^*)_j)^2 = \sum_{j=1}^{p_2} \sum_{i=1}^{p_1} (X_1(\underline{f}_1^*)_i, X_2(\underline{f}_2^*)_j)_d^2, \end{aligned}$$

de donde finalmente por la expresión 2.1

$$(W_1, W_2)_{HS} = \sum_{i=1}^{p_1} \sum_{j=1}^{p_2} ((x_1)^i, (x_2)^j)_d^2. \quad \square$$

Teorema 2.3.4 Si dos matrices W_1, W_2 son HS-ortogonales, entonces las variables del primer estudio tienen covarianza cero con aquéllas del segundo estudio, además $(X_1, X_2)_{HS} = 0$.

Demostración. Dos matrices W_1, W_2 son HS-ortogonales si $(W_1, W_2)_{HS} = 0$, de donde retornando la demostración del teorema 2.3.3 tenemos que

$$(W_1, W_2)_{HS} = \sum_{j=1}^{p_2} \sum_{i=1}^{p_1} (V'_{X_1, X_2}(\underline{f}_1^*)_i, (\underline{f}_2^*)_j)_{m_2}^2 \geq 0,$$

donde la igualdad se da sólo si $V_{X_1, X_2} = 0$, es decir $\forall i \leq p_1$ y $j \leq p_2$ $((x_1)^i, (x_2)^j)_d = 0$, lo que significa que las variables de la matriz observada X_1 tienen covarianza cero con las de X_2 .

Por último, si $0 = V_{X_1, X_2} = X_1' D X_2 \Rightarrow 0 = \text{traza}(M(0)) = \text{traza}(M X_1' D X_2) = (X_1, X_2)_{HS}$. \square

En contraposición a la idea de *Intraestructura*, que busca describir la estructura de los individuos al interior de un estudio, llamaremos *Interestructura* a aquel elemento que pretende describir las relaciones entre estudios. Para esto se construye una forma bilineal simétrica positiva w , llamada *Compromiso*³, que nos permitirá conocer la posición de cada uno de los individuos al compararlos contra un sistema de referencia, que se considera la estructura común al cambiar las condiciones de los estudios.

³Este término se tradujo directamente del francés "compromis" y pretende reflejar la idea de un objeto intermedio entre varios de su tipo.

Por último, a partir de la *Interestructura* se construyen n "trayectorias" que describan la evolución individual de los sujetos durante el período de observación. Por otra parte, el análisis de estas curvas ayuda a detectar algún cambio drástico al pasar de un período a otro.

2.3.2 Estructura común de las variables a través de los estudios.

Esta situación puede encontrarse en condiciones como la descrita en el punto dos de la introducción, en la que un conjunto fijo de p variables se observó sobre k diferentes grupos de individuos, los cuales se interpretan como elementos de un mismo espacio vectorial $(F, m) = \mathbb{R}^p$. Por otra parte, el conjunto de variables formará un espacio vectorial $(E_j, d_j) = \mathbb{R}^{n_j}$, $j \in \{1, \dots, k\}$ cuya dimensión dependerá de la cardinalidad (n_j) del grupo de individuos sobre el cual se mida; así, la información disponible puede resumirse en k ternas del estilo

$$(X_1, m, d_1), \dots, (X_k, m, d_k),$$

La estructura vectorial que soporta este tipo de situaciones se describe en el siguiente *Diagrama de Dualidad* restringido al caso $k = 2$,

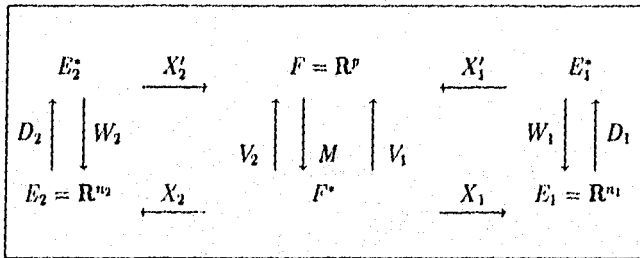


diagrama 4

donde

$$\begin{aligned} V_1 &= X_1' D_1 X_1 & W_1 &= X_1 M X_1' \\ V_2 &= X_2' D_2 X_2 & W_2 &= X_2 M X_2' \end{aligned} \quad \text{y}$$

Lo que buscamos entonces, es comparar las aplicaciones lineales $v_j \in \mathcal{L}(F^*, F)$ asociadas a las matrices de varianzas y covarianzas con una matriz *compromiso* deducida del producto escalar de Hilbert-Schmidt. Puesto que estas matrices (o sus respectivas aplicaciones lineales) son los instrumentos que resumen la información relevante a las relaciones entre variables en cada uno de los *estudios*, el *compromiso* será una matriz de varianzas y covarianzas que describa las relaciones comunes entre variables al ir cambiando de grupo de individuos. Análogamente al caso en el que se buscaba describir el comportamiento de los sujetos, también podremos construir trayectorias para cada una de las variables, que nos permitan conocer su comportamiento a nivel marginal y así detectar cuáles de ellas son las responsables de las desviaciones sufridas con respecto a la matriz *compromiso*.

De esta manera, aplicamos el producto escalar HS a las matrices de varianzas y covarianzas que queremos comparar

$$(V_1, V_2)_{HS} = \text{traza}(V_1^* V_2),$$

y definimos la equivalencia entre dos de ellas, cuando éstas sean iguales y además provengan de matrices observadas centradas previamente en sus respectivos *estudios*. Igualmente al caso anterior, se puede optar por la aplicación del producto escalar HS sobre las matrices normalizadas.

A continuación se presentan algunos resultados análogos a la comparación entre matrices de productos escalares entre individuos.

Teorema 2.3.5

$$d_{HS}^2(V_1, V_2) = \|V_1 - V_2\|_{HS}^2 = \sum_{i=1}^p \sum_{j=1}^p (V_1 - V_2)_{ij}^2,$$

donde p es el número de variables en cada uno de los *estudios*.

Demostración. Es totalmente análoga a la del teorema 2.3.2. \square

Ahora presentamos una expresión alternativa para el producto escalar HS entre dos matrices de varianzas y covarianzas en función de los datos observados

Proposición 2.3.2

$$(V_1, V_2)_{HS} = \sum_{i=1}^p \sum_{j=1}^p \left((\underline{x}_1)^i, (\underline{x}_1)^j \right)_{d_1} \left((\underline{x}_2)^i, (\underline{x}_2)^j \right)_{d_2}.$$

Demostración.

$$\begin{aligned} (V_1, V_2)_{HS} &= \text{traza}(V_1^* V_2) = \text{traza}(M V_1 M V_2) \\ &= \sum_{j=1}^p \left(\underline{f}_j^*, M V_1 M V_2 (\underline{f}_j^*) \right)_{v_1} = \sum_{j=1}^p \left(V_1 (\underline{f}_j^*), V_2 (\underline{f}_j^*) \right)_m \\ &= \sum_{i=1}^p \sum_{j=1}^p (V_1 (\underline{f}_j^*), \underline{f}_i)_m (V_2 (\underline{f}_j^*), \underline{f}_i)_m \\ &= \sum_{i=1}^p \sum_{j=1}^p \left((\underline{x}_1)^i, (\underline{x}_1)^j \right)_{d_1} \left((\underline{x}_2)^i, (\underline{x}_2)^j \right)_{d_2}. \quad \square \end{aligned}$$

2.3.3 Comparación entre las matrices de datos.

Para comparar dos *estudios* (X_1, m, d) , (X_2, m, d) en los que se observó un grupo fijo de variables sobre un mismo conjunto de individuos, a diferentes tiempos calendario o al cambiar algunas condiciones externas, se propone utilizar el producto escalar Hilbert-Schmidt (ver proposición 2.3.1) sobre las matrices X_1 y X_2 , lo que equivale a comparar las aplicaciones lineales a'_1 y a'_2 que las originaron (ver ??)

$$\begin{aligned} (a'_1, a'_2)_{HS} &= \text{traza}((a'_1)^* a'_2) = \text{traza}((a'_1)' a'_2) = \text{traza}((d a'_1 m)' a'_2) \\ &= \text{traza}(m' a_1 d' a'_2) = \text{traza}(m a_1 d a'_2), \end{aligned}$$

de donde se deduce una medida de distancia entre las matrices correspondientes a cada *estudio*

$$\begin{aligned} d(X_1, X_2) &= \|X_1 - X_2\|_{HS} = \sqrt{(X_1 - X_2, X_1 - X_2)_{HS}} \\ &= \sqrt{\text{traza}(M(X_1 - X_2)' D(X_1 - X_2))} \\ &= \sqrt{\text{traza}(M X_1' D X_1 - M X_2' D X_1 - M X_1' D X_2 + M X_2' D X_2)} \\ &= \sqrt{\text{traza}(M X_1' D X_1) - 2 \text{traza}(M X_1' D X_2) + \text{traza}(M X_2' D X_2)} \\ &= \sqrt{\|X_1\|_{HS}^2 - 2(X_1, X_2)_{HS} + \|X_2\|_{HS}^2}. \end{aligned}$$

2.4 Imagen euclidiana de los estudios.

En el capítulo anterior se presentó la construcción de la imagen euclidiana de los individuos y de las variables que participan en un *estudio* que nos permitió contar

con una entidad gráfica de dimensión menor a la original, donde se representaron "lo mejor posible" (en el sentido de pérdida mínima de inercia) los objetos que tratamos de describir.

Ahora que contamos con k estudios, buscamos una representación análoga para los estudios mismos, por lo que pasaremos a describir su estructura algebraica siguiendo las ideas expuestas previamente, cuyo punto central consiste en la interpretación de los objetos de interés como elementos de un espacio vectorial donde el producto escalar asignado defina la métrica que nos convenga usar. Así, llamaremos

(G, s) , espacio de estudios cuya dimensión es igual al número de matrices de datos que intervengan en el análisis, el cual ha sido llamado hasta ahora k con base $\{g_1, \dots, g_k\}$ y definamos el producto escalar s dependiendo de la situación particular de nuestro problema:

$$(g_i, g_j)_s = \begin{cases} (V_i, V_j)_{ns} & \text{si se busca comparar matrices} \\ & \text{de varianzas y covarianzas,} \\ (W_i, W_j)_{ns} & \text{si se comparan matrices de} \\ & \text{productos escalares entre individuos,} \end{cases}$$

de tal manera que su matriz asociada sea

$$S_{ij} = (g_i, g_j)_s,$$

aquella que tiene por entradas los productos escalares Hilbert-Schmidt entre las matrices que estamos interesados en comparar.

Como se probó en el teorema 1.2.1 s tiene asociada una aplicación lineal s que es un elemento de su espacio dual, es decir $s \in G^*$, que le asignaremos un producto escalar conveniente (δ) que refleje los pesos (π_i) de los estudios atendiendo a su importancia relativa en el análisis. Si $\{g_1^*, \dots, g_k^*\}$ es una base de G^* , definimos el producto escalar δ de la siguiente manera

$$(g_i^*, g_j^*)_\delta = \begin{cases} \pi_i & \text{si } i = j \in \{1, \dots, k\} \\ 0 & \text{en otro caso} \end{cases}$$

cuya matriz asociada se denotará por $\Delta = \text{diag}\{\pi_1, \dots, \pi_k\}$.

En el capítulo 1 se explicó con todo detalle, cómo para construir una imagen euclidiana de los individuos, elementos del espacio vectorial (E, m) cuyo dual se denotó por (E^*, v) con matrices asociadas M y V respectivamente, se tuvo que descomponer en valor singular el producto escalar v a través de los valores y vectores propios de la matriz VM . Asimismo, se analizó el caso en el que los elementos de interés fueran las variables, elementos del espacio vectorial (E, d) cuyo dual es (E^*, w) y lo que nuevamente se propuso fue la descomposición en valor singular de w a través de los valores y vectores propios de la matriz WD , haciendo un razonamiento similar. ahora trabajamos sobre los espacios vectoriales (G, s) y (G^*, t) de tal suerte que se propondrá la descomposición en valor singular de s auxiliada por los eigenvalores y eigenvectores de la matriz $S\Delta$.

Se debe notar que la matriz $S\Delta$ es simétrica sólo cuando todos los π_i se consideran de igual importancia, dando lugar a $\Delta = \text{diag}\{1/n, \dots, 1/n\}$, razón por la cual los vectores propios \underline{z}^* (ortonormalizados) tendrán que derivarse de los vectores propios \underline{z} (ortonormalizados) de la matriz simétrica $\Delta^{1/2} S \Delta^{1/2}$ cuya entrada i -ésima es:

$$(\Delta^{1/2} S \Delta^{1/2})_{ij} = S_{ij} \sqrt{\pi_i} \sqrt{\pi_j}.$$

Proposición 2.4.1 Si \underline{z}^* y \underline{z} son vectores propios de la matriz $S\Delta$ y $\Delta^{1/2} S \Delta^{1/2}$ respectivamente, entonces:

$$\underline{z}^* = \Delta^{1/2} \underline{z}$$

si además \underline{z} es unitariedad, \underline{z}^* también lo será pero con respecto a la norma asociada por Δ .

Demostración. Sea λ_i el valor propio asociado al vector \underline{z}_i entonces:

$$\Delta^{1/2} S \Delta^{1/2} \underline{z}_i = \lambda_i \underline{z}_i$$

multiplicando los dos miembros por la matriz $\Delta^{-1/2} = \text{diag}\{1/\pi_1, \dots, 1/\pi_k\}$ tenemos:

$$\begin{aligned} \Delta^{-1/2} \Delta^{1/2} S \Delta^{1/2} \underline{z}_i &= S \Delta^{-1/2} \Delta^{1/2} \underline{z}_i \\ &= \Delta^{-1/2} \Delta^{-1/2} \Delta^{1/2} S \Delta^{1/2} \underline{z}_i = \Delta^{-1} S \Delta \underline{z}_i \end{aligned}$$

cuya matriz asociada se denotará por $\Delta = \text{diag} \{ \pi_1, \dots, \pi_k \}$.

En el capítulo 1 se explicó con todo detalle, cómo para construir una imagen euclidiana de los individuos, elementos del espacio vectorial (F, m) cuyo dual se denotó por (F^*, v) con matrices asociadas M y V respectivamente, se tuvo que descomponer en valor singular el producto escalar v a través de los valores y vectores propios de la matriz VM . Asimismo, se analizó el caso en el que los elementos de interés fueran las variables, elementos del espacio vectorial (E, d) cuyo dual es (E^*, w) y lo que nuevamente se propuso fue la descomposición en valor singular de w a través de los valores y vectores propios de la matriz WD . Siguiendo un razonamiento similar, ahora trabajamos sobre los espacios vectoriales (G, s) y (G^*, δ) de tal suerte que se propondrá la descomposición en valor singular de s auxiliados por los eigenvalores y eigenvectores de la matriz $S\Delta$.

Se debe notar que la matriz $S\Delta$ es simétrica sólo cuando todos los *estudios* se consideran de igual importancia, dando lugar a $\Delta = \text{diag} \{ 1/n, \dots, 1/n \}$, razón por la cual los vectores propios γ_i^* δ -ortonormalizados tendrán que deducirse de los vectores propios z_i (ortonormalizados) de la matriz simétrica $\Delta^{1/2}S\Delta^{1/2}$ cuya entrada ij -ésima es:

$$\left(\Delta^{1/2}S\Delta^{1/2} \right)_{ij} = S_{ij} \sqrt{\pi_i} \sqrt{\pi_j}.$$

Proposición 2.4.1 Si γ_i^* y z_i son vectores propios de la matriz $S\Delta$ y $\Delta^{1/2}S\Delta^{1/2}$ respectivamente, entonces

$$\gamma_i^* = \Delta^{1/2} z_i,$$

si además z_i es ortonormal, γ_i^* también lo será pero con respecto a la norma inducida por δ .

Demostración. Sea λ_i el valor propio asociado al vector z_i , entonces

$$\Delta^{1/2}S\Delta^{1/2} z_i = \lambda_i z_i,$$

multiplicando los dos miembros por la matriz $\Delta^{-1/2} = \text{diag} \{ (\pi_1)^{-1/2}, \dots, (\pi_k)^{-1/2} \}$ tenemos

$$\begin{aligned} S\Delta^{1/2} z_i &= \lambda_i \Delta^{-1/2} z_i \\ \Rightarrow S\Delta \left(\Delta^{-1/2} z_i \right) &= \lambda_i \left(\Delta^{-1/2} z_i \right), \end{aligned}$$

lo que nos dice que $\gamma_i^* = \Delta^{-1/2} z_i$ es un vector propio de la matriz $S\Delta$ asociado al mismo valor característico λ_i .

Si además z_i es ortonormal, tenemos que

$$\|\gamma_i^*\|_s = (\gamma_i^*, \gamma_i^*)_s = \gamma_i^{*'} \Delta \gamma_i^* = (\Delta^{-1/2} z_i)' \Delta (\Delta^{-1/2} z_i) = z_i' z_i = 1. \quad \square$$

Corolario 2.4.1 *El procedimiento descrito en la proposición anterior, genera vectores γ_i^* δ -ortogonales.*

Demostación. Este resultado se hereda de la ortogonalidad de los vectores propios z_i

$$(\gamma_i^*, \gamma_j^*)_s = \gamma_i^{*'} \Delta \gamma_j^* = (\Delta^{-1/2} z_i)' \Delta (\Delta^{-1/2} z_j) = z_i' \Delta^{-1/2} \Delta \Delta^{-1/2} z_j = z_i' z_j = 0. \quad \square$$

Una vez realizada la descomposición en valor singular de s a través de la matriz $S\Delta$ podemos proceder a construir la imagen euclidiana de los estudios, sobre un espacio cuya dimensión se fijará atendiendo a las necesidades del problema, si se busca una representación gráfica ésta tendrá que ser 2 ó 3. Así, cada uno de ellos se representa por un punto M_i ($i = 1, \dots, k$) cuya l -ésima coordenada se obtiene a través de la expresión (consultar sección 1.5.2)

$$\sqrt{\lambda_l} \gamma_i^* = \lambda_l^{-1/2} S \Delta \gamma_i^*,$$

de tal manera que si la imagen euclidiana deseada es de dimensión h , las coordenadas de dicha imagen se calcularán como

$$\left(\sqrt{\lambda_1} \gamma_1^*, \sqrt{\lambda_2} \gamma_2^*, \dots, \sqrt{\lambda_h} \gamma_h^* \right),$$

con $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_h$, lo que puede resumirse matricialmente en la expresión

$$S \Delta \Gamma \Lambda^{-1/2} = \Gamma \Lambda^{1/2},$$

donde Γ es la matriz de eigenvectores $(\gamma_1^*, \dots, \gamma_h^*)$ y $\Lambda = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_h \}$.

Proposición 2.4.2 *Si los pesos asignados a cada uno de los estudios son iguales, basta diagonalizar la matriz S para construir su imagen euclidiana.*

Demostración. Si \underline{z} es un vector propio normado de la matriz S con valor propio asociado $\lambda(S)$, entonces $S\underline{z} = \lambda(S)\underline{z}$ por lo que la matriz $\Delta^{1/2}S\Delta^{1/2} = S\Delta = \frac{1}{k}S$ admite como vector propio a \underline{z} pero asociado al valor propio $\frac{\lambda(S)}{k}$ que denotaremos por $\lambda(S\Delta)$

$$S\Delta\underline{z} = \frac{1}{k}S\underline{z} = \frac{\lambda(S)}{k}\underline{z},$$

así, las i -ésimas coordenadas de los puntos M_i son los elementos del vector

$$\sqrt{\lambda_i(S\Delta)}\underline{\gamma}_i^* = \sqrt{\frac{\lambda(S)}{k}}\underline{\gamma}_i^* = \sqrt{\frac{\lambda(S)}{k}}\sqrt{k}\underline{z} = \sqrt{\lambda}\underline{z}. \quad \square$$

Ahora procederemos a analizar la forma que tiene la imagen euclidiana propuesta.

La longitud del vector \overline{OM}_i que representa la norma del i -ésimo objeto, será la misma si el análisis se basara, ya sea en las matrices de productos escalares entre individuos W_i o bien en las matrices de varianzas y covarianzas V_i , lo anterior se cumple por el resultado probado el teorema 2.3.1, $\|W_i\|_{HS} = \|V_i\|_{HS}$. Esto indica que el punto asociado a un estudio en la imagen euclidiana que se construyó tiene una distancia fija al origen, independientemente de lo que se quiera describir. Cuando los objetos son normados, los vectores asociados también lo serán, por lo que $\|\overline{OM}_i\| = 1$ y la imagen se concentraría en un círculo de radio 1.

Además se vió en la demostración del teorema 2.3.3 que, los productos escalares entre objetos $(W_i, W_j)_{HS}$ son positivos, lo que implica que el coseno del ángulo que forman los vectores en la imagen euclidiana sea agudo y los vectores $\overline{OM}_1, \dots, \overline{OM}_k$ determinan un cono convexo. Si partimos de matrices normalizadas, este cono coincidirá con un semiarco del círculo unitario.

2.5 Estudios Suplementarios.

Para fines prácticos, pensemos que necesitamos comparar las matrices de productos escalares entre individuos W_i , (si nuestro objetivo fuera la comparación de matrices de varianzas y covarianzas, el procedimiento sería completamente análogo) de tal manera que contamos con estudios de la forma (X_i, m_i, d) .

Un *estudio* suplementario (X_{k+1}, m_{k+1}, d) es aquél al que afectamos por un peso nulo para incluirlo en nuestro análisis con fines comparativos, atendiendo posiblemente a la existencia de alguna variable categórica de interés, por lo que el espacio vectorial k -dimensional (G, σ) de *estudios* que habíamos trabajado, ahora será de dimensión $k + 1$, y lo denotaremos por (G_1, s_1) .

La forma de trabajar este tipo de *estudios* es igual que cualquier otro objeto suplementario. Calculamos los productos escalares entre la matriz de productos escalares entre individuos correspondiente al *estudio* suplementario en cuestión W_{k+1} , con aquéllas de los elementos activos

$$(g_i, g_{k+1})_{s_1} = (W_i, W_{k+1})_{HS} \quad \forall i = \{1, \dots, k\}.$$

La representación de este *estudio* en la imagen euclídiana de todo el análisis, tendrá asociado el punto M_{p+1} cuya l -ésima coordenada está dada por

$$\lambda_l^{-1/2} (S\Delta) s'_l(g_{k+1}) \Delta \gamma_l^*,$$

donde $s'_l(g_{k+1})$ es el vector de dimensión k que tiene en cada entrada el producto escalar Hilbert-Schmidt con las k matrices de los *estudios* activos.

2.6 Construcción del *Compromiso*.

Como se mencionó anteriormente, la matriz *compromiso* busca ser un buen resumen de los *estudios* involucrados en el análisis, de tal manera que pueda considerarse como una matriz "representativa" de los sucedido a través de los períodos de observación. Empezaremos por presentar su definición y posteriormente mencionaremos algunas propiedades que posee.

Definición 2.6.1 Sea (g_l, γ_l^*) $l = 1, \dots, k$ las coordenadas del vector propio γ_l^* de la matriz $S\Delta$ asociado al valor propio más grande λ_l en la base $\{g_1^*, \dots, g_k^*\}$. El *compromiso* w es la aplicación lineal de E en E^* definida como combinación lineal de las aplicaciones w_l de la manera siguiente

$$w = \alpha \sum_{l=1}^k \pi_l(g_l, \gamma_l^*) w_l,$$

lo que en notación matricial equivale a

$$W = \alpha \sum_{l=1}^k \sum_{i=1}^k \pi_l(q_l, \gamma_l^*) W_l,$$

donde α es un coeficiente positivo de normalización que se especificará más adelante.

Es decir, cada sumando se obtiene premultiplicando la matriz W_l por la l -ésima coordenada del vector propio γ_l^* y el respectivo peso del estudio. Una de las ventajas de definir de esta manera el *compromiso*, es que hereda la simetría y positividad de las matrices (aplicaciones) que lo generaron.

El coeficiente de normalización α , se escoge de tal manera que la matriz *compromiso* resulte ser del mismo tipo de objetos que intervinieron en su construcción. En el caso de comparar objetos W_l , éste se propone de tal manera que su norma resulte ser igual a la media de las normas de las matrices involucradas, es decir

$$\begin{aligned} \|W\| &= \alpha \sqrt{\lambda_1} = \sum_{i=1}^k \pi_i \|W_i\| \\ \Rightarrow \alpha &= \lambda_1^{-1/2} \sum_{i=1}^k \pi_i \|W_i\| = \lambda_1^{-1/2} \sum_{i=1}^k \pi_i S_i^{1/2}, \end{aligned}$$

de tal forma que la expresión final de la matriz *compromiso* resulta ser

$$W = \left(\lambda_1^{-1/2} \sum_{i=1}^k \pi_i \|W_i\| \right) \sum_{l=1}^k \pi_l(q_l, \gamma_l^*) W_l.$$

Equivalentemente podemos proponer el *compromiso* como $W = \sum_i \alpha_i W_i$, donde los coeficientes α_i son las componentes del vector $\alpha = \lambda_1^{-1/2} \left(\sum_{i=1}^k \pi_i S_i^{1/2} \right) \Delta \gamma_l^*$, lo que se reduce a $\alpha = \lambda_1^{-1/2} \Delta \gamma_l^*$ cuando los objetos son normados.

La representación de esta matriz en la imagen euclidiana, se explica en el siguiente teorema.

Teorema 2.6.1 *El punto asociado a la matriz compromiso en la imagen euclidiana de los estudios tiene coordenadas*

$$\left(\alpha \sqrt{\lambda_1}, 0, \dots, 0 \right).$$

Demostración. La coordenada l -ésima del *compromiso* es una combinación lineal de las coordenadas l -ésimas del punto asociado a cada W_i , las cuales se concentran en el vector $\lambda_l^{1/2} \gamma_l^*$ que siguiendo la notación introducida en la definición, cada una puede expresarse como $\lambda_l^{1/2}(g_i, \gamma_l^*)$, por lo tanto, la l -ésima coordenada de la matriz *compromiso* está dada por

$$\alpha \sum_{i=1}^k \pi_i(g_i, \gamma_l^*) \lambda_l^{1/2}(g_i, \gamma_l^*) = \alpha \lambda_l^{1/2} (\gamma_l^*, \gamma_l^*)_\delta,$$

pero los vectores propios $\{\gamma_l^*\}$ forman una base δ -ortonormada, por lo que

$$(\gamma_l^*, \gamma_l^*)_\delta = \begin{cases} 1 & \text{si } l = 1 \\ 0 & \text{en otro caso} \end{cases}$$

de tal suerte que las coordenadas son las que se propusieron en el enunciado del teorema. \square

En el caso de estar comparando matrices de varianzas y covarianzas V_i , la matriz *compromiso* se define de manera análoga a la matriz W

$$V = \beta \sum_{i=1}^k \pi_i(g_i, \gamma_1^*) V_i,$$

con β el coeficiente de normalización. Ahora bien, si la interpretamos directamente como media ponderada de V_i , se escribirá como

$$V = \sum_{i=1}^k \alpha_i V_i, \quad \text{con } \sum_{i=1}^k \alpha_i = 1,$$

de tal manera que

$$\alpha_i = \frac{\pi_i(g_i, \gamma_1^*)}{\sum_{i=1}^k \pi_i(g_i, \gamma_1^*)},$$

con lo que la expresión final de la matriz *compromiso* V es

$$V = \frac{1}{\sum_{i=1}^k \pi_i(g_i, \gamma_1^*)} \sum_{i=1}^k \pi_i(g_i, \gamma_1^*) V_i.$$

Podríamos decir que hasta ahora, no parece haber alguna razón que justifique las expresiones propuestas para las matrices V y W , sin embargo debemos tener en cuenta que se buscaban matrices altamente correlacionadas con las matrices V_i y W_i respectivamente, de tal manera que los coeficientes de las combinaciones lineales satisfagan este objetivo. Lo anterior se explica en el siguiente teorema.

Teorema 2.6.2 Entre las combinaciones lineales de las matrices W_i , el compromiso $W = \sum_{i=1}^k \alpha_i W_i$ presentado en la definición 2.6.1 tiene la propiedad de ser la de correlación máxima con las matrices W_i .

Demostración. Probar este teorema equivale a probar que el máximo de

$$\frac{\sum_{i=1}^k \pi_i \left(\sum_{l=1}^k \alpha_l W_l, W_i \right)_{HS}^2}{\left\| \sum_{l=1}^k \alpha_l q_l \right\|_s^2} = \frac{\sum_{i=1}^k \pi_i \text{Cov}^2(W, W_i)}{\left\| \sum_{l=1}^k \alpha_l q_l \right\|_s^2},$$

se alcanza precisamente para el compromiso como se definió previamente.

Usando $(W_i, W_i)_{HS} = (q_i, q_i)_s = (q_i, s(q_i))$ y que $\delta(q_i^*) = \pi_i q_i$, vemos que el numerador es igual a

$$\begin{aligned} & \sum_i \pi_i \left(\sum_{l_1} \alpha_{l_1} W_{l_1}, W_i \right)_{HS} \left(\sum_{l_2} \alpha_{l_2} W_{l_2}, W_i \right)_{HS} \\ &= \sum_{l_1} \sum_{l_2} \alpha_{l_1} \alpha_{l_2} \underbrace{\sum_i \pi_i (W_{l_1}, W_i)_{HS} (W_{l_2}, W_i)_{HS}}. \end{aligned}$$

Analizando únicamente el término marcado con la llave, vemos que

$$\begin{aligned} & \sum_i \pi_i (W_{l_1}, W_i)_{HS} (W_{l_2}, W_i)_{HS} = \sum_i \pi_i (q_{l_1}, s(q_{l_1})) (q_{l_2}, s(q_{l_2})) \\ &= \sum_i (q_{l_1}, s(\pi_i q_{l_1})) (q_{l_2}, s(q_{l_2})) = \sum_i (q_{l_1}, s\delta(q_{l_1}^*)) (q_{l_2}, s(q_{l_2})) \\ &= (q_{l_1}, s\delta(\sum_i (q_{l_1}, s(q_{l_2})) q_i^*)) = (q_{l_1}, s\delta s(q_{l_2})), \end{aligned}$$

por lo que el numerador se reduce a

$$\begin{aligned} & \sum_{l_1} \sum_{l_2} \alpha_{l_1} \alpha_{l_2} (q_{l_1}, s\delta s(q_{l_2})) = \left(\sum_{l_1} \alpha_{l_1} q_{l_1}, \sum_{l_2} \alpha_{l_2} s\delta s(q_{l_2}) \right) \\ &= \left(\sum_{l_1} \alpha_{l_1} q_{l_1}, s\delta s \left(\sum_{l_2} \alpha_{l_2} q_{l_2} \right) \right). \end{aligned}$$

Si llamamos $\underline{x} = \sum_{i=1}^k \alpha_i q_i$, lo anterior se reduce a

$$(\underline{x}, s\delta s(\underline{x})) = (\underline{x}, \delta s\delta s(\underline{x}))_s = (\underline{x}, s^* s(\underline{x}))_s = (s^* s(\underline{x}), \underline{x})_s.$$

Analizando ahora el denominador, vemos que

$$\left\| \sum_{l=1}^k \alpha_l q_l \right\|_s^2 = \left(\sum_{l=1}^k \alpha_l q_l, \sum_{l=1}^k \alpha_l q_l \right)_s = (\underline{x}, \underline{x})_s,$$

de tal manera que el cociente original no es más que el cociente de Rayleigh para la matriz S^*S (para el operador s^*s) y alcanza el máximo cuando \underline{x} es el vector propio

de la matriz S^*S asociado al valor propio más grande de dicha matriz, el cual se denotará por $\lambda_1(S^*S) = \lambda_1^2(S\Delta)$ y puede deducirse de los vectores propios $\{\xi_i^*\}$ de $S\Delta$ de la manera siguiente:

Si $\{\xi_i^*\}$ forma una base de vectores propios ordenados en forma decreciente atendiendo a la magnitud de su respectivo valor propio de $S\Delta$, entonces el conjunto $\{\delta(\xi_i^*) = \xi_i\}$ forma una base dual de vectores propios ordenados en el mismo sentido, de la matriz transpuesta ΔS por lo que

$$\begin{aligned}\delta s(\xi_1) &= \lambda_1 \xi_1 \\ \Rightarrow \delta s \delta s(\xi_1) &= \lambda_1 \delta s(\xi_1) \\ \Rightarrow s^* s(\xi_1) &= \lambda_1^2 \xi_1,\end{aligned}$$

lo que indica que ξ_i es vector propio de s^*s asociado también al valor propio máximo (si $\max_i \{\lambda_i\} = \lambda_1$ entonces $\max_i \{\lambda_i^2\} = \lambda_1^2$ pues $\lambda_i \geq 0$), lo que implica que $\xi_1 = \underline{x} = \delta(\xi_1^*)$, es decir, $\delta^{-1}(\underline{x}) = \xi_1$ es vector propio de $s\delta$ asociado a λ_1 , de tal suerte que éste tiene que ser forzosamente γ_1^* o bien $\alpha\gamma_1^*$.

Si identificamos el espacio vectorial G con su bidual y hacemos notar que la matriz Δ es simétrica y $W = \sum \alpha_l W_l$, podemos ver que

$$\alpha_l = (\underline{x}, q_l^*) = \alpha (\delta(\gamma_1^*), q_l^*) = \alpha (\gamma_1^*, \delta(q_l^*)) = \alpha \pi_l(q_l, \gamma_1^*) \cdot \square$$

Para concluir esta sección, sólo conviene notar que la forma propuesta de construcción del *compromiso* permite garantizar que éste resultará ser una matriz centrada cuando provenga de matrices $\{W_i\}_{i=1}^k$ que también lo sean

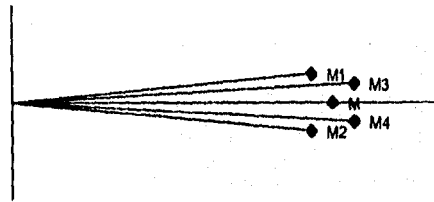
$$\sum_{i=1}^n p_i W_{ij} = \sum_{i=1}^n p_i \sum_{l=1}^k \alpha_l (W_l)_{ij} = \sum_{l=1}^k \alpha_l \sum_{i=1}^n p_i (W_l)_{ij} = \sum_{l=1}^k \alpha_l (0) = 0.$$

2.6.1 Interpretación del *compromiso*.

Centremos nuestro análisis en el caso en que contamos con *estudios* de la forma (X_i, m_i, d) y se comparan matrices de productos escalares entre individuos W_i , es decir, tenemos un conjunto fijo de individuos sobre los cuales se han medido diferentes grupos de variables.

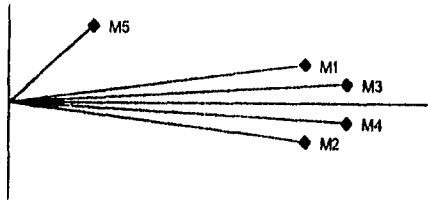
En la sección anterior se presentó la forma de construir una imagen euclidiana para los *estudios*, en la cual también pudiera representarse la matriz *compromiso*; la notación empleada fue la siguiente: llamamos M_i al punto en el espacio euclidiano asociado a la matriz W_i y el punto M al correspondiente para el *compromiso* W . Esto se hizo de tal forma que dos matrices W_i similares, tendrían normas en la imagen también similares.

A continuación presentaremos algunos ejemplos de imágenes que podríamos encontrarnos en alguna situación práctica y la interpretación que deberíamos hacer de ella.



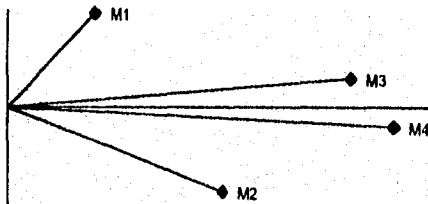
Gráfica No. 1

Esta gráfica muestra evidencia de que existe estructura común de los individuos al cambiar de conjunto de variables observadas, es decir, al pasar de un *estudio* a otro; esta estructura puede resumirse en una sola matriz, el *compromiso*. Las coordenadas en el primer eje son muy parecidas, lo que refleja aportaciones semejantes de cada uno de los estudios que intervinieron en la construcción de W .



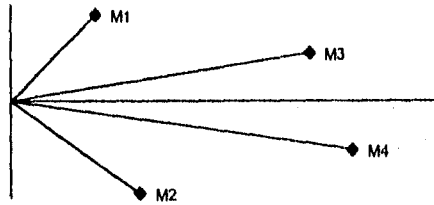
Gráfica No. 2

La gráfica 2 presenta una situación en la que el quinto estudio es diferente a los demás, pues se encuentra casi-perpendicular al conjunto restante, lo cual podría interpretarse como evidencia de que el conjunto de variables que se midieron en este quinto estudio son casi ortogonales a las medidas en el resto, además contribuye muy poco en la construcción del *compromiso* pues su primera coordenada es notablemente menor que la de los demás. En este caso, el W resumirá el comportamiento de la "mayoría" de los estudios.



Gráfica No. 3

En la gráfica 3, las normas de los *estudios* son semejantes, no obstante el primero y segundo tienen una primera coordenada muy pequeña lo que indica que contribuyen muy poco a la construcción del *compromiso*. Se recomienda normalizarlos primero y construir la matriz W con $\frac{W_i}{\|W_i\|_{H^3}}$.



Gráfica No. 4

En la gráfica 4 se puede observar que los cosenos de los ángulos formados entre la representación de cada *estudio*, son muy pequeños, lo que indica una posible falta de estructura común de los individuos a lo largo de la investigación. El *compromiso* simplemente es la media ponderada de los objetos que lo conformaron y es el más correlacionado con ellos.

Al considerar que la matriz W es un buen resumen de lo que acontece al interior de cada W_i , entonces podemos utilizarlo para construir una "Imagen Euclidiana *Compromiso*" de los individuos, es decir, una representación gráfica que resuma todo su comportamiento a lo largo del período de observación.

Para obtener esta imagen, basta repetir lo propuesto para la construcción correspondiente a una matriz W_i cualquiera (ver capítulo 1), es decir, descomponer en valor singular el *compromiso* W a través de los valores propios no nulos $\lambda_1 \geq \dots \geq \lambda_p$ y de los vectores propios d -ortonormales asociados, $\{\underline{e}_i\}_{i=1}^p$, del operador WD de (E, d) . A partir de esta descomposición obtener los puntos O, M_1, \dots, M_n en el espacio

euclidiano que le corresponde a cada uno de los individuos.

La distancia del punto M_i a M_j , d_{M_i, M_j} , representa la distancia *compromiso* entre los individuos i y j coincidiendo su cuadrado con el cuadrado de las distancias entre los individuos en cada uno de los *estudios*, ponderadas por los coeficientes α_l . Llamemos $(\underline{x}_w)_i$ al i -ésimo individuo en el *compromiso*, entonces

$$\begin{aligned} d_{M_i, M_j}^2 &= \|(\underline{x}_w)_i - (\underline{x}_w)_j\|_d^2 = ((\underline{x}_w)_i - (\underline{x}_w)_j, (\underline{x}_w)_i - (\underline{x}_w)_j)_d \\ &= ((\underline{x}_w)_i, (\underline{x}_w)_i)_d + ((\underline{x}_w)_j, (\underline{x}_w)_j)_d - 2((\underline{x}_w)_i, (\underline{x}_w)_j)_d \\ &= W_{ii} + W_{jj} - 2W_{ij} = \sum_{l=1}^k \alpha_l ((W_l)_{ii} + (W_l)_{jj} - 2(W_l)_{ij}) \\ &= \sum_{l=1}^k \alpha_l \|(\underline{x}_l)_i - (\underline{x}_l)_j\|_d^2. \end{aligned}$$

Como es de suponerse, bajo este contexto también podremos construir algo equivalente al "Círculo de Correlaciones" introducido en la sección 1.5.2. Esta gráfica presenta la correlación entre las variables y los ejes obtenidos para la imagen euclidiana propuesta, por lo que siguiendo la metodología discutida previamente vemos que si la variable \underline{x} es centrada y reducida, este valor está dado por

$$\text{corr}(\underline{x}, \underline{\epsilon}_l) = \underline{x}' D \underline{\epsilon}_l,$$

y será la l -ésima coordenada del punto asociado a la variable \underline{x} en dicha imagen. Asimismo, podemos representar en la misma gráfica alguna variable exógena (centrada y reducida) \underline{y} que se hubiera medido sobre el mismo conjunto de individuos que participa en el análisis, su l -ésima coordenada está dada por

$$\text{corr}(\underline{y}, \underline{\epsilon}_l) = \underline{y}' D \underline{\epsilon}_l.$$

2.7 Trayectorias de los individuos.

El concepto de "trayectoria" de un individuo surge al buscar representar gráficamente el desarrollo o evolución de uno de los individuos involucrados en el análisis. Esto nos permitiría además de detectar aquellos sujetos que se distingan de los restantes, hacer un seguimiento de alguno en particular. La idea consiste en asociar un punto en la imagen euclidiana *compromiso* por cada individuo en cada uno de

los *estudios*, de tal manera que la l -ésima coordenada del punto asociado al i -ésimo sujeto en el j -ésimo *estudio* está dada por

$$\lambda_l^{-1/2} (w_j(\underline{\varepsilon}_i), \varepsilon_{i,l}) = \lambda_l^{-1/2} (W_j)'_i D\varepsilon_l,$$

donde ε_l es, como se mencionó anteriormente, el l -ésimo vector propio del operador w_d asociado al valor propio λ_l del conjunto ordenado en forma decreciente. Al correr $j = 1, \dots, k$ obtendremos los puntos asociados a este individuo al variar los *estudios*; si buscamos una representación bidimensional entonces $l = 1, 2$ y finalmente, al mover $i = 1, \dots, n$ obtenemos las trayectorias de cada uno de los individuos.

Todo esto se resumirá en la notación

M_i^j : punto asociado al i -ésimo individuo en el j -ésimo *estudio*.

M_i : representación *compromiso* del i -ésimo individuo.

Usando la definición de W , podemos deducir que M_i será el centro de gravedad de la trayectoria $\{M_i^j\}_{j=1}^k$ ponderada por los coeficientes α_j .

En lo que se refiere a las trayectorias de los individuos suplementarios se procederá de manera análoga, fijemos ideas con el primer sujeto adicional en el *estudio* j -ésimo, calculamos el vector n -dimensional $\underline{w}_j(\underline{\varepsilon}_{n+1})$ que contiene en cada entrada el producto escalar del individuo suplementario con cada uno de los activos en ese *estudio*, de tal manera que la l -ésima coordenada del j -ésimo punto de su trayectoria estará dado por

$$\lambda_l^{-1/2} (\underline{w}_j(\underline{\varepsilon}_{n+1}))' D\varepsilon_l,$$

con $\{\lambda_l\}$ y $\{\varepsilon_l\}$ como se definieron previamente. Cabe hacer notar que debido a que este tipo de individuos tienen asociado un peso nulo, no participaron en la construcción de W , por lo que no existirá su posición *compromiso*.

2.7.1 Información de las variables en las trayectorias de los individuos.

Cuando las variables involucradas en cada período de observación son altamente correlacionadas entre sí, el "Círculo de Correlaciones" presentará agrupación primero

por *estudio* que por variable, de tal suerte que no será fácil describir los ejes *compromiso* en función de ellas ni interpretar las trayectorias.

Por el contrario, cuando las variables están fuertemente correlacionadas entre los *estudios*, los ejes del *compromiso* tendrán una interpretación clara en función de ellas. Este es el caso en que los vectores propios de $W_k D$ son muy similares a los de $W D$, en particular el asociado al valor propio máximo. Así, la explicación del primer eje sería estable de un *estudio* a otro y las variables que lo explican serían las mismas, lo que proporcionaría un criterio general que permitiera asignar un nombre común a dicho eje, de tal manera que las diferencias en las coordenadas serían evidencia de diferencias en las variables asociadas a dicho eje.

2.8 Representación de las variables en *estudios*

(X_l, m, d_l) .

Esta es la situación que corresponde a la medición de un conjunto fijo de variables sobre diferentes grupos de individuos, de tal suerte que el espacio de variables (E_l, d_l) , es de dimensión n_l dependiendo del número de sujetos en cada período de observación y buscamos comparar las matrices de varianzas y covarianzas definidas en cada uno de ellos.

En la sección anterior definimos la matriz *compromiso* $V = \sum_{l=1}^k \alpha_l V_l \in \mathcal{L}(F^*, F)$ con $\sum_l \alpha_l = 1$, que puede considerarse como una matriz de varianzas medias entre variables, pues observemos que si partimos de *estudios* donde las variables se centran previamente, entonces

$$V_{ij} = \sum_{l=1}^k \alpha_l (V_l)_{ij} = \sum_{l=1}^k \alpha_l \left((\bar{x}_l)^i, (\bar{x}_l)^j \right)_{d_l},$$

que puede interpretarse como la covarianza *compromiso* entre las variables i, j .

Como puede deducirse fácilmente, ahora trabajaremos con la matriz V de la misma manera que lo hicimos con W . Se descompondrá en valor singular a través de los valores y vectores propios de VM y obtendremos una imagen euclidiana O, N_1, \dots, N_m para las variables. El producto escalar $(\overline{ON}_i, \overline{ON}_j)$ denotará la covarianza *compromiso*.

por *estudio* que por variable, de tal suerte que no será fácil describir los ejes *compromiso* en función de ellas ni interpretar las trayectorias.

Por el contrario, cuando las variables están fuertemente correlacionadas entre los *estudios*, los ejes del *compromiso* tendrán una interpretación clara en función de ellas. Este es el caso en que los vectores propios de $W_k D$ son muy similares a los de $W D$, en particular el asociado al valor propio máximo. Así, la explicación del primer eje sería estable de un *estudio* a otro y las variables que lo explican serían las mismas, lo que proporcionaría un criterio general que permitiera asignar un nombre común a dicho eje, de tal manera que las diferencias en las coordenadas serían evidencia de diferencias en las variables asociadas a dicho eje.

2.8 Representación de las variables en *estudios*

(X_t, m, d_t) .

Esta es la situación que corresponde a la medición de un conjunto fijo de variables sobre diferentes grupos de individuos, de tal suerte que el espacio de variables (E_t, d_t) , es de dimensión n_t dependiendo del número de sujetos en cada período de observación y buscamos comparar las matrices de varianzas y covarianzas definidas en cada uno de ellos.

En la sección anterior definimos la matriz *compromiso* $V = \sum_{l=1}^k \alpha_l V_l \in \mathcal{L}(F^*, F)$ con $\sum_l \alpha_l = 1$, que puede considerarse como una matriz de varianzas medias entre variables, pues observemos que si partimos de *estudios* donde las variables se centran previamente, entonces

$$V_{ij} = \sum_{l=1}^k \alpha_l (V_l)_{ij} = \sum_{l=1}^k \alpha_l \left((\underline{x}_l)^i, (\underline{x}_l)^j \right)_{d_t},$$

que puede interpretarse como la covarianza *compromiso* entre las variables i, j .

Como puede deducirse fácilmente, ahora trabajaremos con la matriz V de la misma manera que lo hicimos con W . Se descompondrá en valor singular a través de los valores y vectores propios de VM y obtendremos una imagen euclidiana O, N_1, \dots, N_m para las variables. El producto escalar $(\overline{ON}_i, \overline{ON}_j)$ denotará la covarianza *compromiso*

miso entre las variables i y j ; así como el coseno del ángulo que forman representará su correlación *compromiso*.

Para construir trayectorias asociadas a cada una de las variables, consideremos el conjunto m -normalizado de vectores propios de VM : $\{\underline{\phi}_1, \dots, \underline{\phi}_m\}$ asociado a los valores propios no nulos $\lambda_1 \geq \dots \geq \lambda_p$. La l -ésima coordenada del punto N_j que representa a la variable j en la matriz *compromiso* está dada por

$$\lambda_l^{-1/2} (V(\underline{f}_j^*), \underline{\phi}_l)_m = \lambda_l^{-1/2} (\underline{V}^j)' M \underline{\phi}_l,$$

y la l -ésima coordenada del punto N_i^j que representa a la variable j en el estudio i estará dada por:

$$\lambda_l^{-1/2} (V_i(\underline{f}_j^*), \underline{\phi}_l)_m = \lambda_l^{-1/2} (V_i)^j M \underline{\phi}_l,$$

lo que nos permitirá construir una trayectoria para cada variable y así detectar cuáles son las responsables de diferencias importantes entre matrices V_i, V_j .

Capítulo 3

Aplicación del Método Statis a la dinámica demográfica en México (1930-1990).

3.1 Introducción.

Este tercer capítulo busca ilustrar el desarrollo, el tipo de resultados e interpretaciones que se pueden obtener con el *Método Statis* cuando se aplica a un problema práctico. Dada la gran cantidad de información numérica que se puede resumir, así como la complejidad de los cálculos, decidimos implementar una rutina en el lenguaje *S-Plus* que llevara a cabo este trabajo; la razón de haber escogido precisamente este lenguaje sobre la gran cantidad de opciones de que se disponen actualmente, fue pensando en su manejo interno de matrices y la forma sencilla de operarlas, además de estar ampliamente difundido en el medio estadístico. El código de las rutinas y todas las corridas obtenidas se anexan en el apéndice final de la tesis.

3.2 El Problema.

Contamos con la información de algunos aspectos demográficos de la República Mexicana entre los años de 1930 a 1990 y se busca, a partir de ellos, describir la

dinámica poblacional que ha sufrido el país en este período, tanto a nivel nacional como estatal. Esta información se obtuvo de los censos nacionales publicados por el Instituto Nacional Estadística, Geografía e Informática (INEGI) [10], de los cuales decidimos resumir el problema a las siguientes variables:

- número de nacimientos de hombres,
- número de nacimientos de mujeres,
- número de defunciones de niños menores de un año,
- número de defunciones generales,
- matrimonios,
- divorcios.

Es importante señalar que nuestro objetivo consiste en describir el comportamiento de una *población*, y la información disponible también es de carácter poblacional, de tal manera que estamos ante un caso en que el interés no es hacer inferencia sino *describir* lo que presentan los censos; contamos con los valores de estas 6 variables para los 32 estados de la República durante 7 censos: 1930, 1940, 1950, 1960, 1970, 1980 y 1990, que se resumieron en 7 matrices de 32 renglones por 6 columnas.

3.3 Selección de métricas y coeficientes para el uso del *Método Statis*.

Traduciendo el problema que tratamos de resolver en el lenguaje manejado en los dos primeros capítulos de la tesis, podemos notar que contamos con 7 *estudios*, en los cuales tanto el espacio vectorial definido por individuos (estados de la República), como aquél definido por las variables, permanecieron fijos en el período de interés. Esto nos permite entonces, analizar tanto la evolución demográfica de cada uno de los estados a través del tiempo, así como el comportamiento a nivel nacional de las variables que estamos estudiando.

Decidimos realizar todo el análisis utilizando matrices previamente centradas y estandarizadas, con el objeto de eliminar desde un principio cualquier problema debido a las unidades de medición y magnitud de las variables.

Tomando en cuenta que para fines de nuestro estudio, los 32 estados de la República tienen igual importancia, consideramos que lo más adecuado era asignarle al espacio vectorial de variables E , el producto escalar estándar d_p definido por

$$(\xi_i, \xi_j)_{d_p} = \begin{cases} \frac{1}{32} & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

cuya matriz asociada es $D_p = (1/32) * I_{32 \times 32}$. Asimismo, el espacio de individuos F fue provisto del correspondiente producto escalar estándar, el cual le asigna igual importancia relativa a cada una de las variables

$$(\underline{f}_i, \underline{f}_j)_{m_p} = \begin{cases} \frac{1}{6} & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

con matriz asociada $M_p = (1/6) * I_{6 \times 6}$. Así, la estructura con la cual trabajaremos consiste de 7 *estudios* de la forma $(X_k, (1/6) * I_{6 \times 6}, (1/32) * I_{32 \times 32})$, ($k = 1, \dots, 7$), en los cuales la matriz X_k es una matriz de datos centrada y reducida. Los *estudios* serán todos del mismo peso, $1/7$, definiéndose así, de manera análoga a los casos anteriores, el producto escalar en el espacio vectorial formado por los *estudios*.

Recordemos que este estudio tiene dos objetivos: primero, describir la evolución de cada uno de los estados de la República para resaltar los cambios demográficos sufridos a través del tiempo y segundo, describir el comportamiento de las variables durante ese período, lo que nos sugiere comparar tanto las matrices de distancias entre individuos W_k como aquéllas de distancias entre variables V_k , siendo éstas últimas en realidad, matrices de correlación puesto que los datos fueron reducidos previamente.

3.4 Evolución demográfica de los estados de la República.

Para estudiar el comportamiento de los Estados, primero se procedió a encontrar las matrices de productos escalares entre individuos W_k en cada uno de los censos, para

posteriormente "compararlas" construyendo la matriz de productos escalares Hilbert-Schmidt entre estos objetos. Una vez obtenida esta matriz, la cual llamaremos S atendiendo a la notación del capítulo anterior, se descompuso en valor singular con el objeto de construir una imagen euclidiana de los Estados que nos permitiera detectar la posible existencia de una estructura común a lo largo del período de observación, lo que dió lugar a la Gráfica No. 1: "Dinámica con respecto a los estados".

Esta gráfica pone de manifiesto el comportamiento temporal del fenómeno bajo estudio, ya que los puntos asociados a cada uno de los censos siguen precisamente el orden cronológico que les dio origen. Las normas de los vectores asociados a los puntos W_k , son del mismo orden de magnitud, lo que sugiere la existencia de una posible estructura común al interior de cada tabla, que evoluciona con el paso del tiempo, sin embargo, los ángulos que forman entre sí los vectores asociados a cada uno de los censos, detectan cambios bruscos en la dinámica demográfica en los censos de 1950 y 1990, puesto que los ángulos que definen con los censos precedentes respectivos son ligeramente más grandes que los restantes, de tal manera que se tienen indicios de comportamientos diferentes en las décadas de los 40's y 80's que resumen estos censos. El porcentaje de inercia que explica esta representación gráfica resulta ser del 87.53%, lo cual nos habla de una muy alta calidad de representación global del fenómeno.

3.4.1 Posición *compromiso* de los Estados.

Hasta ahora, hemos detectado la existencia de una estructura evolutiva común de los Estados, pero ¿cómo puede resumirse o presentarse esta estructura?. Recordemos que en la propuesta del *Método Statis*, una vez detectada la presencia de *interestructura*, se sugiere la construcción de una matriz *compromiso* W , que nos permita entenderla. Esta matriz, que no es más que una media ponderada de las matrices W_k , goza de la importante propiedad de ser aquella que maximiza la correlación con las matrices que la originaron, de tal suerte que puede interpretarse como el mejor "resumen" de lo acontecido a nivel individual durante el período de estudio. Por lo anterior, tendremos que construir ahora una *imagen euclidiana compromiso* de los estados partiendo de la matriz W . El resultado de esta construcción se presenta en

la Gráfica No. 2: "Representación compromiso de los estados".

Para poder interpretar adecuadamente esta gráfica, es preciso entender perfectamente bien lo que representan cada uno de los ejes *compromiso* en términos de las variables que estamos estudiando. Para esto calcularemos sus correlaciones con los ejes en cada uno de los censos. Los resultados se presentan a continuación y se plasmaron en las Gráficas No. 3, 4, 5, 6, 7, 8, y 9: "Correlaciones con los ejes compromiso: " 1930, 1940, 1950, 1960, 1970, 1980 y 1990 respectivamente:

		<i>Con eje 1</i>	<i>Con eje 2</i>	<i>Con eje 1</i>	<i>Con eje 2</i>
		1930	1930	1940	1940
<i>var 1</i>	Nac. de Hombres	0.8471	-0.2338	0.8438	-0.1682
<i>var 2</i>	Nac. de Mujeres	0.8444	-0.2352	0.8405	-0.1695
<i>var 3</i>	Def. Generales	0.8390	-0.2568	0.8538	-0.2030
<i>var 4</i>	Def. de Niños	0.8427	-0.2221	0.8687	-0.1399
<i>var 5</i>	Matrimonios	0.8659	-0.0208	0.7806	-0.0292
<i>var 6</i>	Divorcios	0.2572	0.5449	0.3230	0.7679

	<i>Con eje 1</i>	<i>Con eje 2</i>	<i>Con eje 1</i>	<i>Con eje 2</i>	<i>Con eje 1</i>	<i>Con eje 2</i>
	1950	1950	1960	1960	1970	1970
<i>var 1</i>	0.8948	-0.0582	0.9061	0.05985	0.9072	0.0576
<i>var 2</i>	0.8999	-0.0552	0.9033	0.0685	0.9032	0.0772
<i>var 3</i>	0.8958	-0.1223	0.9034	-0.0483	0.9104	-0.0235
<i>var 4</i>	0.9056	-0.0332	0.8964	0.0491	0.8951	0.0163
<i>var 5</i>	0.8682	0.1493	0.8510	0.2007	0.8799	0.1692
<i>var 6</i>	0.2033	0.8409	0.1344	0.8045	0.0852	0.7768

		<i>Con eje 1</i>	<i>Con eje 2</i>	<i>Con eje 1</i>	<i>Con eje 2</i>
		1980	1980	1990	1990
<i>var 1</i>	Nac. de Hombres	0.9039	-0.0120	0.8364	-0.1181
<i>var 2</i>	Nac. de Mujeres	0.9042	-0.0167	0.8240	-0.1240
<i>var 3</i>	Def. Generales	0.9013	-0.0634	0.8849	-0.0518
<i>var 4</i>	Def. de Niños	0.8556	-0.1247	0.7581	-0.1796
<i>var 5</i>	Matrimonios	0.8987	0.0816	0.8623	-0.0382
<i>var 6</i>	Divorcios	0.7427	0.4209	0.7368	0.4734

Esto nos permite ver que las primeras cinco variables: nacimientos de ambos sexos, defunciones y matrimonios son las que se encuentran representadas primordialmente en el primer eje, mientras que la variable divorcios es aquella que define principalmente el segundo. Regresando entonces a la Gráfica No. 2, podemos notar que existen dos estados que difieren claramente del grupo restante, el número 8, Chihuahua y el 9, que representa al Distrito Federal. Chihuahua tiene una abscisa semejante a los demás estados siendo la ordenada la que lo distingue de ellos; esto nos indica que el comportamiento demográfico de Chihuahua es similar al del resto de la República excepto en lo referente al número de divorcios; por otro lado, el Distrito Federal se distingue en ambas coordenadas, lo cual nos da evidencia de un dinámica totalmente diferente de esta entidad, tanto en lo que se refiere a nacimientos y defunciones, así como en matrimonios y divorcios.

3.4.2 Trayectorias de cada Estado.

Tomando en cuenta que en la imagen euclidiana *compromiso*, la evolución de cada estado a lo largo del período de observación fue reducida a un único "punto promedio", nos preguntamos ahora, cómo hubieran sido sus trayectorias si se hubiera sintetizado a un punto pero ahora por cada uno de los censos trabajados; esto nos permitiría conocer a detalle la forma cómo se construyó ese "punto promedio" y así detectar momentos importantes de cada entidad, tales como aquéllos en los cuales hubo cambios de tendencia en la dinámica, o los períodos de estabilidad, etc. Todo

lo anterior entonces, justifica la construcción de trayectorias individuales en el *plano compromiso*, que podrán interpretarse en referencia a un "estado promedio" el cual sería una entidad ficticia con la propiedad de tener valores medios de cada variable en todos los censos, de tal suerte que la trayectoria correspondiente se reduciría a un punto que será precisamente el origen del *plano compromiso* donde se representarán dichas trayectorias.

Así llegamos a la Gráfica No. 10: "Trayectorias de los estados: 1930-1990", la cual nos presenta las 32 trayectorias asociadas a cada uno de los estados simultáneamente en el *plano compromiso*, por lo que la interpretación de los ejes es exactamente la misma que en la Gráfica No. 2. Es importante señalar que esta representación tiene el inconveniente de incluir, como se dijo antes, a los 32 estados simultáneamente, dificultando notablemente su lectura, no obstante se consideró importante incluirla, porque nos da un panorama general de ellas. Aquí es donde podemos detectar la presencia de tres estados cuya inercia demográfica ha sido más accidentada que la de los demás: Chihuahua, Edo. de México y Distrito Federal

Confirmando la lectura de la *imagen compromiso*, Chihuahua tiene una abscisa muy similar a la de los estados restantes, la cual se mantiene a lo largo de toda la trayectoria, es decir, en lo referente a nacimientos y defunciones podría considerarse un estado típico, sin embargo en la década de los 40's se presentó un fuerte movimiento en cuanto a divorcios se refiere, fenómeno que se repitió en los 70's pero en sentido inverso, puesto que se tendió a regresar a la posición inicial.

En lo que se refiere al Distrito Federal, la lectura de su trayectoria nos aporta información muy valiosa adicional a la que ya teníamos en su representación *compromiso*, pues nos dice que su diferencia principal con respecto al resto de la República, radica en el número de nacimientos y defunciones; en lo que a divorcios se refiere, vemos que, como se detectó en la Gráfica No. 2, existen diferencias en ordenadas pero no tan llamativas como aquéllas de las abscisas.

El Estado de México, no había sido detectado antes, lo cual se debe a que, en promedio, tuvo un comportamiento similar al de los estados restantes, no obstante al analizar su trayectoria, notamos que sufrió cambios muy bruscos en su inercia demográfica, sobre todo en lo referente a nacimientos y defunciones, lo que podría

ser producto de su cercanía al Distrito Federal y el continuo crecimiento de la ciudad que se ha extendido hasta este estado. Cabe señalar que las décadas de los años 70's y 80's son las que presentan mayor desviación con respecto a la nube formada por las entidades restantes.

Como pudimos darnos cuenta, la lectura de esta última gráfica fue bastante complicada por representar simultáneamente todos los estados, dificultándose aun más, si hubiéramos necesitado distinguir alguna trayectoria en particular. Esto nos sugirió reducir esta representación a grupos de ocho estados, lo cual tiene la desventaja de quitarnos el panorama global que teníamos, pero a cambio nos permite conocer a detalle cada una de las curvas expuestas. Ver Gráficas 11, 12, 13, 14 y 14': "Trayectorias de los estados I, II, III, IV y IV'" respectivamente.

Utilizando estos acercamientos podemos hacer algunas observaciones adicionales. En el grupo I, además de Chihuahua, ya comentado ampliamente, los estados de Baja California y Chiapas sufrieron cambios bruscos en cuanto a nacimientos y defunciones se refiere, no así en lo concerniente a divorcios; Aguascalientes, Baja California Sur, Campeche y Colima, presentan tendencias similares y de pocos cambios.

Analizando el segundo grupo, vemos que junto con el Estado de México, discutido previamente, Jalisco, Guanajuato y Michoacán también presentan cambios fuertes en la dinámica referente a divorcios, no así en cuanto a nacimientos y defunciones. Este mismo fenómeno se presenta en Puebla, Veracruz y Oaxaca (ver Gráficas 13, 14 y 14') y en sentido inverso en Morelos, Sonora y Tlaxcala.

Puesto que tanto el conjunto de individuos (estados) como el de variables permanecieron fijos a lo largo del período de estudio, podemos realizar un análisis similar al que hicimos sobre los estados, pero ahora para las variables.

3.5 Evolución de las variables.

Repetiendo el procedimiento que describimos anteriormente, iniciamos este enfoque calculando las matrices de distancias entre variables en cada uno de los censos, que en realidad serán de correlación por haberse estandarizado desde un principio.

Descriptivamente hablando, podemos notar que la variable número de divorcios,

es aquélla que se encuentra menos correlacionada con las restantes, no obstante podríamos decir en términos genéricos, que va aumentando con el paso del tiempo, llegando a ser muy alta con el número de matrimonios.

Con el objeto de detectar una posible estructura común al interior de estas matrices, se calculó la matriz de productos escalares Hilbert-Schmidt entre objetos V_i , que nos permitió construir la imagen euclidiana de los censos con respecto a las variables, presentada en la Gráfica No. 15: "Dinámica con respecto a las variables". Aquí podemos observar cómo los censos de 1980 y 1990 son muy semejantes entre sí, pero notablemente distintos a los restantes; a diferencia del análisis anterior, la evidencia de interestructura resulta ser vaga o quizá fragmentada, antes de 1970 y después de 1980, lo que puede interpretarse como señal de un cambio demográfico muy drástico en la década de los 70's resumido en el censo de 1980.

A pesar de haberse detectado esta falta de estructura, decidimos construir la matriz *compromiso* V que resumiera lo acontecido en las variables estudiadas. Con ella generamos la imagen euclidiana de las variables donde se resalta lo descubierto en el primer análisis; el comportamiento radicalmente diferente de los divorcios evidenciando también, una ligera separación del número de matrimonios con respecto a las demás variables. Esto puede consultarse en la Gráfica No. 15: "Representación compromiso de variables demográficas". También se pone de manifiesto, como era de esperarse, el comportamiento prácticamente idéntico del nacimiento de hombres y mujeres.

En este segundo análisis, cobra mucha mayor importancia la construcción de trayectorias asociadas a cada una de las características estudiadas para tratar de explicarnos el cambio detectado en 1980. El resultado de esta construcción se presenta en las gráficas No. 17 y 17': "Trayectorias de las variables" las cuales muestran el gran salto que dieron todas las variables entre el censo de 1970 y 1980 tendiendo a ocupar una posición relativamente semejante en estos dos últimos períodos. Aunque claramente el mayor brinco lo dieron los divorcios, también podemos notar que nacimientos de hombres, de mujeres, defunciones generales y matrimonios tuvieron una alteración muy semejante pero ligeramente diferente a la de defunciones de niños, que durante los censos iniciales era prácticamente indistinguible de ellas,

pero también hacia la década de los 70's tiende a separarse.

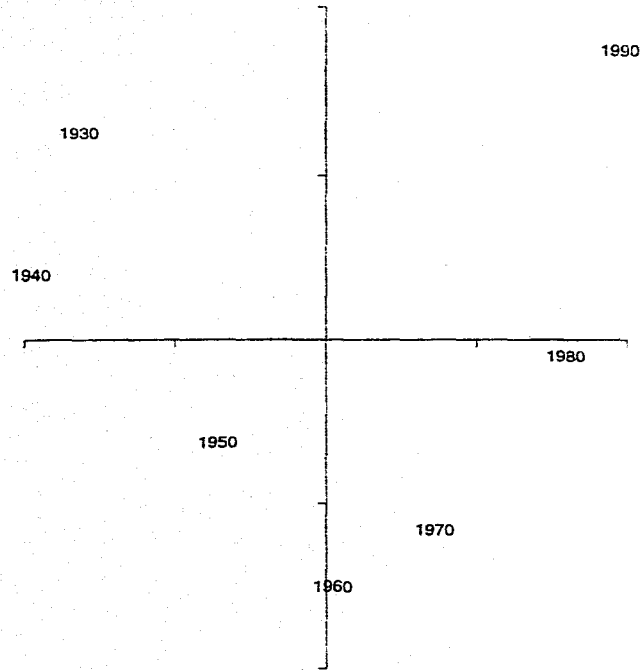
3.6 Conclusiones.

Esta aplicación nos permitió observar con detalle la propuesta del *Método Statis* para reducir una base muy grande de información y cuyo análisis era de entrada, bastante complejo. El conjunto inicial de datos era verdaderamente enorme y, ayudados de algunas rutinas de cómputo, pudimos describir y entender el comportamiento del fenómeno resumido en unas cuantas gráficas.

Claramente este método es de carácter descriptivo, razón por la cual no se impusieron supuestos distribucionales de ninguna especie para poderse aplicar de manera adecuada; se enfocó fundamentalmente a "leer" lo que los datos dictaron. Algo interesante de este ejemplo, fue que al estudiar el conjunto de estados, la existencia de interestructura fue sumamente clara, permitiéndonos realizar un análisis bastante detallado a nivel estatal. En cambio, esto no ocurrió cuando tratamos de hacer el análisis correspondiente para las variables, donde las evidencias de esta estructura fueron muy vagas; no obstante pudimos detectar desde un enfoque descriptivo, el punto de cambio de tendencia donde se presentó esta fractura, lo cual podría tener interpretaciones muy interesantes desde el punto de vista demográfico, al contextualizarse en las condiciones generales que rodearon al país en ese momento.

Gráficas

Dinamica con respecto a los estados
87.53% de la inercia explicada

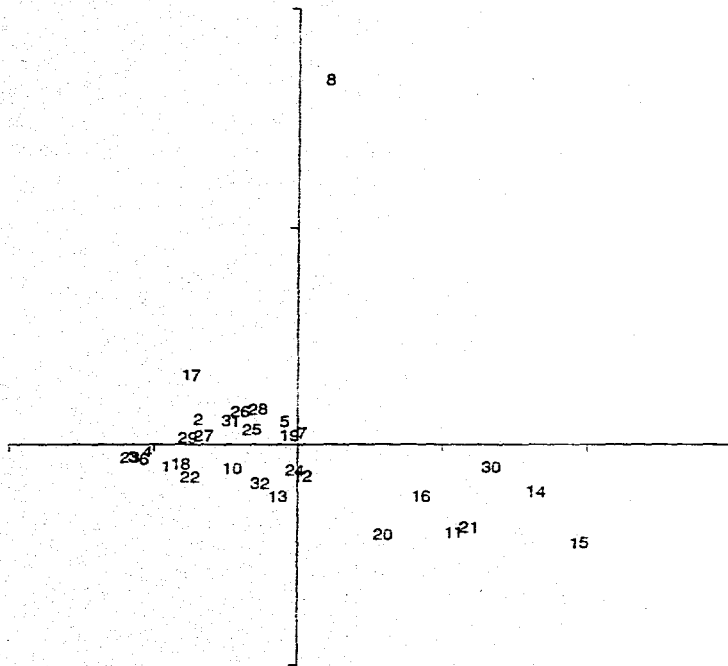


Grafica No. 1

Representación compromiso de los estados
87.96% de la inercia explicada

- 1 AGS
- 2 BC
- 3 BCS
- 4 CAMP
- 5 COAH
- 6 COL
- 7 CHIS
- 8 CHIH
- 9 DF
- 10 DGO
- 11 GTO
- 12 GRC
- 13 HGO
- 14 JAL
- 15 MEX
- 16 MICH

- 17 MOR
- 18 NAY
- 19 NL
- 20 OAX
- 21 PUE
- 22 QRO
- 23 QR
- 24 SLP
- 25 SIN
- 26 SON
- 27 TAB
- 28 TAM
- 29 TLAX
- 30 VER
- 31 YUC
- 32 ZAC

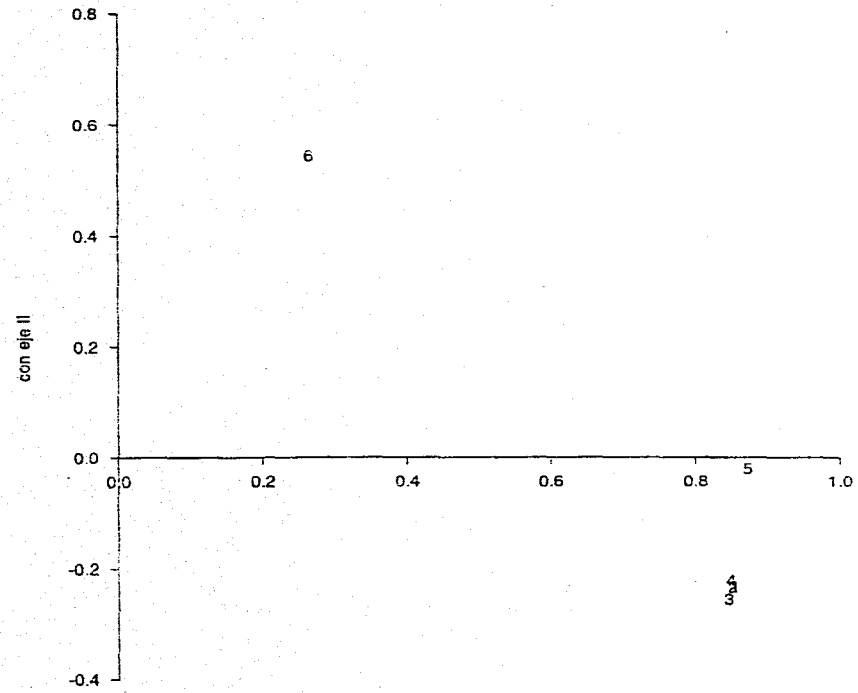


9

79

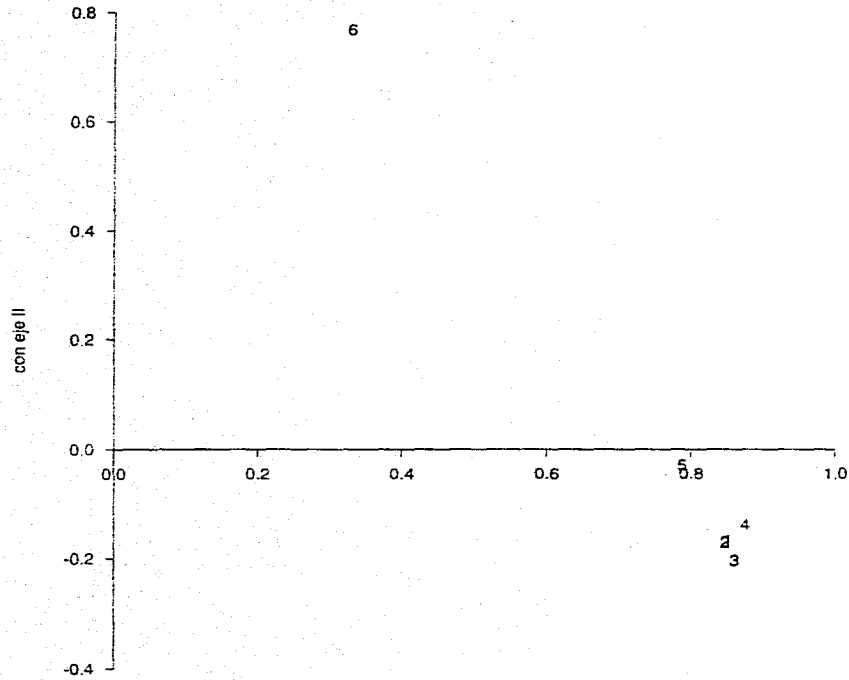
Grafica No. 2

Correlaciones con los ejes compromiso: 1930



con eje I
Grafica No. 3

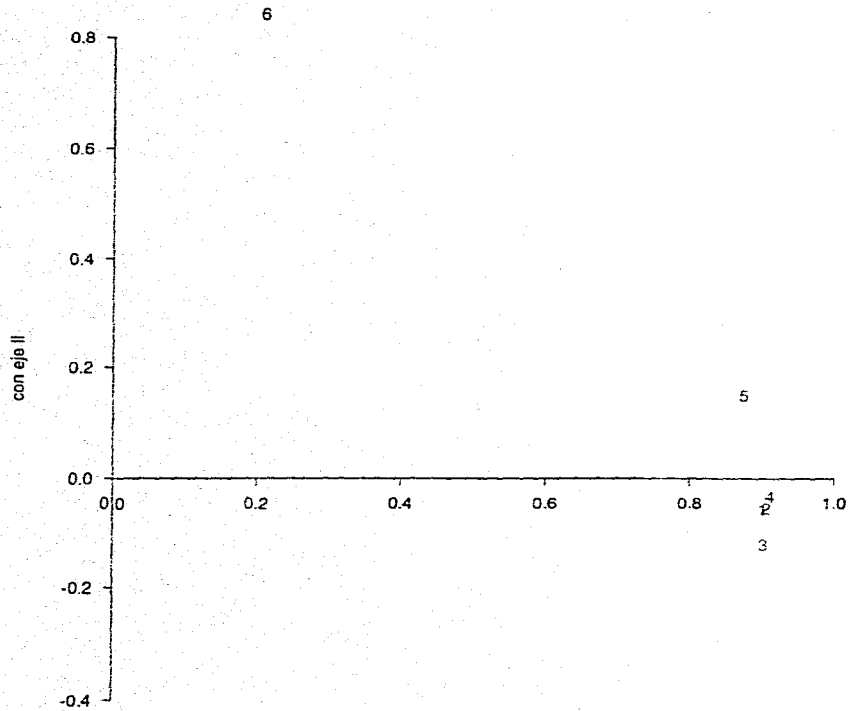
Correlaciones con los ejes compromiso: 1940



- 1 nac. hombres
- 2 nac. mujeres
- 3 def. generales
- 4 def. niños
- 5 matrimonios
- 6 divorcios

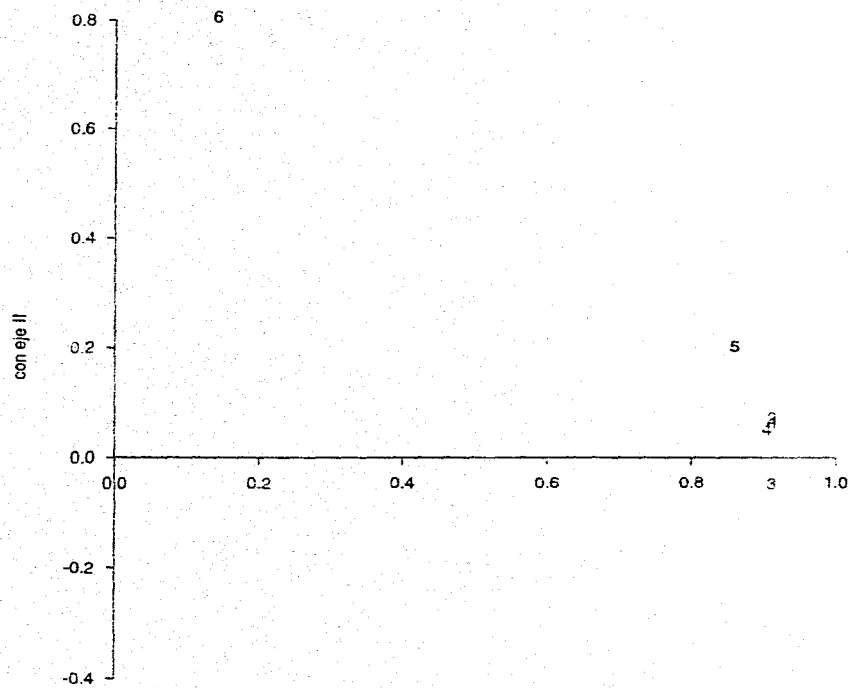
con eje I
Grafica No. 4

Correlaciones con los ejes compromiso: 1950



con eje I
Grafica No. 5

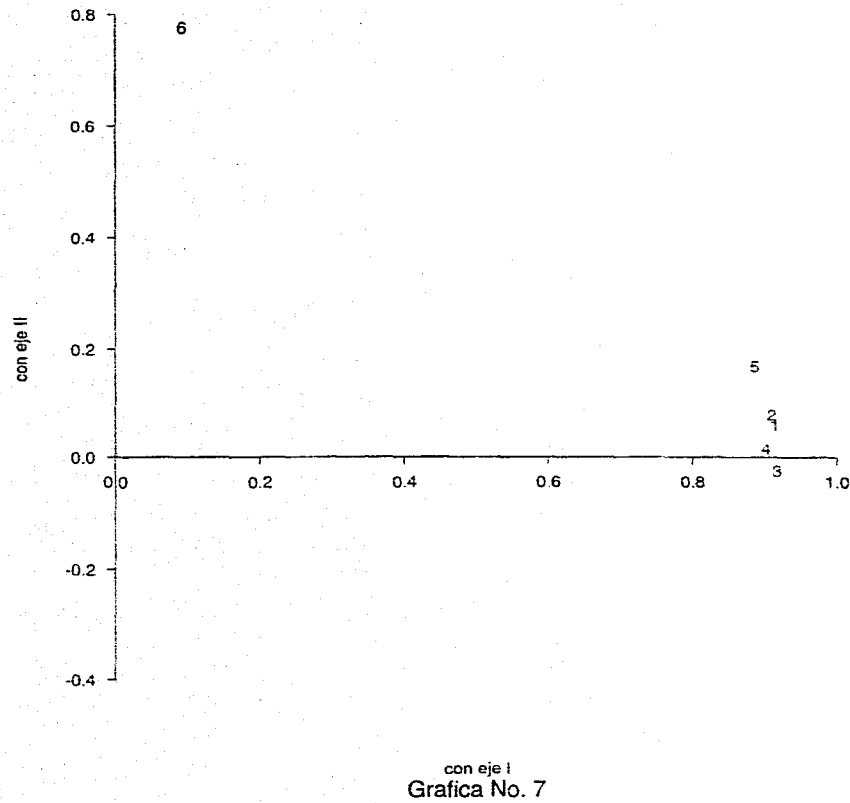
Correlaciones con los ejes compromiso: 1960



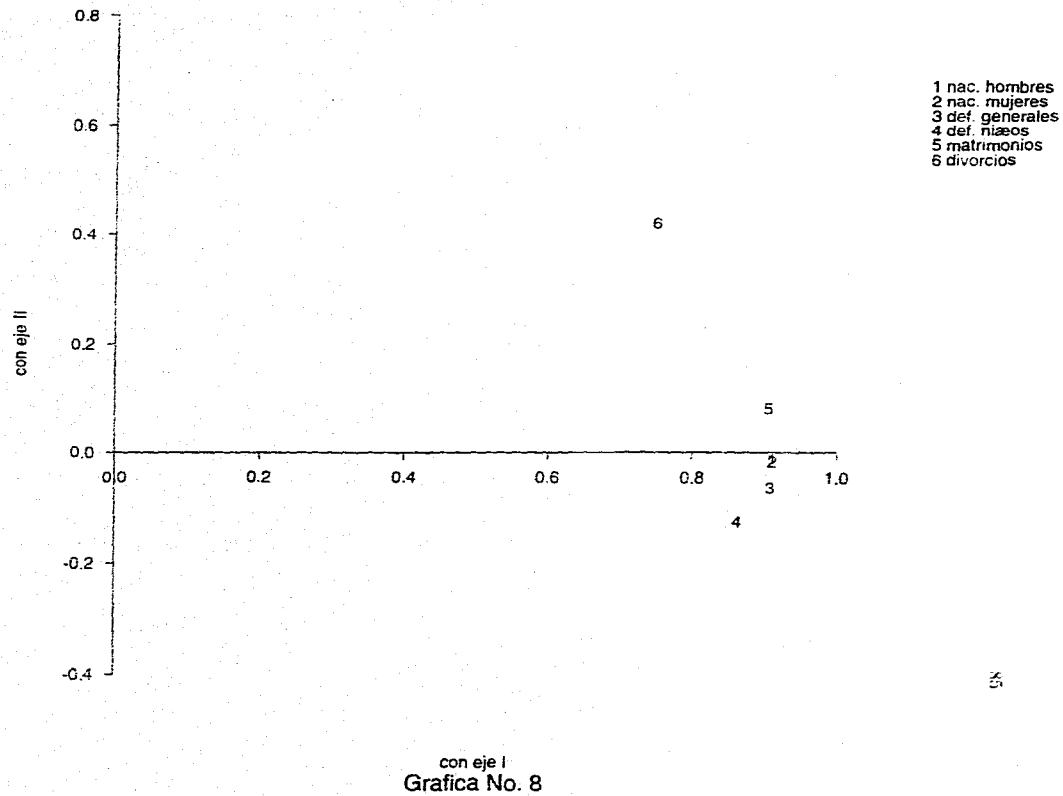
- 1 nac. hombres
- 2 nac. mujeres
- 3 def. generales
- 4 def. niños
- 5 matrimonios
- 6 divorcios

con eje I
Grafica No. 6

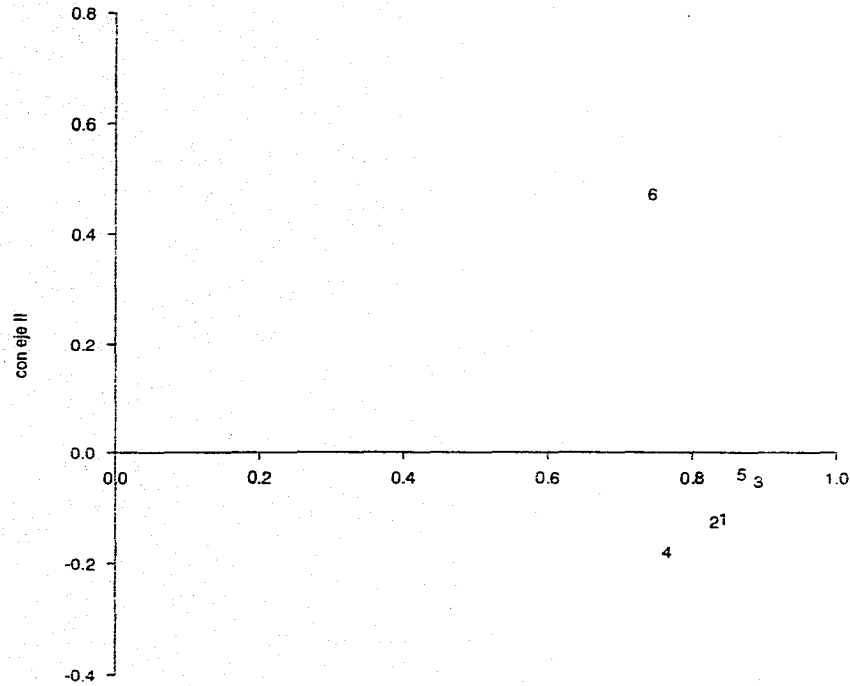
Correlaciones con los ejes compromiso: 1970



Correlaciones con los ejes compromiso: 1980



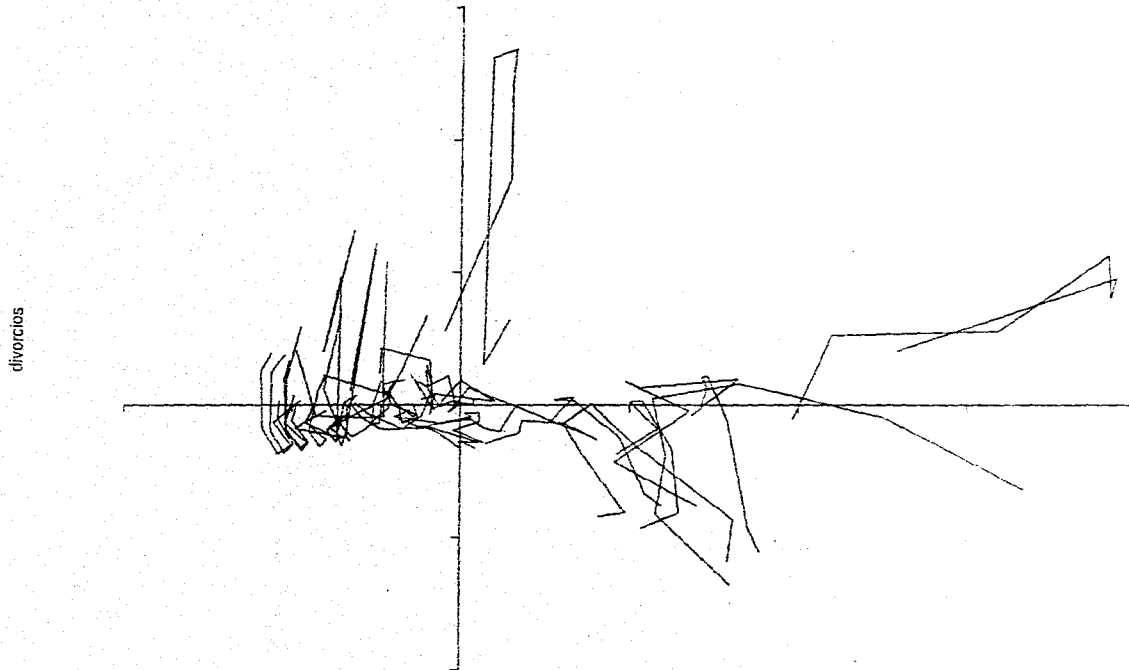
Correlaciones con los ejes compromiso: 1990



- 1 nac. hombres
- 2 nac. mujeres
- 3 def. generales
- 4 def. niños
- 5 matrimonios
- 6 divorcios

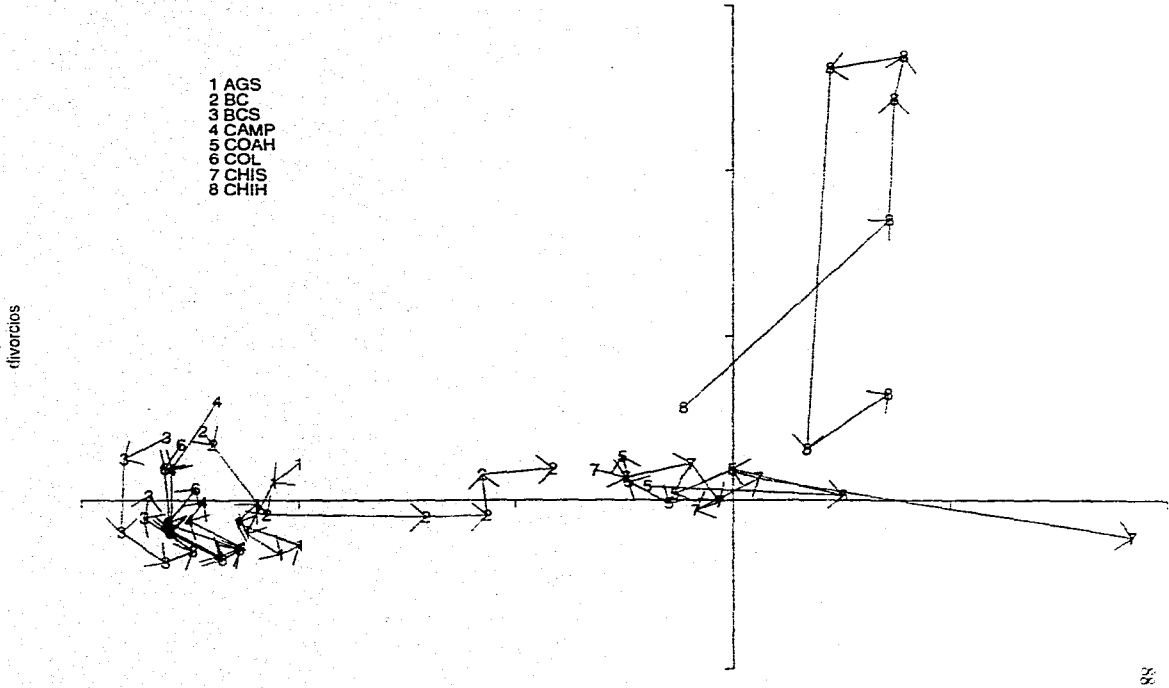
con eje I
Grafica No. 9

Trayectorias de estados: 1930-1990



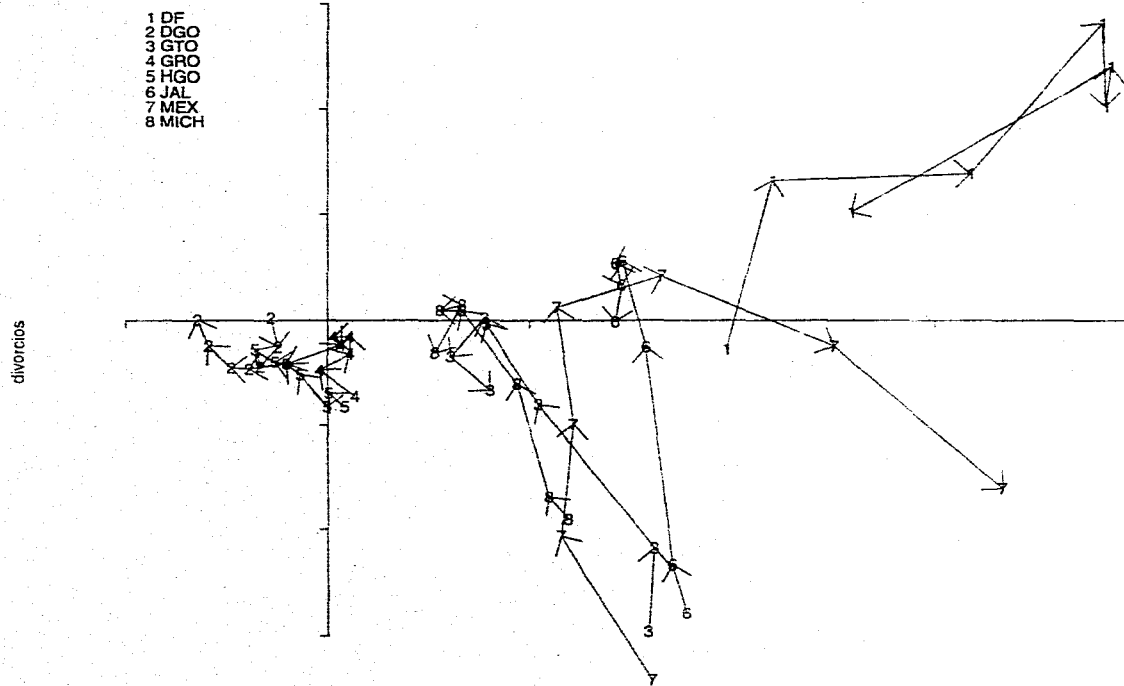
nacimientos, defunciones y matrimonios
Grafica No. 10

Trayectorias de los estados I



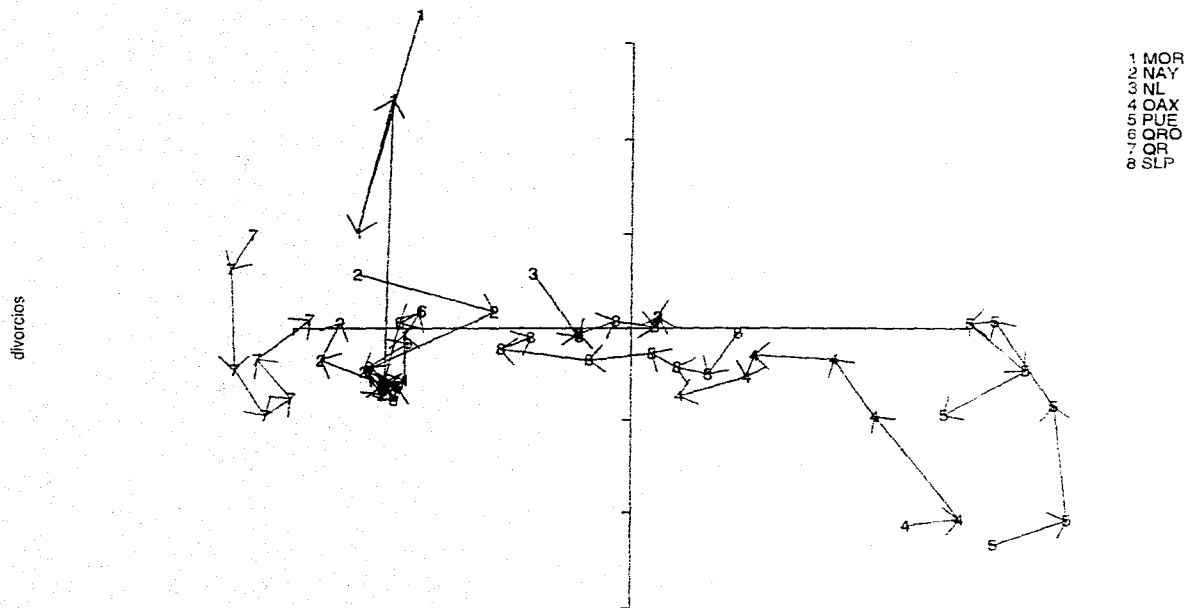
nacimientos, defunciones y matrimonios
Grafica No. 11.

Trayectorias de los estados II



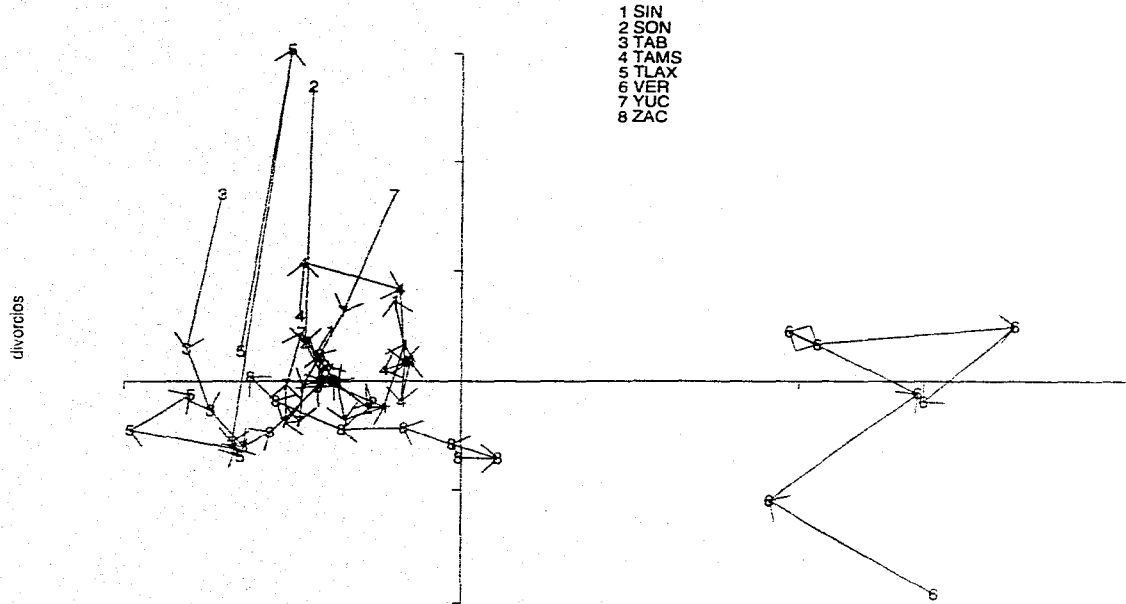
nacimientos, defunciones y matrimonios
Grafica No. 12

Trayectorias de los estados III



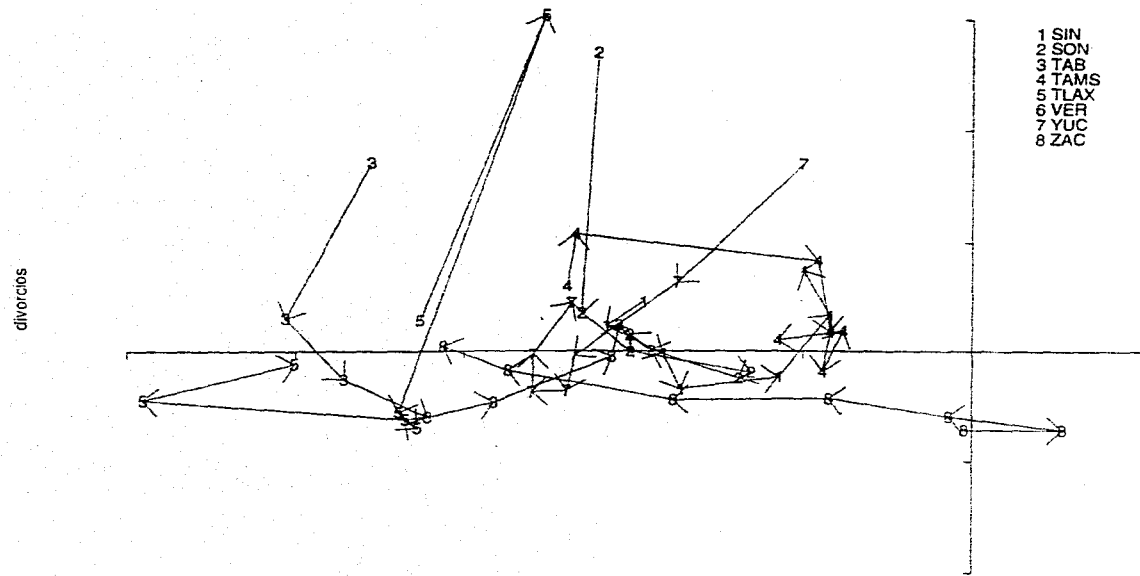
nacimientos, defunciones y matrimonios
Grafica No. 13

Trayectorias de los estados IV



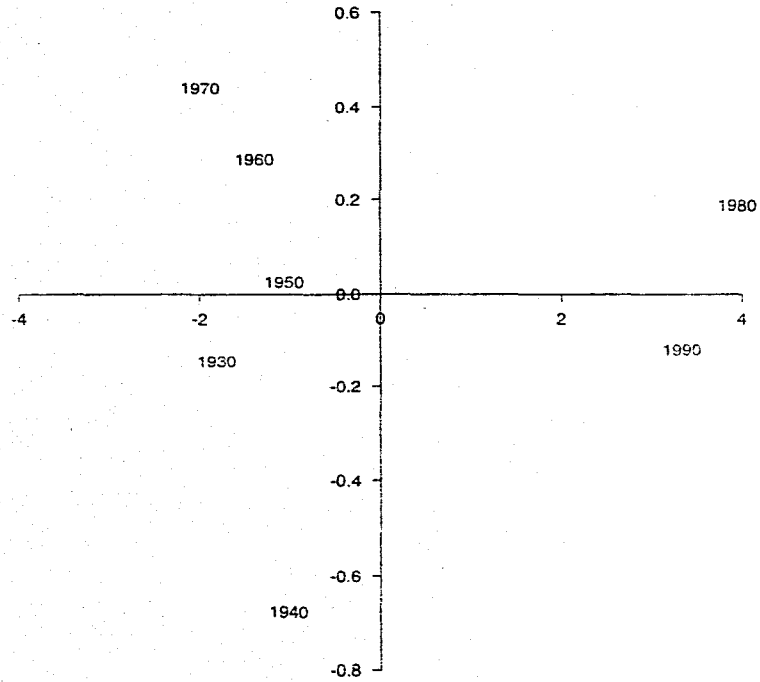
nacimientos, defunciones y matrimonios
Grafica No. 14

Trayectorias de los estados IV. Acercamiento



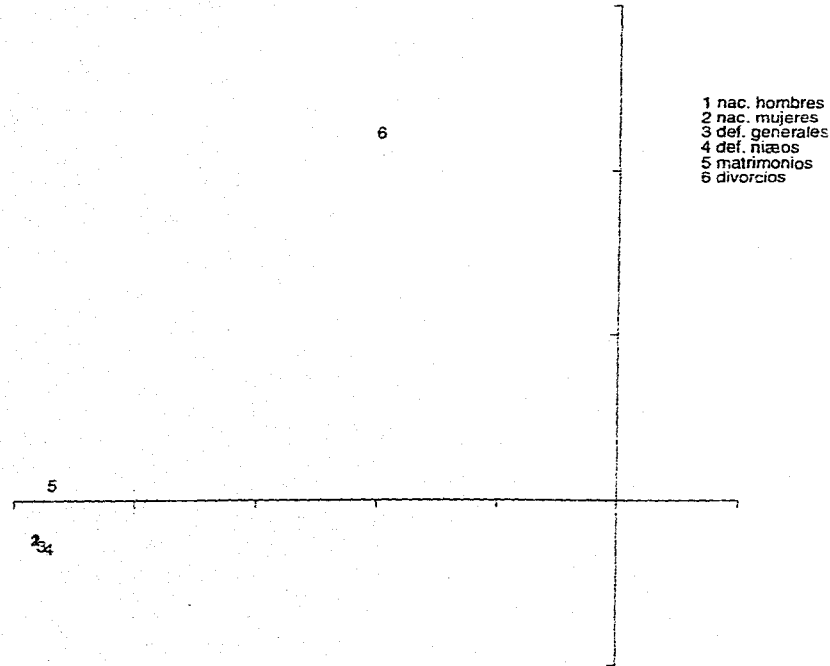
nacimientos, defunciones y matrimonios
Grafica No. 14'

Dinamica con respecto a las variables
99.14% de la inercia explicada



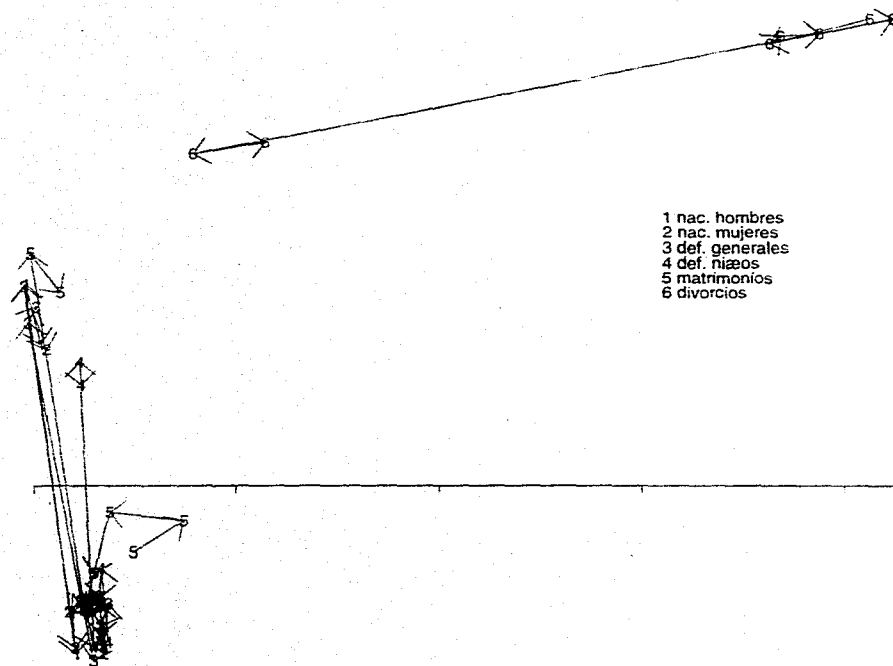
Grafica No. 15

Representacion compromiso de variables demograficas



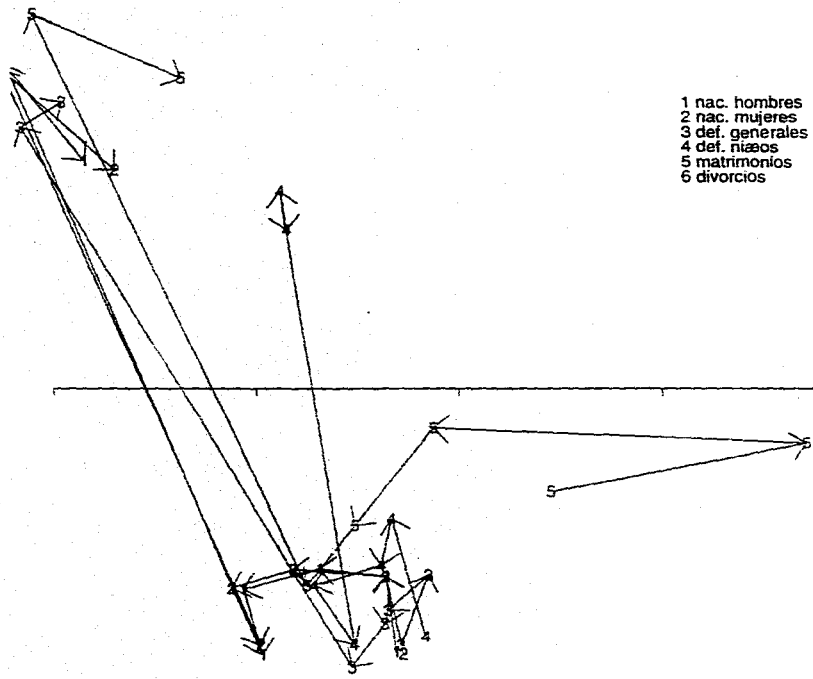
Grafica No. 16

Trayectorias de las variables



Grafica No. 17

Trayectorias de las variables. Acercamiento



Grafica No. 17'

Rutinas en S-Plus

Rutinas en S-Plus

Este programa implementa el *Método Stats* en este lenguaje, para el problema concreto que se trabaja en esta tesis, en caso de que se vaya a utilizar sobre algún otro problema de interés para el lector, deberán modificarse situaciones particulares como la ruta en disco donde se tienen grabadas las matrices a analizar, los nombres de los archivos correspondientes y el número de ellas, así como el vector de nombres de individuos, *estudios* y variables.

Los renglones que inician con "#" son aquéllos que contienen comentarios que indican cuál es la tarea de cada parte del programa y donde se podrán identificar los segmentos a cambiar para otro problema particular.

El programa se dividió en cuatro rutinas únicamente por razones de eficiencia y deberán de ser corridas en ese mismo orden. Las rutinas 1 y 2 atacan el problema en el que el conjunto de individuos observados permanece constante durante el período de observación y el que pudo haber sido modificado fue el de variables; las rutinas 3 y 4 analizan la situación inversa.

RUTINA No. 1

Definición de algunas funciones útiles

traza:

```
trace.x_function(x){
  trace.x_0
  end_sqrt(length(x))
  for(i in 1:end) trace.x_trace.x+x[i,i]
}
```

Definición del número de matrices

k_7

Lectura de tablas de datos

```
x1_read.table("c:\mara\tesis\19302.txt", header=T)
x2_read.table("c:\mara\tesis\19402.txt", header=T)
x3_read.table("c:\mara\tesis\19502.txt", header=T)
x4_read.table("c:\mara\tesis\19602.txt", header=T)
x5_read.table("c:\mara\tesis\19702.txt", header=T)
x6_read.table("c:\mara\tesis\19802.txt", header=T)
x7_read.table("c:\mara\tesis\19902.txt", header=T)
```

Conversión de cada tabla a una matriz

```
x1.mat_as.matrix(x1)
x2.mat_as.matrix(x2)
x3.mat_as.matrix(x3)
x4.mat_as.matrix(x4)
x5.mat_as.matrix(x5)
x6.mat_as.matrix(x6)
x7.mat_as.matrix(x7)
```

```

# Reducción y estandarización de cada matriz
x1.std_sweep(x1.mat,2,apply(x1.mat,2,mean))
x1.std_sweep(x1.std,2,sqrt(apply(x1.mat,2,var)),FUN="/")
write(x1.std,"x1.std",ncol=6)
x2.std_sweep(x2.mat,2,apply(x2.mat,2,mean))
x2.std_sweep(x2.std,2,sqrt(apply(x2.mat,2,var)),FUN="/")
write(x2.std,"x2.std",ncol=6)
x3.std_sweep(x3.mat,2,apply(x3.mat,2,mean))
x3.std_sweep(x3.std,2,sqrt(apply(x3.mat,2,var)),FUN="/")
write(x3.std,"x3.std",ncol=6)
x4.std_sweep(x4.mat,2,apply(x4.mat,2,mean))
x4.std_sweep(x4.std,2,sqrt(apply(x4.mat,2,var)),FUN="/")
write(x4.std,"x4.std",ncol=6)
x5.std_sweep(x5.mat,2,apply(x5.mat,2,mean))
x5.std_sweep(x5.std,2,sqrt(apply(x5.mat,2,var)),FUN="/")
write(x5.std,"x5.std",ncol=6)
x6.std_sweep(x6.mat,2,apply(x6.mat,2,mean))
x6.std_sweep(x6.std,2,sqrt(apply(x6.mat,2,var)),FUN="/")
write(x6.std,"x6.std",ncol=6)
x7.std_sweep(x7.mat,2,apply(x7.mat,2,mean))
x7.std_sweep(x7.std,2,sqrt(apply(x7.mat,2,var)),FUN="/")
write(x7.std,"x7.std",ncol=6)

# Matrices de pesos de individuos/variables en cada estudio
# En este caso se asignaron pesos idénticos a cada uno de ellos
n_nrow(x1.std)
p_ncol(x1.std)
d_(1/n)*diag(n)
m_(1/p)*diag(p)

# Cálculo y escritura de las matrices de productos escalares entre individuos
w1_x1.std%*%m%*%t(x1.std)
w2_x2.std%*%m%*%t(x2.std)
w3_x3.std%*%m%*%t(x3.std)
w4_x4.std%*%m%*%t(x4.std)
w5_x5.std%*%m%*%t(x5.std)
w6_x6.std%*%m%*%t(x6.std)
w7_x7.std%*%m%*%t(x7.std)

write(w1,"w1",ncol=32)
write(w2,"w2",ncol=32)
write(w3,"w3",ncol=32)
write(w4,"w4",ncol=32)
write(w5,"w5",ncol=32)
write(w6,"w6",ncol=32)
write(w7,"w7",ncol=32)

```

```

# Cálculo de la matriz de productos escalares entre objetos Wi, Wj. (tr(w(i)*w(j)))
s_matrix(0,k,k)
for(i in 1:k) {
  for(j in 1:k) {
    s[i,j]_trace.x(scan(paste("w",i,sep="")))%*%scan(paste("w",j,sep="")))
  }
}
write(s,"s",ncol=7)

#Cálculo de la norma de la matriz Wi raiz(tr(w(i)*w(i)))
norma.w_matrix(0,k,1)
for(i in 1:k) norma.w[i]_sqrt(s[i,i])
write(norma.w,"norma.w")

# Matriz de pesos de los estudios; en este ejemplo se asignaron pesos idénticos.
delta_(1/k)*diag(k)

# Valores y vectores propios de s.delta
s.delta_s%*%delta
sd.valp_eigen(s.delta)$values[1:2]
sd.vccp_sqrt(k)*eigen(s.delta)$vectors[,1:2]
write(sd.valp,"sdvalps")
write(t(sd.vccp),"sdvccps",ncol=2)

# Nombres de individuos, variables y estudios. Aquí los estados también se agruparon de
# diferentes maneras para las gráficas posteriores.
estados_c(1:32)
edos1_c("1 AGS", "2 BC", "3 BCS", "4 CAMP", "5 COAH", "6 COL", "7 CHIS",
"8 CHIH", "9 DF", "10 DGO", "11 GTO", "12 GRO", "13 HGO", "14 JAL", "15 MEX",
"16 MICH")
edos2_c("17 MOR", "18 NAY", "19 NL", "20 OAX", "21 PUE", "22 QRO", "23 QR",
"24 SLP", "25 SIN", "26 SON", "27 TAB", "28 TAMS", "29 TLAX", "30 VER", "31 YUC",
"32 ZAC")
estados8_c("1 AGS", "2 BC", "3 BCS", "4 CAMP", "5 COAH", "6 COL", "7 CHIS",
"8 CHIH", "1 DF", "2 DGO", "3 GTO", "4 GRO", "5 HGO", "6 JAL", "7 MEX",
"8 MICH", "1 MOR", "2 NAY", "3 NL", "4 OAX", "5 PUE", "6 QRO", "7 QR",
"8 SLP", "1 SIN", "2 SON", "3 TAB", "4 TAMS", "5 TLAX", "6 VER", "7 YUC", "8 ZAC")
titvar_c("1 nac. hombres", "2 nac. mujeres", "3 def. generales", "4 def. niños", "5 matrimonios",
"6 divorcios")
variables_c("nac. hombres", "nac. mujeres", "def. generales", "def. niños", "matrimonios",
"divorcios")
censos_c("1930", "1940", "1950", "1960", "1970", "1980", "1990")

```

```

# Generación de coordenadas de la imagen Euclidiana de las tablas
coord_matrix(0,k,2)
coord[,1]_sqrt(sd.valp[1])*sd.vecp[,1]
coord[,2]_sqrt(sd.valp[2])*sd.vecp[,2]
write(t(coord),"coords",ncol=2)
win.graph()
par(cex=1,pty="s")
plot(coord[,1:2],xlab="cje I",ylab="", type="n",col=1)
par(cex=0.75)
text(coord[,1:2], censos)
par(bty="n")
legend(locator(), legend="Gráfica No. 1",density=-99)
par(cex=1)
title(main="Dinámica con respecto a los estados",sub="93.83% de la inercia explicada")

# Imagen Euclidiana centrada de las tablas
ident_diag(k)
cc_matrix(0,k,k)
unos_matrix(1,k,k)
cc_(ident-unos%%delta)%s%(ident-delta%%unos)
cc.valp_svd(cc%%delta)$d #[1:2]
cc.vecp_sqrt(k)*svd(cc)$v[,1:2]
write(cc.valp,"ccvalps",ncol=k)
write(cc.vecp,"ccvecps",ncol=2)
coord.c_matrix(0,k,2)
coord.c[,1]_sqrt(cc.valp[1])*cc.vecp[,1]
coord.c[,2]_sqrt(cc.valp[2])*cc.vecp[,2]
write(t(coord.c),"coord.c",ncol=2)
win.graph()
par(cex=.75,pty="s")
plot(coord.c[,1:2],xlab="",ylab="", type="n",axes=F,xlim=c(-12,12),ylim=c(-10,10))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
text(coord.c[,1:2], censos)
par(bty="n")
par(bty="o",cex=1)
title(main="Dinámica con respecto a los estados",sub="87.53% de la inercia explicada", sub="Gráfica No. 1")

```

```

# Cálculo de la matriz compromiso W
alfa_matrix(0,k,1)
alfa_(1/sqrt(sd.valp[1]))*(sum(delta%*%norma.w))*delta%*%sd.vecp[,1]
w_matrix(0,n,n)
w_alfa[1]*w1+alfa[2]*w2+alfa[3]*w3+alfa[4]*w4+alfa[5]*w5+alfa[6]*w6+alfa[7]*w7
write(alfa,"alfa")
write(t(w),"w",ncol=n)
wd_w
wd.valp_svd(wd)$d
wd.vecp_sqrt(n)*svd(wd)$v[,1:2]
write(wd.valp,"wdvalps",ncol=n)
write(wd.vecp,"wdvecps",ncol=2)

# Generación de coordenadas de la imagen euclidiana compromiso
coord.comp_matrix(0,n,2)
coord.comp[,1]_sqrt(wd.valp[1])*wd.vecp[,1]
coord.comp[,2]_sqrt(wd.valp[2])*wd.vecp[,2]
write(t(coord.comp),"coordcom",ncol=2)
win.graph()
par(cex=0.75,pty="s")
plot(coord.comp[,1:2],xlab="",ylab="",type="n",axes=F,xlim=c(-10,15),ylim=c(-5,10))
axis(1,pos=0,lty=1,labels=F)
axis(2,pos=0,lty=1,labels=F)
text(coord.comp[,1:2],estados)
par(bty="n")
legend(-20,8,legend=c(dos1,density=-99))
legend(20,8,legend=c(dos2,density=-99))
par(bty="o",cex=1)
title("Representación compromiso de los estados
87.96% de la inercia explicada",sub="Gráfica No. 2")

# Correlaciones de las variables con los ejes compromiso

cot1_matrix(0,p,2)
cot1[,1:2]_(1/p)*t(x1.std)%*%svd(wd)$v[,1:2]
write(t(cot1),"cot1",ncol=2)
cot2_matrix(0,p,2)
cot2[,1:2]_(1/p)*t(x2.std)%*%svd(wd)$v[,1:2]
write(t(cot2),"cot2",ncol=2)
cot3_matrix(0,p,2)
cot3[,1:2]_(1/p)*t(x3.std)%*%svd(wd)$v[,1:2]
write(t(cot3),"cot3",ncol=2)
cot4_matrix(0,p,2)
cot4[,1:2]_(1/p)*t(x4.std)%*%svd(wd)$v[,1:2]
write(t(cot4),"cot4",ncol=2)

```

```

cot5_matrix(0,p,2)
cot5[,1:2]_(1/p)*t(x5.std)%*%svd(wd)$v[,1:2]
write(t(cot5),"cot5",ncol=2)
cot6_matrix(0,p,2)
cot6[,1:2]_(1/p)*t(x6.std)%*%svd(wd)$v[,1:2]
write(t(cot6),"cot6",ncol=2)
cot7_matrix(0,p,2)
cot7[,1:2]_(1/p)*t(x7.std)%*%svd(wd)$v[,1:2]
write(t(cot7),"cot7",ncol=2)
variables_c(1:6)

win.graph()
par(cex=0.75,pty="s")
plot(cot1[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot1[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1930",sub="Gráfica No. 3")

win.graph()
par(cex=0.75,pty="s")
plot(cot2[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot2[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1940",sub="Gráfica No. 4")

win.graph()
par(cex=0.75,pty="s")
plot(cot3[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot3[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1950",sub="Gráfica No. 5")

```

```

cot5_matrix(0,p,2)
cot5[,1:2]_(1/p)*t(x5.std)%*%svd(wd)$v[,1:2]
write(t(cot5),"cot5",ncol=2)
cot6_matrix(0,p,2)
cot6[,1:2]_(1/p)*t(x6.std)%*%svd(wd)$v[,1:2]
write(t(cot6),"cot6",ncol=2)
cot7_matrix(0,p,2)
cot7[,1:2]_(1/p)*t(x7.std)%*%svd(wd)$v[,1:2]
write(t(cot7),"cot7",ncol=2)
variables_c(1:6)

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot1[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot1[,1:2],variables,adj=0)
par(bty="n")
legend(1.1,0.7,legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1930",sub="Gráfica No. 3")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot2[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot2[,1:2],variables,adj=0)
par(bty="n")
legend(1.1,0.7,legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1940",sub="Gráfica No. 4")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot3[,1:2],xlab="con eje I",ylab="con eje II",type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot3[,1:2],variables,adj=0)
par(bty="n")
legend(1.1,0.7,legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1950",sub="Gráfica No. 5")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot4[,1:2],xlab="con eje I",ylab="con eje II", type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot4[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1960",sub="Gráfica No. 6")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot5[,1:2],xlab="con eje I",ylab="con eje II", type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1, pos=0, lty=1, las=1)
axis(2, pos=0, lty=1, las=1)
text(cot5[,1:2], variables, adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1970",sub="Gráfica No. 7")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot6[,1:2],xlab="con eje I",ylab="con eje II", type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot6[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1980",sub="Gráfica No. 8")

```

```

win.graph()
par(cex=0.75,pty="s")
plot(cot7[,1:2],xlab="con eje I",ylab="con eje II", type="n",axes=F,xlim=c(0,1),ylim=c(-0.4,0.8))
axis(1,pos=0,lty=1,las=1)
axis(2,pos=0,lty=1,las=1)
text(cot7[,1:2],variables,adj=0)
par(bty="n")
legend(1.1, 0.7, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Correlaciones con los ejes compromiso: 1990",sub="Gráfica No. 9")

```


RUTINA No. 2

Construcción de trayectorias de individuos

```

tray.tab1_matrix(0,n,2)
tray.tab1[,1]_(1/sqrt(wd.valp[1]))*w1%*%wd.vccp[,1]
tray.tab1[,2]_(1/sqrt(wd.valp[2]))*w1%*%wd.vccp[,2]
write(t(tray.tab1),"traytab1",ncol=2)
tray.tab2_matrix(0,n,2)
tray.tab2[,1]_(1/sqrt(wd.valp[1]))*w2%*%wd.vccp[,1]
tray.tab2[,2]_(1/sqrt(wd.valp[2]))*w2%*%wd.vccp[,2]
write(t(tray.tab2),"traytab2",ncol=2)
tray.tab3_matrix(0,n,2)
tray.tab3[,1]_(1/sqrt(wd.valp[1]))*w3%*%wd.vccp[,1]
tray.tab3[,2]_(1/sqrt(wd.valp[2]))*w3%*%wd.vccp[,2]
write(t(tray.tab3),"traytab3",ncol=2)

```

```

tray.tab4_matrix(0,n,2)
tray.tab4[,1]_(1/sqrt(wd.valp[1]))*w4%*%wd.vccp[,1]
tray.tab4[,2]_(1/sqrt(wd.valp[2]))*w4%*%wd.vccp[,2]
write(t(tray.tab4),"traytab4",ncol=2)
tray.tab5_matrix(0,n,2)
tray.tab5[,1]_(1/sqrt(wd.valp[1]))*w5%*%wd.vccp[,1]
tray.tab5[,2]_(1/sqrt(wd.valp[2]))*w5%*%wd.vccp[,2]
write(t(tray.tab5),"traytab5",ncol=2)
tray.tab6_matrix(0,n,2)
tray.tab6[,1]_(1/sqrt(wd.valp[1]))*w6%*%wd.vccp[,1]
tray.tab6[,2]_(1/sqrt(wd.valp[2]))*w6%*%wd.vccp[,2]
write(t(tray.tab6),"traytab6",ncol=2)
tray.tab7_matrix(0,n,2)
tray.tab7[,1]_(1/sqrt(wd.valp[1]))*w7%*%wd.vccp[,1]
tray.tab7[,2]_(1/sqrt(wd.valp[2]))*w7%*%wd.vccp[,2]
write(t(tray.tab7),"traytab7",ncol=2)

```

Gráficas de las trayectorias

```

tray.todo_cbind(tray.tab1,tray.tab2,tray.tab3,tray.tab4,tray.tab5,tray.tab6,tray.tab7)
flecha_c(1, 3, 5, 7, 9, 11)
primcoor_c(1,3,5,7,9,11,13)
segcoor_c(2,4,6,8,10,12,14)

xx.todo_matrix(0,32,7)
yy.todo_matrix(0,32,7)
xx.todo_tray.todo[,primcoor]
yy.todo_tray.todo[,segcoor]

```

```

xx1_matrix(0,8,7)
yy1_matrix(0,8,7)
xx1_tray.todo[1:8,primcoor]
yy1_tray.todo[1:8,segcoor]

xx2_matrix(0,8,7)
yy2_matrix(0,8,7)
xx2_tray.todo[9:16,primcoor]
yy2_tray.todo[9:16,segcoor]

xx3_matrix(0,8,7)
yy3_matrix(0,8,7)
xx3_tray.todo[17:24,primcoor]
yy3_tray.todo[17:24,segcoor]

xx4_matrix(0,8,7)
yy4_matrix(0,8,7)
xx4_tray.todo[25:32,primcoor]
yy4_tray.todo[25:32,segcoor]

```

```

win.graph()
par(cex=0.75, pty="m")
matplot(t(xx.todo), t(yy.todo), xlab="nacimientos, defunciones y matrimonios",
ylab="divorcios", type="l", col=1, lty=1, axes=F, xlim=c(-10,20), ylim=c(-10,15))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
par(cex=1, bty="o")
title("Trayectorias de estados: 1930-1990", sub="Gráfica No. 10")

```

```

win.graph()
par(cex=0.75, pty="m")
matplot(t(xx1), t(yy1), xlab="nacimientos, defunciones y matrimonios", ylab="divorcios",
type="p", col=1, lty=1, axes=F, ylim=c(-5,15))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
arrows(tray.todo[1:8,flecha], tray.todo[1:8,flecha+1], tray.todo[1:8,flecha+2],
tray.todo[1:8,flecha+3], size=0.2, open=TRUE)
par(bty="n")
legend(-5, 13, legend=estados8[1:8], density=-99)
par(cex=1, bty="o")
title("Trayectorias de los estados I", sub="Gráfica No. 11")

```

```

win.graph()
par(cex=0.75, pty="n")
matplot( t(xx2), t(yy2), xlab="nacimientos, defunciones y matrimonios", ylab="divorcios",
type="p", col=1, lty=1, axes=F, xlim=c(-5,20))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
arrows(tray.todo[9:16, flecha], tray.todo[9:16, flecha+1], tray.todo[9:16, flecha+2],
tray.todo[9:16, flecha+3], size=0.2, open=TRUE)
par(bty="n")
legend(-5, 6, legend=estados8[9:16], density=-99)
par(cex=1, bty="o")
title("Trayectorias de los estados II", sub="Gráfica No. 12")

```

```

win.graph()
par(cex=0.75, pty="m")
matplot( t(xx3), t(yy3), xlab="nacimientos, defunciones y matrimonios", ylab="divorcios",
type="p", col=1, lty=1, axes=F, xlim=c(-7.5,7.5), ylim=c(-7,7))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
arrows(tray.todo[17:24, flecha], tray.todo[17:24, flecha+1], tray.todo[17:24, flecha+2],
tray.todo[17:24, flecha+3], size=0.2, open=TRUE)
par(bty="n")
legend(7, 6, legend=estados8[17:24], density=-99)
par(cex=1, bty="o")
title("Trayectorias de los estados III", sub="Gráfica No. 13")

```

```

win.graph()
par(cex=0.75, pty="m")
matplot( t(xx4), t(yy4), xlab="nacimientos, defunciones y matrimonios", ylab="divorcios",
type="p", col=1, lty=1, axes=F, xlim=c(-5,10), ylim=c(-5,7))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
arrows(tray.todo[25:32, flecha], tray.todo[25:32, flecha+1], tray.todo[25:32, flecha+2],
tray.todo[25:32, flecha+3], size=0.2, open=TRUE)
par(bty="n")
legend(2,7, legend=estados8[25:32], density=-99)
par(cex=1, bty="o")
title("Trayectorias de los estados IV", sub="Gráfica No. 14")

```

```

win.graph()
par(cex=0.75, pty="m")
matplot( t(xx4), t(yy4), xlab="nacimientos, defunciones y matrimonios", ylab="divorcios",
type="p", col=1, lty=1, axes=F, xlim=c(-5,1), ylim=c(-5,7))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
arrows(tray.todo[25:32,flecha], tray.todo[25:32,flecha+1], tray.todo[25:32,flecha+2],
tray.todo[25:32,flecha+3], size=0.2, open=TRUE)
par(bty="n")
legend(0.25,6, legend=estados8[25:32], density=-99)
par(cex=1, bty="o")
title("Trayectorias de los estados IV.
Acercamiento", sub="Gráfica No. 14")

```

RUTINA No. 3

Caso en el que el espacio de variables se mantiene fijo

Matrices de productos escalares entre variables

```

v1_(1/(n-1))*t(x1.std)%*%x1.std
v2_(1/(n-1))*t(x2.std)%*%x2.std
v3_(1/(n-1))*t(x3.std)%*%x3.std
v4_(1/(n-1))*t(x4.std)%*%x4.std
v5_(1/(n-1))*t(x5.std)%*%x5.std
v6_(1/(n-1))*t(x6.std)%*%x6.std
v7_(1/(n-1))*t(x7.std)%*%x7.std

```

Salvar

```

write(v1,"v1",ncol=6)
write(v2,"v2",ncol=6)
write(v3,"v3",ncol=6)
write(v4,"v4",ncol=6)
write(v5,"v5",ncol=6)
write(v6,"v6",ncol=6)
write(v7,"v7",ncol=6)

```

Cálculo de la matriz de productos escalares entre objetos vi, vj: (tr(v(i)*v(j)))

```

s2_matrix(0,k,k)
for(i in 1:k) {
  for(j in 1:k) {
    s2[i,j]_trace.x(scan(paste("v",i,sep="")))%*%scan(paste("v",j,sep="")))
  }
}
write(s2,"s2",ncol=7)

```

```

# Cálculo de la norma de la matriz vi: raiz(tr(v(i)*v(i)))
norma.v_nmatrix(0,k,1)
for(i in 1:k) norma.v[i]_sqrt(s2[i,i])
write(norma.v,"norma.v")

# Matriz de pesos de los estudios; en este ejemplo se asignaron
# pesos idénticos para cada uno de ellos.
delta2_(1/k)*diag(k)

# Valores y vectores propios de s2.delta
s2.delta_s2%%delta2
sd2.valp_eigen(s2.delta)$values #[1:2]
sd2.vecp_sqrt(k)*eigen(s2.delta)$vectors[,1:2]
write(sd2.valp,"sd2valps")
write(t(sd2.vecp),"sd2vecps",ncol=2)

# Generación de coordenadas de la imagen Euclidiana de las tablas
coord2_matrix(0,k,2)
coord2[,1]_sqrt(sd2.valp[1])*sd2.vecp[,1]
coord2[,2]_sqrt(sd2.valp[2])*sd2.vecp[,2]
write(t(coord2),"coords2",ncol=2)

# Imagen Euclidiana centrada de las tablas
ident_diag(k)
cc2_matrix(0,k,k)
cc2_(ident-unos%%delta2)%%s2%%(ident-delta2%%unos)
cc2.valp_svd(cc2)$d #[1:2]
cc2.vecp_sqrt(k)*svd(cc2)$v[,1:2]
write(cc2.valp,"cvalps2")
write(cc2.vecp,"cvecps2",ncol=2)
coord.c2_matrix(0,k,2)
coord.c2[,1]_sqrt(cc2.valp[1])*cc2.vecp[,1]
coord.c2[,2]_sqrt(cc2.valp[2])*cc2.vecp[,2]
write(t(coord.c2),"coordse2",ncol=2)

win.graph()
par(cex=0.75, pty="s")
plot(coord.c2[,1:2],xlab="",ylab=" ", type="n", axes=F, xlim=c(-4,4), ylim=c(-0.8,0.6))
axis(1, pos=0, lty=1, las=1)
axis(2, pos=0, lty=1, las=1)
text(coord.c2[,1:2], censos)
par(bty="n")
par(cex=1,bty="o")
title(main="Dinámica con respecto a las variables
99.14% de la inercia explicada", sub="Gráfica No. 15")

```

```

# Cálculo de la matriz compromiso v
alfa2_matrix(0,k,1)
alfa2_delta2%*%sd2.vecp[,1]
v_matrix(0,p,p)
v_(1/sum(alfa2))*(alfa2[1]*v1 + alfa2[2]*v2 + alfa2[3]*v3 + alfa2[4]*v4 + alfa2[5]*v5 +
alfa2[6]*v6 + alfa2[7]*v7)
write(t(v),"v",ncol=n)
vm_v
vm.valp_svd(vm)$d
vm.vecp_sqrt(p)*svd(vm)$v[,1:2]
write(vm.valp,"vmvalps",ncol=n)
write(vm.vecp,"vmvecps",ncol=2)

# Generación de coordenadas de la imagen euclidiana compromiso
coord2.comp_matrix(0,p,2)
coord2.comp[,1]_sqrt(vm.valp[1])*vm.vecp[,1]
coord2.comp[,2]_sqrt(vm.valp[2])*vm.vecp[,2]
write(t(coord.comp),"coordcom.2",ncol=2)

win.graph()
par(cex=0.75,pty="s")
plot(coord2.comp[,1:2],xlab="",ylab="",type="n",axes=F,xlim=c(-2.5,0.5),ylim=c(-1,3))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
text(coord2.comp[,1:2], variables)
par(bty="n")
legend(0.25, 2.5, legend=titvar,density=-99)
par(cex=1,bty="o")
title("Representación compromiso de
variables demográficas",
sub="Gráfica No. 16")

# Correlaciones de las variables con los ejes compromiso
cotv1_matrix(0,n,2)
cotv1[,1:2]_(1/n)*(x2.std)%*%svd(vm)$v[,1:2]
write(t(cotv1),"cotv2",ncol=2)
cotv2_matrix(0,n,2)
cotv2[,1:2]_(1/n)*(x3.std)%*%svd(vm)$v[,1:2]
write(t(cotv2),"cotv2",ncol=2)
cotv3_matrix(0,n,2)
cotv3[,1:2]_(1/n)*(x4.std)%*%svd(vm)$v[,1:2]
write(t(cotv3),"cotv3",ncol=2)
cotv4_matrix(0,n,2)
cotv4[,1:2]_(1/n)*(x5.std)%*%svd(vm)$v[,1:2]
write(t(cotv4),"cotv4",ncol=2)
cotv5_matrix(0,n,2)
cotv5[,1:2]_(1/n)*(x5.std)%*%svd(vm)$v[,1:2]

```

```

write(t(cotv5),"cotv5",ncol=2)
cotv6_matrix(0,n,2)
cotv6[,1:2]_(1/n)*(x6.std)%*%svd(vn)$v[,1:2]
write(t(cotv6),"cotv6",ncol=2)
cotv7_matrix(0,n,2)
cotv7[,1:2]_(1/n)*(x7.std)%*%svd(vn)$v[,1:2]
write(t(cotv7),"cotv7",ncol=2)

```

RUTINA No. 4

Construcción de trayectorias de individuos

```

tr2.tab1_matrix(0,p,2)
tr2.tab1[,1]_(1/sqrt(vn.valp[1]))*v1%*%vm.vecp[,1]
tr2.tab1[,2]_(1/sqrt(vn.valp[2]))*v1%*%vm.vecp[,2]
write(t(tr2.tab1),"tr2tab1",ncol=2)
tr2.tab2_matrix(0,p,2)
tr2.tab2[,1]_(1/sqrt(vn.valp[1]))*v2%*%vm.vecp[,1]
tr2.tab2[,2]_(1/sqrt(vn.valp[2]))*v2%*%vm.vecp[,2]
write(t(tr2.tab2),"tr2tab2",ncol=2)
tr2.tab3_matrix(0,p,2)
tr2.tab3[,1]_(1/sqrt(vn.valp[1]))*v3%*%vm.vecp[,1]
tr2.tab3[,2]_(1/sqrt(vn.valp[2]))*v3%*%vm.vecp[,2]
write(t(tr2.tab3),"tr2tab3",ncol=2)
tr2.tab4_matrix(0,p,2)
tr2.tab4[,1]_(1/sqrt(vn.valp[1]))*v4%*%vm.vecp[,1]
tr2.tab4[,2]_(1/sqrt(vn.valp[2]))*v4%*%vm.vecp[,2]
write(t(tr2.tab4),"tr2tab4",ncol=2)
tr2.tab5_matrix(0,p,2)
tr2.tab5[,1]_(1/sqrt(vn.valp[1]))*v5%*%vm.vecp[,1]
tr2.tab5[,2]_(1/sqrt(vn.valp[2]))*v5%*%vm.vecp[,2]
write(t(tr2.tab5),"tr2tab5",ncol=2)

tr2.tab6_matrix(0,p,2)
tr2.tab6[,1]_(1/sqrt(vn.valp[1]))*v6%*%vm.vecp[,1]
tr2.tab6[,2]_(1/sqrt(vn.valp[2]))*v6%*%vm.vecp[,2]
write(t(tr2.tab6),"tr2tab6",ncol=2)
tr2.tab7_matrix(0,p,2)
tr2.tab7[,1]_(1/sqrt(vn.valp[1]))*v7%*%vm.vecp[,1]
tr2.tab7[,2]_(1/sqrt(vn.valp[2]))*v7%*%vm.vecp[,2]
write(t(tr2.tab7),"tr2tab7",ncol=2)

```

Gráficas de las trayectorias

```

tray2.todo_cbind(tr2.tab1,tr2.tab2,tr2.tab3,tr2.tab4,tr2.tab5,tr2.tab6,tr2.tab7)
flecha_c(1, 3, 5, 7, 9, 11)
primcoor_c(1, 3, 5, 7, 9, 11, 13)
segcoor_c(2, 4, 6, 8, 10, 12, 14)

```

```

xx2.todo_matrix(0,6,7)
yy2.todo_matrix(0,6,7)
xx2.todo_tray2.todo[,primcoor]
yy2.todo_tray2.todo[,segcoor]

```

```

win.graph()
par(cex=0.75, pty="m")
matplot(t(xx2.todo), t(yy2.todo), type="b", lty=1, xlab="", ylab="", col=1, axes=F, xlim=c(-2.5,0), ylim=c(-1,2.5))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=0, lty=1, labels=F)
ss_c(1,2,3,4,5,6)
arrows(xx2.todo[,ss], yy2.todo[,ss],xx2.todo[,ss+1], yy2.todo[,ss+1], size=0.2, open=TRUE)
par(bty="n")
legend(-1, 1.5, legend=titvar, density=-.99)
par(bty="o",cex=1)
title("Trayectorias de las variables",sub="Gráfica No. 17")

```

```

win.graph()
par(cex=0.75, pty="m")
matplot(t(xx2.todo), t(yy2.todo), type="b", lty=1, xlab="", ylab="", col=1, axes=F, xlim=c(-2.5,-2), ylim=c(-1,1.2))
axis(1, pos=0, lty=1, labels=F)
axis(2, pos=-2, lty=3, labels=F)
ss_c(1,2,3,4,5,6)
arrows(xx2.todo[,ss], yy2.todo[,ss],xx2.todo[,ss+1], yy2.todo[,ss+1], size=0.2, open=TRUE)
par(bty="n")
legend(-2.2, 1, legend=titvar, density=-.99)
par(bty="o",cex=1)
title("Trayectorias de las variables.
Acercamiento",sub="Gráfica No. 17 ")

```


Bibliografía

- [1] Becker, R. A., Chambers, J. M. & Wilks, A. R.. *The New Language S Plus. A Programming Environment for Data Analysis and Graphics*. Wadsworth & Brooks/Cole Advanced Books & Software. USA 1988.
- [2] Bove, G., Ciaccio, A., *A User-oriented Overview of Multiway Methods and Software*. Computational Statistics & Data Analysis. Volumen 18, 1994.
- [3] Cailliez, F. & Pages, J. P. *Introduction à l'Analyse des Données*. Smash. París, 1976.
- [4] Coppi, R., *An Introduction to Multiway Data and their Analysis*. Computational Statistics & Data Analysis. Volumen 18, 1994.
- [5] Lavit, C., *Analyse Conjointe de Tableaux Quantitatifs*, Editorial Masson. París, 1988.
- [6] Lavit, C., Escoufier, Y. et. al., *The ACT (Statis Method)*. Computational Statistics & Data Analysis. Volumen 18, 1994.
- [7] Friedberg, S. H., Insel, A. J., Spence, L. E. *Algebra Lineal*. Publicaciones Cultural, S.A. 1a. edición. México, 1982.
- [8] González J., Rodríguez, O., *Algoritmos e Implementación del Método Statis*. IX Simposio de Métodos Matemáticos Aplicados a las Ciencias. Costa Rica, febrero 1995.
- [9] Hernández, C., Rubén. *Análisis Multivariado Exploratorio*. Instituto Tecnológico Autónomo de México; Documentos de Trabajo. México, 1995.

- [10] *Anuario Estadístico de los Estados Unidos Mexicanos*. Instituto Nacional de Estadística, Geografía e Informática (INEGI). México 1994.