



CIMPA-UCR

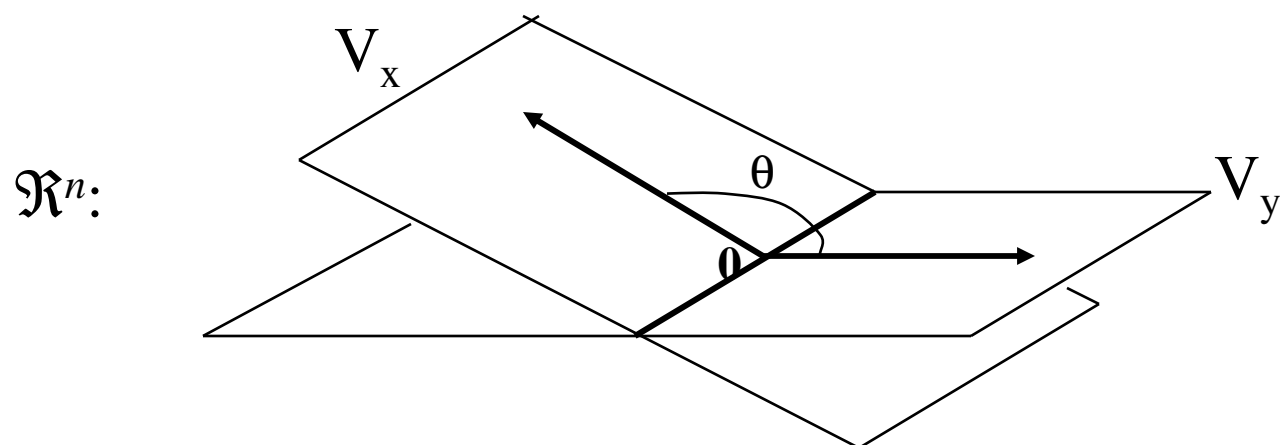
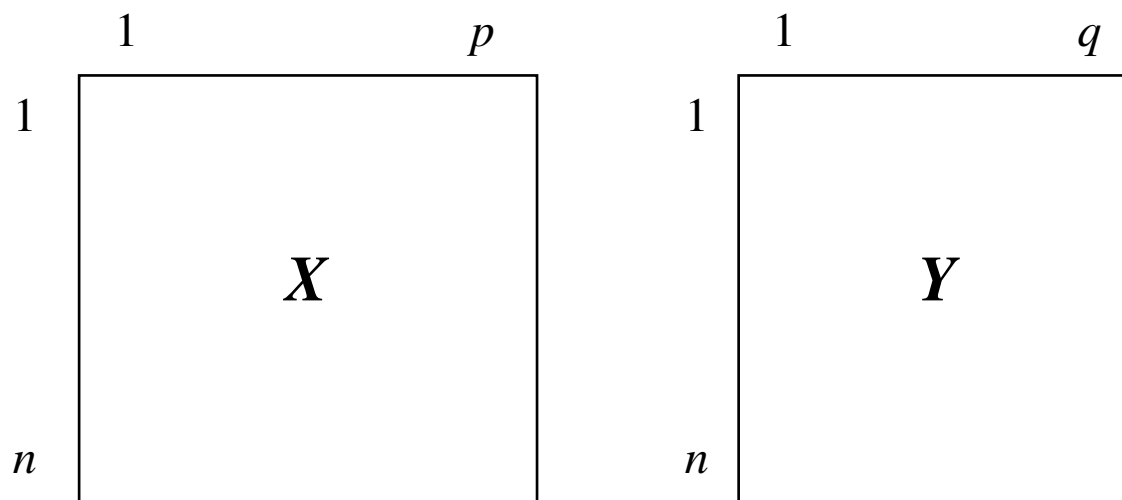
Análisis de Tablas Múltiples

Análisis Canónico

Método Statis

Análisis de Factorial Múltiple

Diagrama del Analisis Canónico





CIMPA-UCR

Situación del Análisis Canónico

- Situación: se tiene dos paquetes de variables cuantitativas observadas o medidas en n individuos

$$x^1, x^2, \dots, x^p \in \mathbb{R}^n$$

$$y^1, y^2, \dots, y^q \in \mathbb{R}^n$$

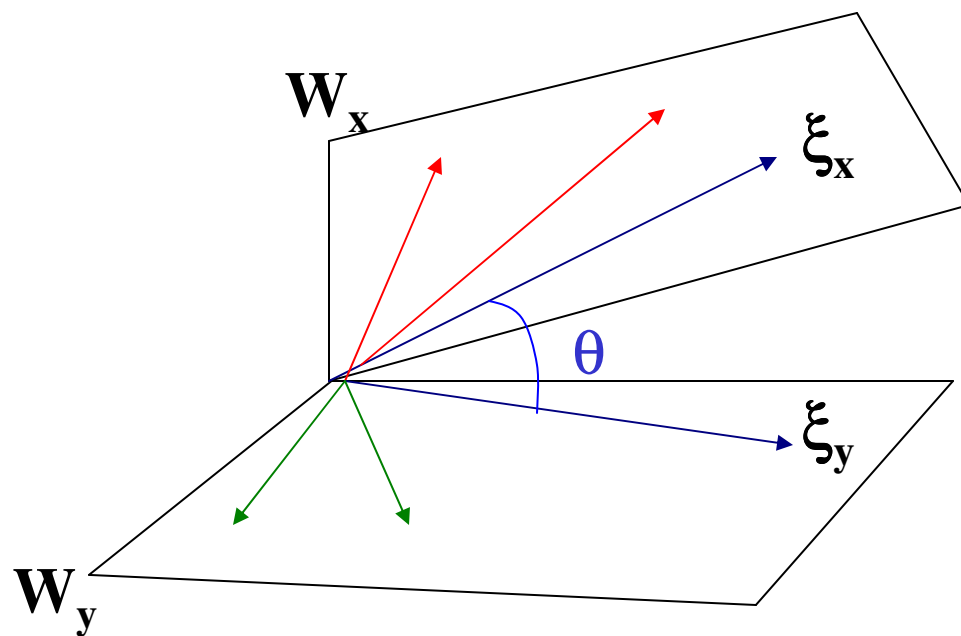
- Objetivo: hallar $\xi_x = a_1 \vec{x}^1 + \dots + a_q \vec{x}^q$
tales que $\xi_y = b_1 \vec{y}^1 + \dots + b_q \vec{y}^q$
 $\rho^2(\xi_x, \xi_y)$ sea máxima

Análisis de Tablas Múltiples



CIMPA-UCR

	x^1	x^j	x^p		y^1	y^2	y^q
1	X			1	y		
i				i			
n				n			



Objetivo

Hallar $\mathbf{a}^1$ generada por las variables de $\mathbf{X}$ y
 $\mathbf{b}^1$ generada por las variables de $\mathbf{Y}$ tales que

$$\text{Max } r^2(\mathbf{X}\mathbf{a}^1, \mathbf{Y}\mathbf{b}^1), \quad \|\mathbf{a}^1\| = 1 = \|\mathbf{b}^1\|$$

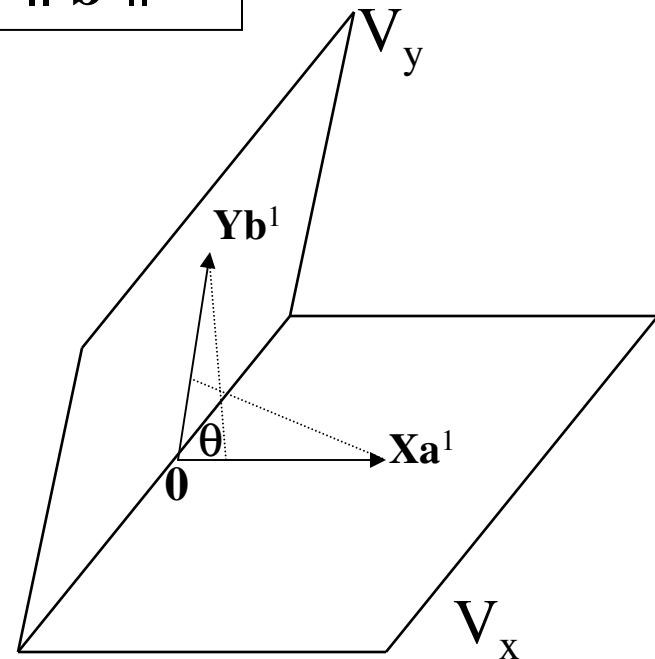
Solución:

$$\mathbf{a}^1 = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y} \mathbf{b}^1 / \lambda$$

$$\mathbf{b}^1 = (\mathbf{Y}^t \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{X} \mathbf{a}^1 / \lambda$$

$$\text{Sqrt}(\lambda) = \cos(\theta) = \cos(\mathbf{X}\mathbf{a}^1, \mathbf{Y}\mathbf{b}^1)$$

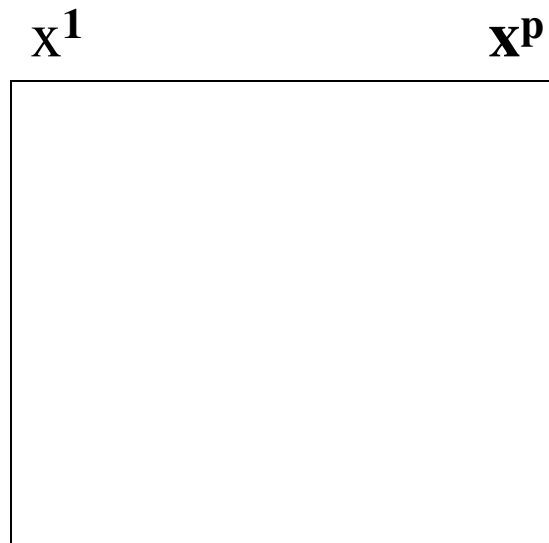
Correlación canónica





Casos particulares del Análisis Canónico

Regresión lineal Múltiple



P cuantitativas



1 cuantitativa



CIMPA-UCR

Discriminación

x^1

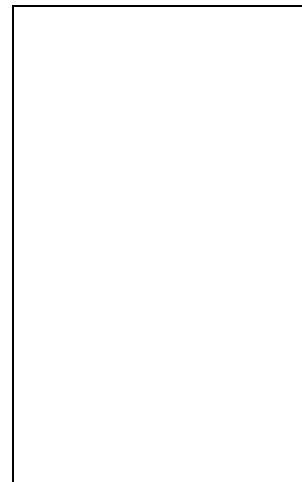
x^p



P cuantitativas

y^1

y^q



1 cualitativa con
q modalidades
(indicatrices)

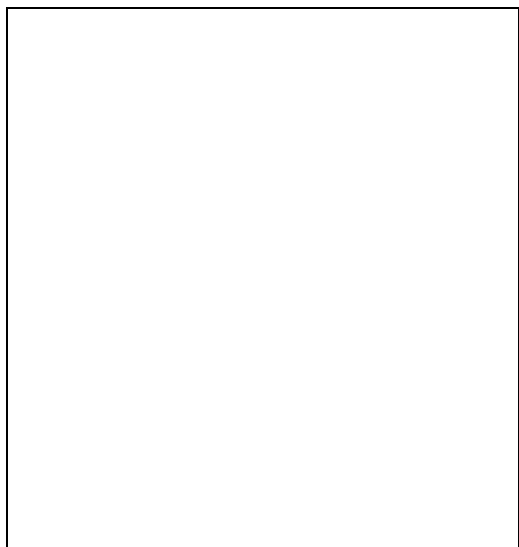


CIMPA-UCR

Análisis de Correspondencias

x^1

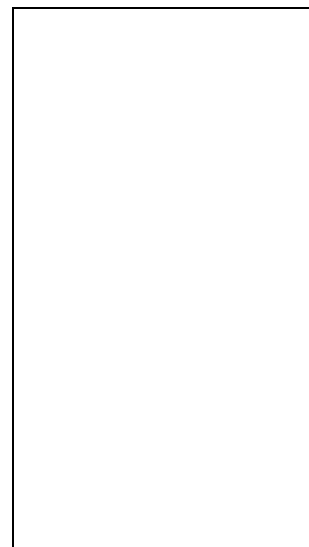
x^p



1 cualitativa
con p modalidades

y^1

y^q



1 cualitativa
con q modalidades

indicatrices



Análisis Canónico: Formulación Geométrica

$$\left. \begin{array}{l} x^1, \dots, x^p \in R^n \\ y^1, \dots, y^q \in R^n \end{array} \right\} \text{centrados}$$

$$\text{Ponderación} \\ p_i > 0, \sum_{i=1}^n p_i = 1$$

Espacios generados: $W_x = \{\xi_x \in R^n / x = Xa, a \in R^p\}$

$$W_y = \{\xi_y \in R^n / y = Yb, b \in R^q\}$$

$$\text{Supongamos: } \begin{cases} \text{rang}(X) = p \Rightarrow \dim(W_x) = p \\ \text{rang}(Y) = q \Rightarrow \dim(W_y) = q \end{cases}$$

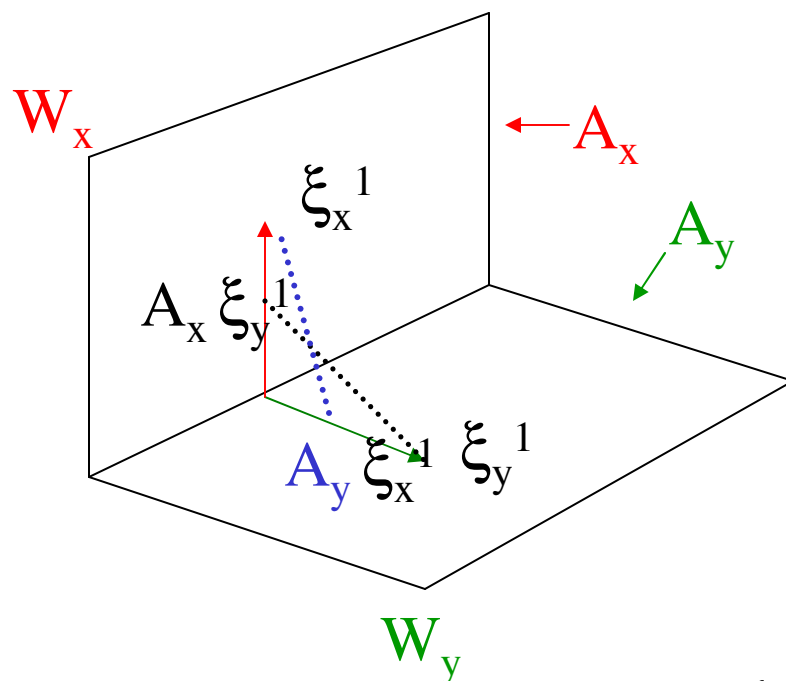


Problema:

CIMPA-UCR

Encontrar $\xi_x^1 \in W_x$, $\xi_y^1 \in W_y$ tales que

$$\text{Max } \cos^2(\xi_x^1, \xi_y^1) = r^2(\xi_x^1, \xi_y^1)$$



A_x : op proyección
ortogonal sobre W_x

A_y : op proyección
ortogonal sobre W_y

Se puede tomar

$$\|\xi_x^1\| = 1, \|\xi_y^1\| = 1$$

Se tiene: $A_x = X(X^t D X)^{-1} X^t D$

$$A_y = Y(Y^t D Y)^{-1} Y^t D$$

(ξ_x^1, ξ_y^1) se llama pareja
canónica y $\lambda = r^2(\xi_x^1, \xi_y^1)$

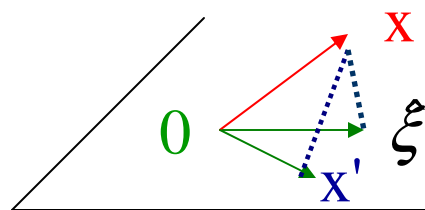
primera correlación
canónica



CIMPA-UCR

Análisis Canónico: deducción de solución

Propiedad: Sea $x \in W_x$ con $\|x\| = 1$. Entonces el vector de W_y con ángulo mínimo con x es $\hat{\xi} = A_y x$



$$\text{Sen}^2(x, x') = \frac{\|x - x'\|^2}{\|x\|^2}$$

Por T. Pitágoras: $\|x - x'\|^2 = \|x - \hat{\xi}\|^2 + \|\hat{\xi} - x'\|^2$

$$\text{Luego } \text{sen}^2(x, x') = \frac{\|\hat{\xi}\|^2}{\|x\|^2} + \frac{\|\hat{\xi} - x'\|^2}{\|x\|^2}$$



CIMPA-UCR

Hay un mínimo canónico $\xi_x^1 = \hat{\xi}$, la proyección ortogonal

Luego, ξ_y^1 debe ser colineal con $A_y \xi_x^1$:

$$A_y \xi_x^1 = r_1 \xi_y^1$$

donde $r_1 = \langle \xi_x^1, \xi_y^1 \rangle = \cos(\xi_x^1, \xi_y^1) = r(\xi_x^1, \xi_y^1)$

Se deduce

$$\left. \begin{aligned} A_x A_y \xi_x^1 &= \lambda_1 \xi_x^1 \\ A_y A_x \xi_y^1 &= \lambda_1 \xi_y^1 \end{aligned} \right\} \text{con } \lambda_1 = r_1^2 = \cos^2(\xi_x^1, \xi_y^1)$$

Por lo tanto ξ_x^1 es vector propio de $A_x A_y$

ξ_y^1 es vector propio de $A_y A_x$

asociados al mayor valor propio $\lambda_1 = r^2(\xi_x^1, \xi_y^1)$



CIMPA-UCR

Observaciones

$$\xi_y^1 = \frac{1}{\sqrt{\lambda_1}} A_y \xi_x^1 \quad \xi_x^1 = \frac{1}{\sqrt{\lambda_1}} A_x \xi_y^1$$

Siguientes variables canónicas: $(\xi_x^2, \xi_y^2), (\xi_x^3, \xi_y^3), \dots$

con máxima correlación y tales que $\xi_x^1 \perp \xi_x^2, \xi_y^1 \perp \xi_y^2, \dots$

son vectores propios de $A_x A_y, A_y A_x$ resp. asociados

a los valores propios siguientes: $\lambda_2 \geq \lambda_3 \geq \dots$



Cálculo de Factores Canónicos

$$\text{Como } \xi_x \in W_x \Rightarrow \xi_x = a_1 x^1 + \dots + a_p x^p = Xa \text{ con } a = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}$$

$$\xi_y \in W_y \Rightarrow \xi_y = b_1 y^1 + \dots + b_q y^q = Yb$$

$$\text{Tenemos } \left. \begin{array}{l} A_x = X(X^t DX)^{-1} X^t D \\ A_y = Y(Y^t DY)^{-1} Y^t D \end{array} \right\} \text{ op. de proyección ortogonal}$$



Análisis de Tablas Múltiples

Como $A_x A_y \xi_x^1 = \lambda_1 \xi_x^1$, $A_y A_x \xi_y^1 = \lambda_1 \xi_y^1$ poniendo $V_x = X^t D X$ $V_y = Y^t D Y$ $V_{xy} = X^t D Y = V_{yx}^t$

entonces

$$X V_x^{-1} V_{xy} V_y^{-1} Y^t D \xi_x^1 = X V_x^{-1} V_{xy} V_y^{-1} V_{yx} a = \lambda X a$$

Análogamente

$$V_y^{-1} V_{yx} V_x^{-1} V_{xy} b = \lambda b$$

Los factores canónicos a , b son vectores propios de

$$V_x^{-1} V_{xy} V_y^{-1} V_{yx}$$

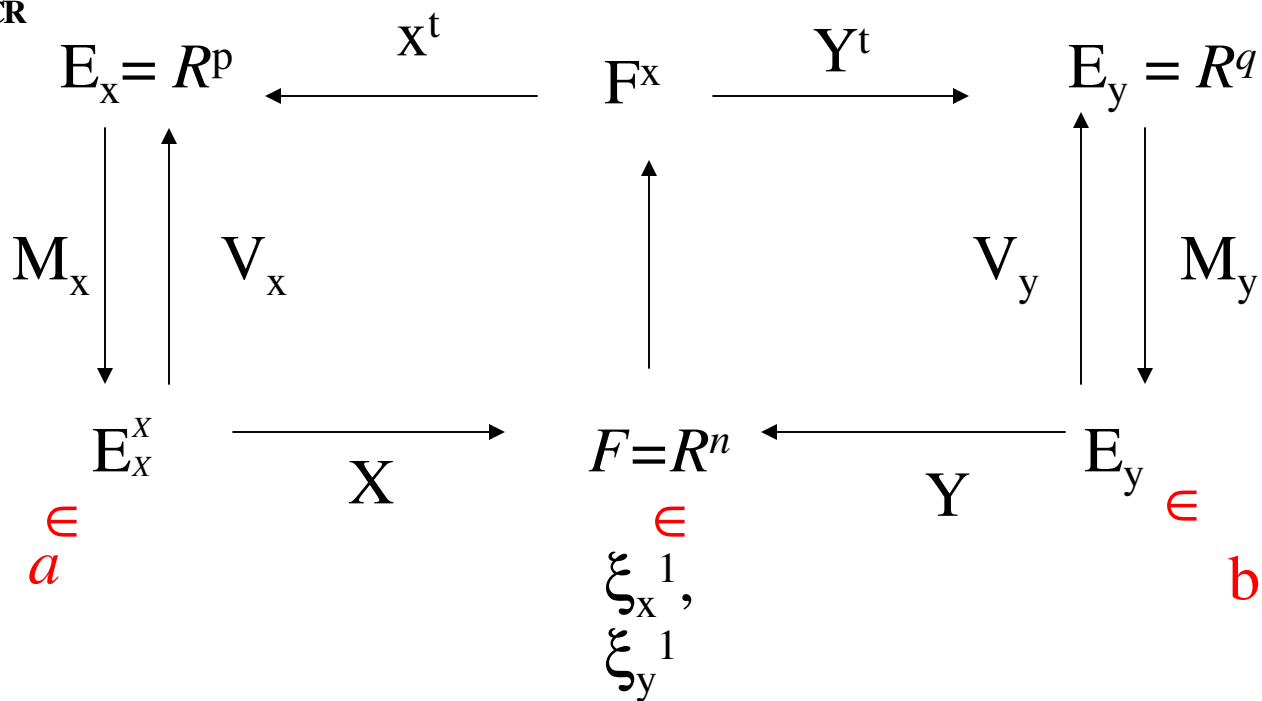
$$V_y^{-1} V_{yx} V_x^{-1} V_{xy} \text{ respectivamente asociados}$$

al mayor valor propio λ



CIMPA-UCR

Análisis de Tablas Múltiples





CIMPA-UCR

Análisis Canónico: Ejemplo

$n=40$ Fuente: R. Thomas (1975) y Bouroche- Saporta (1981)

X: características físicas y dinámicas de los atletas

TAM: estatura (cms)

PESO: peso (kgs)

DTH: salto largo, sin impulso, con pies juntos (cms)

DTV: salto alto, sin impulso, manos levantadas, talón al suelo (cms)

FPIE: fuerza piernas (kgs): peso levantado sobre los hombros

VELO: velocidad (decena de seg): tiempo para recorrer 30 mts

SALR: salto largo (cms): mejor resultado

3SAL: triple salto (mts): mejor resultado



Análisis de Tablas Múltiples

Y promedio de nota (sobre 20) dada por 3 jueces al rendimiento de los atletas

CIMPA-UCR

NSAL: nota en salto

NELA: nota en impulso (estilo)

NIMP: nota en impulso

NISUR: nota de suspensión recepción

	$\bar{X}$	$\diamond_x$		$\bar{X}$	$\diamond_x$
TAM	178	6.1	NSAL	10.1	1.8
PESO	72.5	7.6	NELA	9.9	1.8
DTH	261	15.7	NINP	10.1	1.1
DTV	65.5	5.1	NSUR	10	1.7
FPIE	109	17.8			
VELO	33.5	1.3			
SALR	583	39.1			
SAL3	11.4	0.9			



CIMPA-UCR

Análisis Canónico: Atletas

Correlaciones Canónicas:

	Valor propio	Correlación canónica
1	0.707	0.841
2	0.309	0.556
3	0.177	0.421
4	0.060	0.246



Correlaciones entre variables canónicas y variables originales

CIMPA-UCR

	$\boxed{\times}$	$\boxed{\times}^2$	η_0	η_0^2
TAM	0.073	-0.025	0.061	-0.014
PESO	-0.208	0.290	-0.175	0.161
DTH	0.468	0.197	0.394	0.109
DTV	0.324	0.183	0.273	0.101
FPIE	-0.061	0.328	-0.051	0.182
VELO	-0.705	-0.012	-0.593	-0.006
SALR	0.918	-0.066	0.772	-0.036
3SAL	0.741	0.436	0.628	0.242
NSAL	0.809	-0.027	0.962	-0.049
NELA	0.768	-0.177	0.913	-0.318
NIMP	0.762	0.174	0.906	0.313
NSUR	0.667	-0.013	0.743	-0.023



CIMPA-UCR

Análisis Canónico: Atletas

Correlaciones

	TAM	PESO	DTH	DTV	FPIE	VELO	SALR	3SAL
TAM	1							
PESO	<u>0.77</u>	1						
DTH	0.51	0.27	1					
DTV	0.16	0.04	0.62	1				
FPIE	0.47	<u>0.74</u>	0.36	0.23	1			
VELO	-0.23	-0.09	-0.43	-0.33	-0.05	1		
SALR	0.29	0.05	0.59	0.39	0.06	-0.63	1	
3SAL	0.31	-0.02	0.64	0.47	-0.05	-0.54	<u>0.67</u>	1



CIMPA-UCR

Análisis de Tablas Múltiples

NSAL

NELA

NIMP

NSUR

NSAL

1

NELA

0.83

1

NIMP

0.80

0.79

1

NSUR

0.82

0.69

0.77

1



CIMPA-UCR

Análisis de Tablas Múltiples

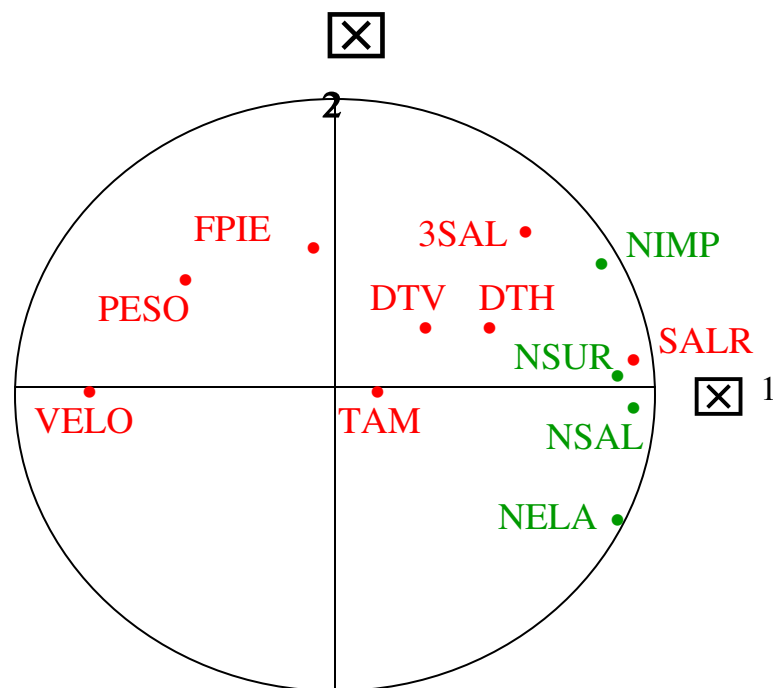
	NSAL	NELA	NIMP	NSUR
TAM	0.03	0.08	0.05	-0.05
PESO	-0.19	-0.20	-0.10	-0.18
DTH	0.31	0.38	0.42	0.18
DTV	0.23	0.24	0.26	0.06
FPIE	-0.09	-0.07	0.03	-0.11
VELO	-0.53	-0.58	-0.57	-0.41
SALR	<u>0.75</u>	<u>0.71</u>	<u>0.68</u>	0.61
3SAL	0.58	0.50	0.63	0.43



Análisis Canónico: Atletas

CIMPA-UCR

Círculo de correlaciones (81%)



- Relación: -Notas de los jueces
-Rendimiento en salto (SALR, 3SAL)
- Velocidad parece oponerse a las notas y rendimiento en salto



Análisis Canónico

Ventajas:

- valor teórico
- generaliza

(ACP, AFC, ACM, AFD, Regresión L.M., ANOVA)

Limitaciones:

- se trabaja sólo en R^n : espacio de variables
- se estudia únicamente la estructura de las correlaciones y no una representación fiel, una proyección, de las tablas de datos originales.

Propiedad: El análisis canónico de x, y equivale al ACP de $Z = [x : y]$ con la métrica

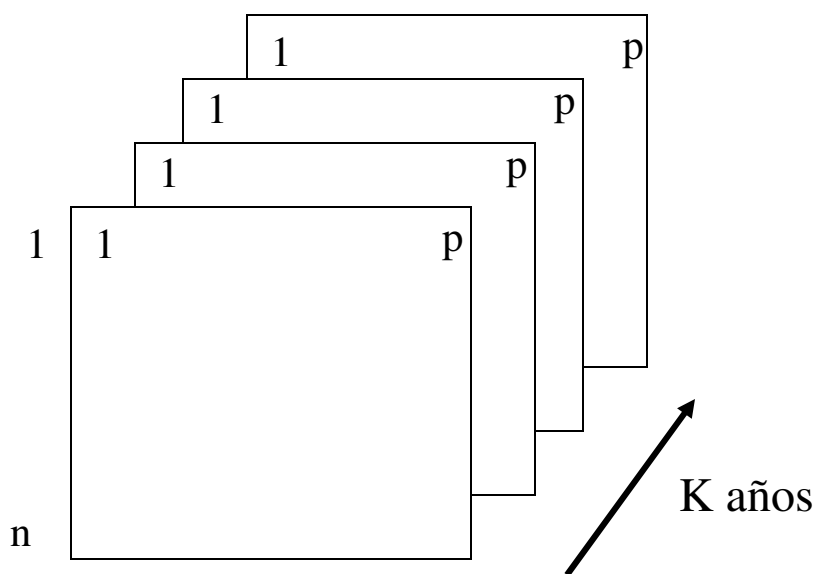
$$M_z = \begin{bmatrix} V_x^{-1} & 0 \\ 0 & V_y^{-1} \end{bmatrix}$$



CIMPA-UCR

MÉTODO STATIS

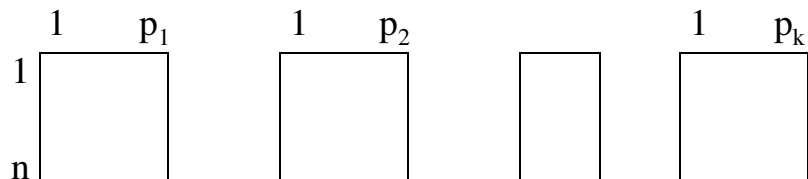
Estructuración de tablas con Tres Indices de la Estadística
“*Structuration des Tableaux à Trois Indices de la Statistique*”
(Y. Escoufier et al.)



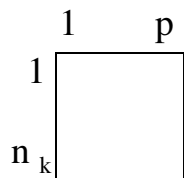
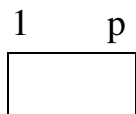
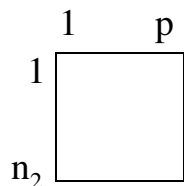
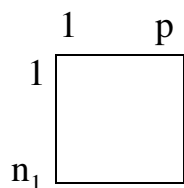
P. ej.: para analizar
tablas múltiples
(datos evolutivos)



Análisis conjunto de varias tablas de datos



- Mismo individuo
- Varios grupos de variables



- Mismas variables
- Varios grupos de individuos
(Statis dual)



MÉTODO STATIS

Situación:

K tablas $X_1, X_2, \dots, X_k$ de variables cuantitativas (centradas, estandarizadas) observadas sobre n individuos

Objetivos:

- Comparar las tablas
- Describir la estructura común a las diferentes tablas
- Conocer las diferencias entre las tablas

Notación:

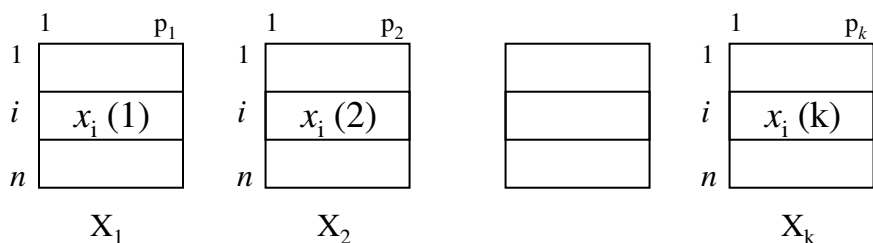
p_k : # de variables del grupo k

$$p = p_1 + \dots + p_k$$

$$X = [X_1 \mid X_2 \mid \dots \mid X_k]$$

x_i : Fila i de X

$$x_i(k): \text{fila } i \text{ de } X_k \Rightarrow [x_i(1) \mid x_i(2) \mid \dots \mid x_i(k)]$$



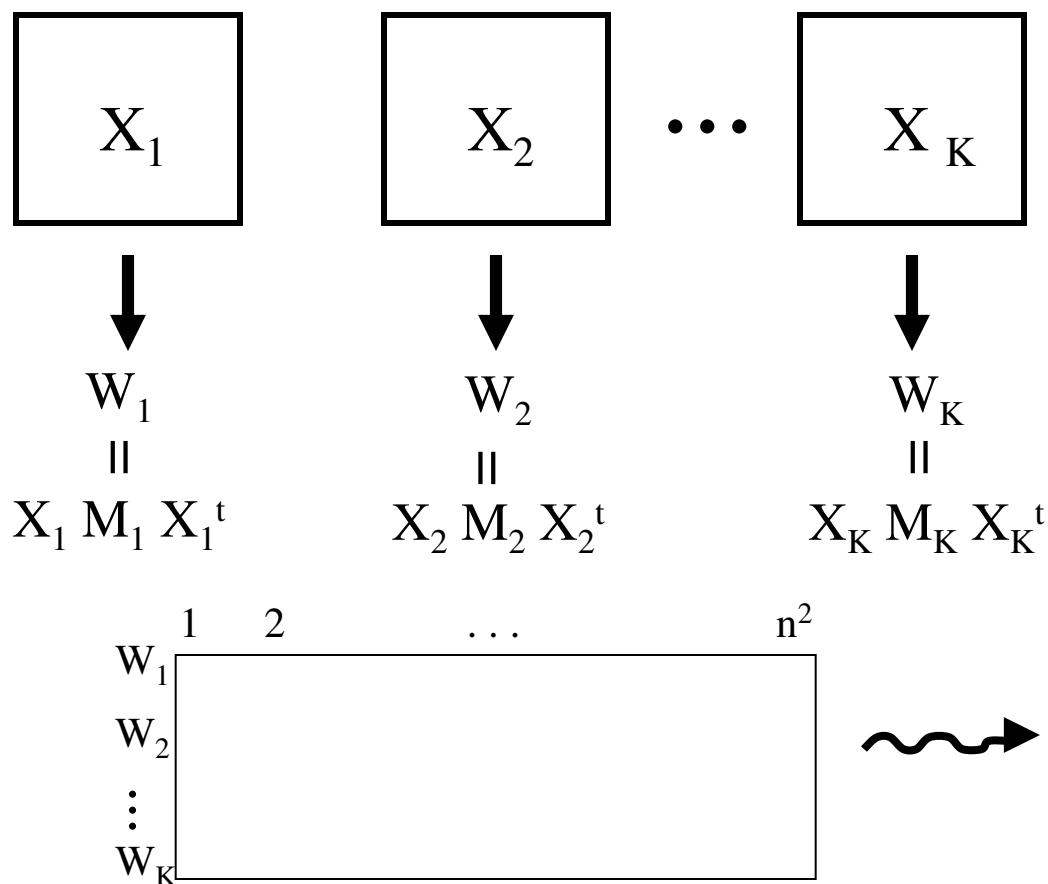


STATIS: LA INTERESTRUCTURA

CIMPA-UCR

Objetivo:

Comparar globalmente las tablas





CIMPA-UCR

STATIS: LA INTERESTRUCTURA

Métrica:

Producto interno de Hilbert-Schmidt:

$$\langle W_k, W_{k'} \rangle = \text{Tr} (DW_k DW_{k'})$$

- norma
- distancia

Estudios normalizados: $\frac{W_k}{\|W_k\|_{HS}}$

Coeficiente RV de Robert - Escoufier

$$RV(k, k') = \left\langle \frac{W_k}{\|W_k\|_{HS}}, \frac{W_{k'}}{\|W_{k'}\|_{HS}} \right\rangle_{HS}$$

STATIS: LA INTERESTRUCTURA (2)

Propiedades de RV:

$$d_{HS} \left(\frac{W_k}{\|W_k\|}, \frac{W_{k'}}{\|W_{k'}\|} \right) = \sqrt{2 [1 - RV(k, k')]}$$

$$\text{Si } RV(k, k') = 1 \Rightarrow \frac{W_k}{\|W_k\|} = \frac{W_{k'}}{\|W_{k'}\|}$$

$\Rightarrow k, k'$ son estudios (ACP) equivalentes

Si $RV(k, k') = 0 \Rightarrow$ Las variables de X_k tienen covarianzas nulas con las de $X_{k'}$

STATIS: LA INTERESTRUCTURA (2)

Representación Euclídea:

S : matriz de productos escalares entre tablas

$\Delta = \text{diag} (\pi_k)$ ponderación de las tablas

$\Rightarrow$ Diagonalizar $S \Delta$

Obtenemos:

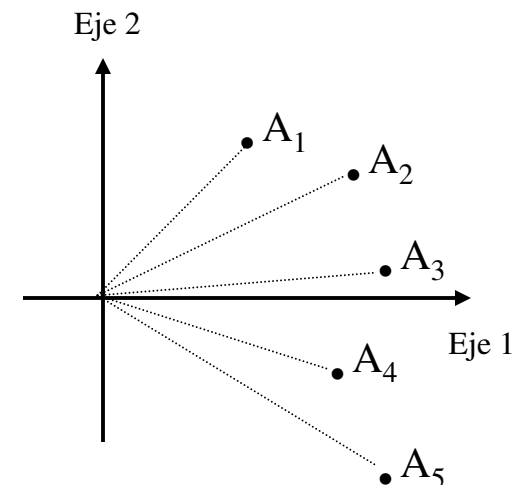
$$S \Delta z = \lambda z$$

puntos $A_1, \dots, A_k$ que representan a $W_1, \dots, W_k$

$\rightsquigarrow$ Coordenadas de A_k por $\sqrt{\lambda} z$

$$RV(k, k') = \cos (DA_k, DA_{k'})$$

Tablas complementarias con peso nulo





STATIS: COMPROMISO

Objetivo:

Encontrar una nube que sea “término medio” de todas las K nubes

$$W = \sum_{k=1}^K \alpha_k \frac{W_k}{\|W_k\|_{HS}}$$

Criterios:

- W tiene correlación Máxima (HS) con los W_k
- $\|W\|_{HS} = 1$

Sea $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix}$ el vector propio asociado al mayor valor propio

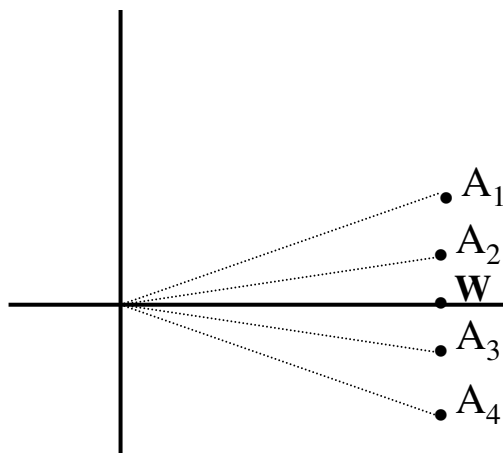
$$\lambda_1, \text{ de } S\Delta \text{ (se toma con } z : >0 \text{) } \Rightarrow \alpha_k = \frac{1}{\sqrt{\lambda_1}} \pi_k z_k$$

$$\text{así } W = \sum_{k=1}^K \frac{\pi_k z_k}{\sqrt{\lambda_1}} \frac{W_k}{\|W_k\|_{HS}}$$

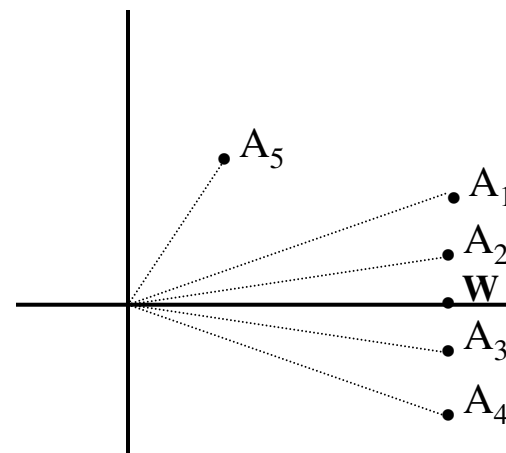


CIMPA-UCR

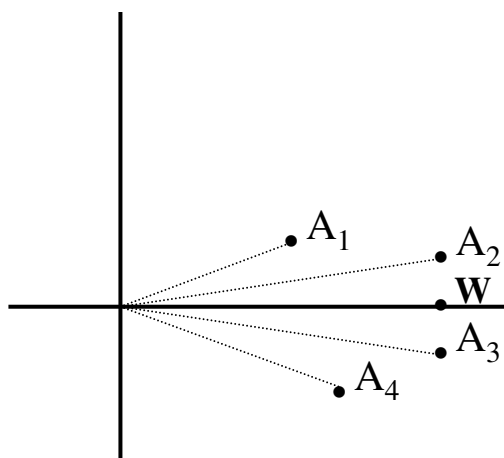
STATIS: COMPROMISO



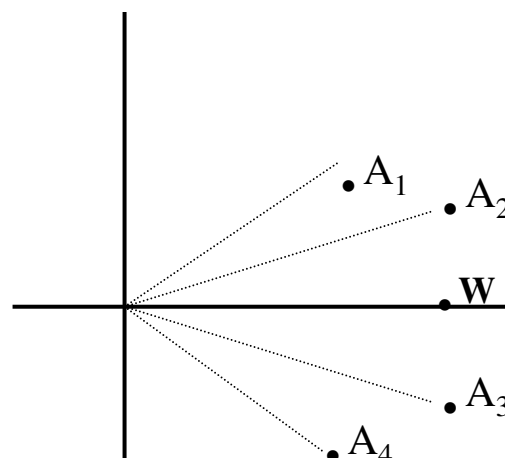
El compromiso refleja la estructura común



X_5 posee una estructura diferente
W es resumen mayoritario



W no es buen resumen (normas $\neq$)



No hay estructura común



CIMPA-UCR

STATIS: LA INTRAESTRUCTURA

Objetivos:

- Representar los n puntos por el conjunto de K tablas posición media
 - Representar las correlaciones de las variables con los ejes del compromiso
- Permite interpretar ejes e individuos

Individuos:

- W : Matriz $n \times n$
- ACP de la nube de individuos cuyos productos escalares están en W
- Diagonalizar WD
- ~> Componentes principales $C^1, C^2, \dots, C^n$
Valores propios $\mu_1, \mu_2, \dots, \mu_n$ } → individuos



CIMPA-UCR

STATIS: LA INTRAESTRUCTURA

Calidad de Representación: $\frac{\mu_1 + \mu_2 + \dots + \mu_n}{\sum_{i=1}^n \mu_i}$ % Inercia

Prop: distancia $d^2(i, i') = \sum \alpha_k \|i - i'\|_{HS}^2$: distancia promedio

Variables:

Para toda variable original $x^j(k)$, se calcula la correlación: $\rho(x^j(k), C^l)$

→ Coordenadas en el círculo de correlaciones

Obs: habría sido equivalente hacer el ACP de la tabla
 $[\sqrt{\alpha_1} X_1 \mid \sqrt{\alpha_2} X_2 \mid \dots \mid \sqrt{\alpha_K} X_K]$

STATIS: TRAYECTORIAS DE LOS INDIVIDUOS (Caso en el que la k es un índice de tiempo)

Objetivo:

Representar en los ejes del compromiso los individuos descritos por cada tabla.

Se proyecta cada fila $[0 \parallel x_i(k) \parallel 0]$ como elemento suplementario del ACP de WD

$$\frac{1}{\sqrt{\mu_j}} W_k DC^j : \text{da las coordenadas de los } x_i(k)$$

Propiedades:

Coordenada de i en el compromiso es el baricentro de todos los i ponderados por x_i

Trayectorias permiten descubrir individuos responsables de las diferencias entre tablas X_k $X_{k'}$



STATIS: TRAYECTORIAS DE LOS INDIVIDUOS *(Caso en el que la k es un índice de tiempo)*

Formas de las trayectorias:

- *Trayectoria poco extendida, gira alrededor de sí misma:* Evolución sigue la media (misma diferencia con la media de tabla a tabla).
- *Trayectoria de gran amplitud:* cambio de estructura del individuo a lo largo de las tablas ($\neq$ evolución media)

Obs: *se pueden graficar individuos suplementarios.*

Si un individuo aparece sólo en algunas tablas, se pone

Como suplementario (su compromiso no existe)



Statis: Proyecto Hidroeléctrico Ventas-Garita

*Con el objetivo de construir una planta Hidroeléctrica en Costa Rica se utilizarán las aguas de tres ríos. Para analizar la calidad del agua y la agresividad de esta a los diferentes materiales necesarios para la construcción de la represa, se definieron cinco puntos de muestreo llamados: Reventazón1 (**R1**), Reventazón3 (**R3**), Tuis4 (**T4**), Tuis5 (**T5**) y Turr2 (**T2**). Durante cuatro años (Marzo de 1995-Diciembre de 1998) se realizaron muestreos mensuales para obtener mediciones de 19 variables, las cuales son: **Ph**, **Temp** (Temperatura), **Na**, **K**, **Ca**, **Mg**, **SiO₂**, **OD**, **DBO**, **SD** (sólidos sedimentables), **ST** (sólidos totales), **PO₄**, **Cl**, **NO₃**, **SO₄**, **HCO₃**, **DT** (dureza total), **POD** (porcentaje de oxígeno disuelto), y **CAL**. **CAL** es un índice de calidad del agua expresado con un número entre 0 y 100, obtenido como una función (no lineal) de algunas de las variables anteriores.*



CIMPA-UCR

Análisis de Tablas Múltiples

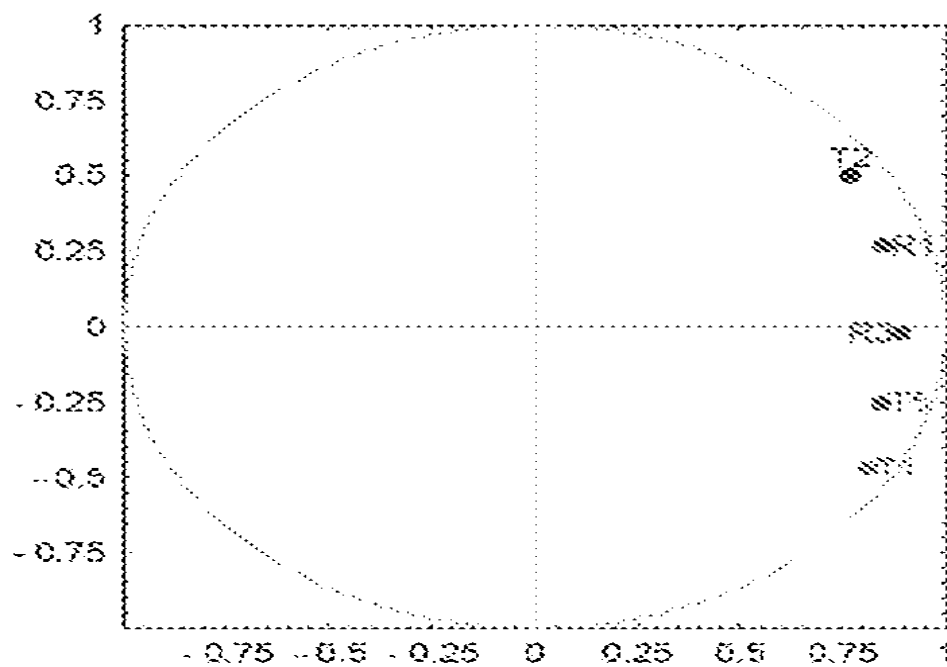


Figura 8.1: Interestructura.



CIMPA-UCR

Variables de Opinión

- *El seguro social debe privatizarse (SP).*
- *Se debe propiciar la libre elección médica (LM).*
- *La democracia es el sufragio (DS).*
- *Se debe defender la banca nacionalizada (BN).*
- *Se debe confiar en la justicia (CJ).*
- *Se debe aumentar el presupuesto en seguridad (PS).*
- *Ya no nos sentimos seguros (IS).*
- *Los enfermos de sida deben aislarse (SD).*



CIMPA-UCR

Variables de Señalización

- Sexo: Hombres (H) y Mujeres (M).
- Escolaridad: Ninguna (EN), Primaria (EP), Secundaria (ES), Técnica (ET), Universitaria incompleta (EUI) y Universitaria completa (EUC).
- Profesión: Agricultor (Bag), Patrón o empleado de comercio (Bec), Profesión liberal (Bpl), Empleado (Bem), Profesión intermedia (Bpi), maestro-profesor (Bmp), pensionado (Bpen), Ama de casa (Bac) y Desempleado (Bdes)
- Partido político: Partido Liberación Nacional (PLN), Partido Unidad Social Cristiana (PUSC), Izquierda (Izq), Otros partidos (Potr) y Ningún partido (Pnin).
- Ingreso Familiar: Ingresos Bajos (IB), Ingresos Medios (IM) e Ingresos Altos (IA).
- Provincia: San José (Sjos), Alajuela (Ala), Cartago (Car), Guanacaste (Gua), Puntarenas (Aren), Limón (Lim) y Heredia (Her).
- Edad, codificada en las siguientes clases: de 18 a 25 (A18), de 25 a 35 (A25), de 35 a 45 (A35), de 45 a 55 (A45) y más de 55 (A55).



Evolución de la Opinión Pública

Los datos de base de este ejemplo corresponden a las encuestas de opinión pública de los años 1988, 1989, 1990 y 1991 realizadas por el proyecto de investigación Estructuras de la Opinión Pública de la Escuela de Matemática de la Universidad de Costa Rica. Estas encuestas son anuales sobre una muestra de unos 1000 individuos, mayores de 18 años, preservando cuotas por sexo, edad y ocupación. Los entrevistados se pronuncian sobre unas 100 afirmaciones relativas a los principales conflictos sociales ventilados por la prensa local [64], [34].



CIMPA-UCR

El Compromiso

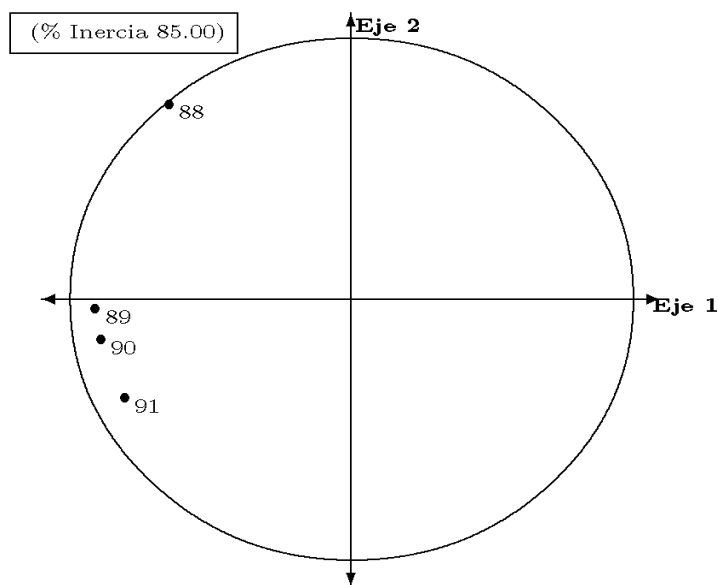


Figura 8.11: Interestructura para la evolución de la opinión.



CIMPA-UCR

Círculo de Correlaciones

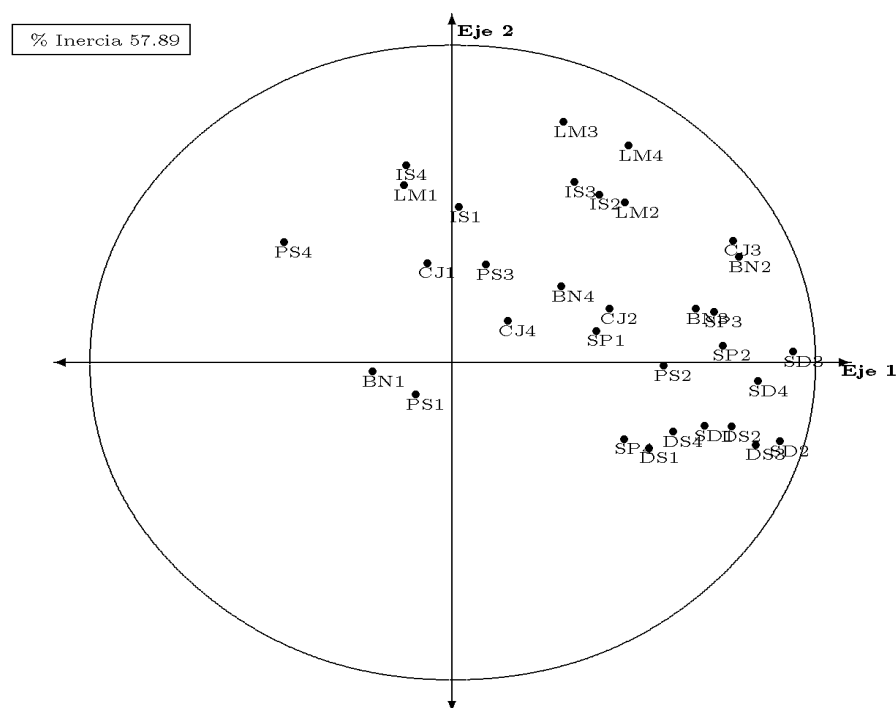


Figura 8.12: Correlación entre variables y los ejes del compromiso para la evolución de la opinión.



Grupos Sociales

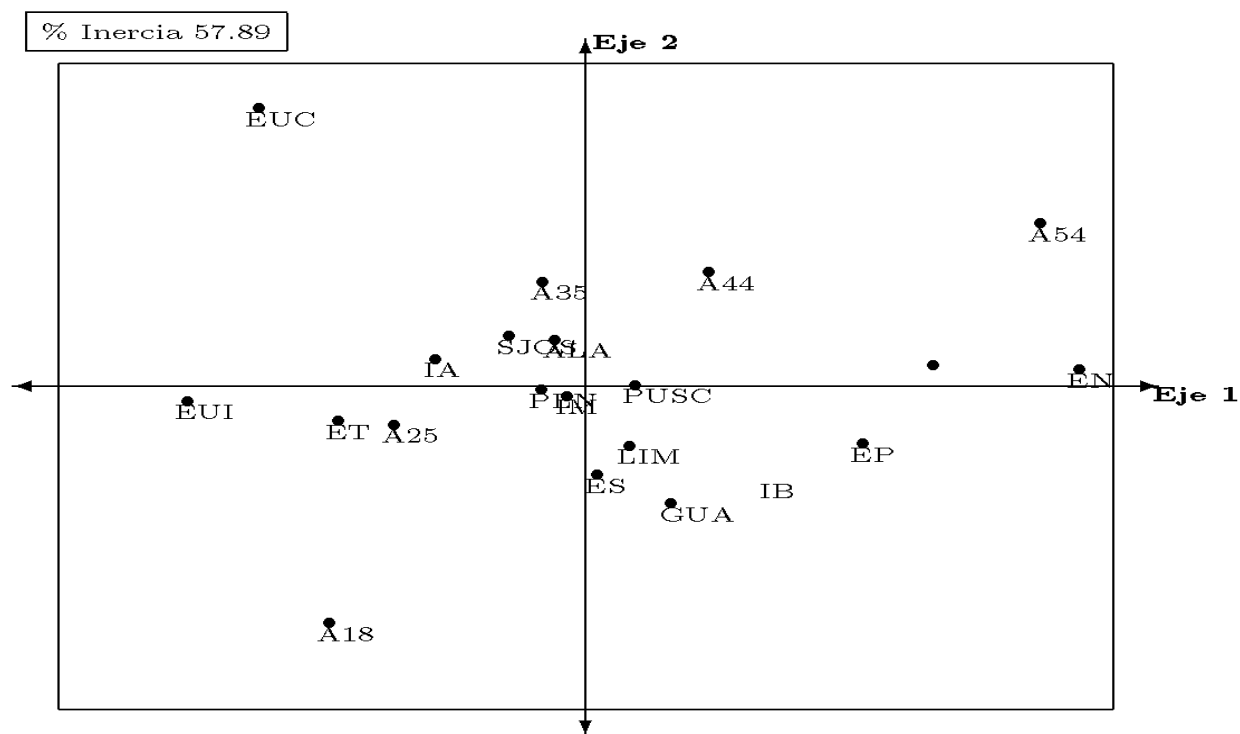
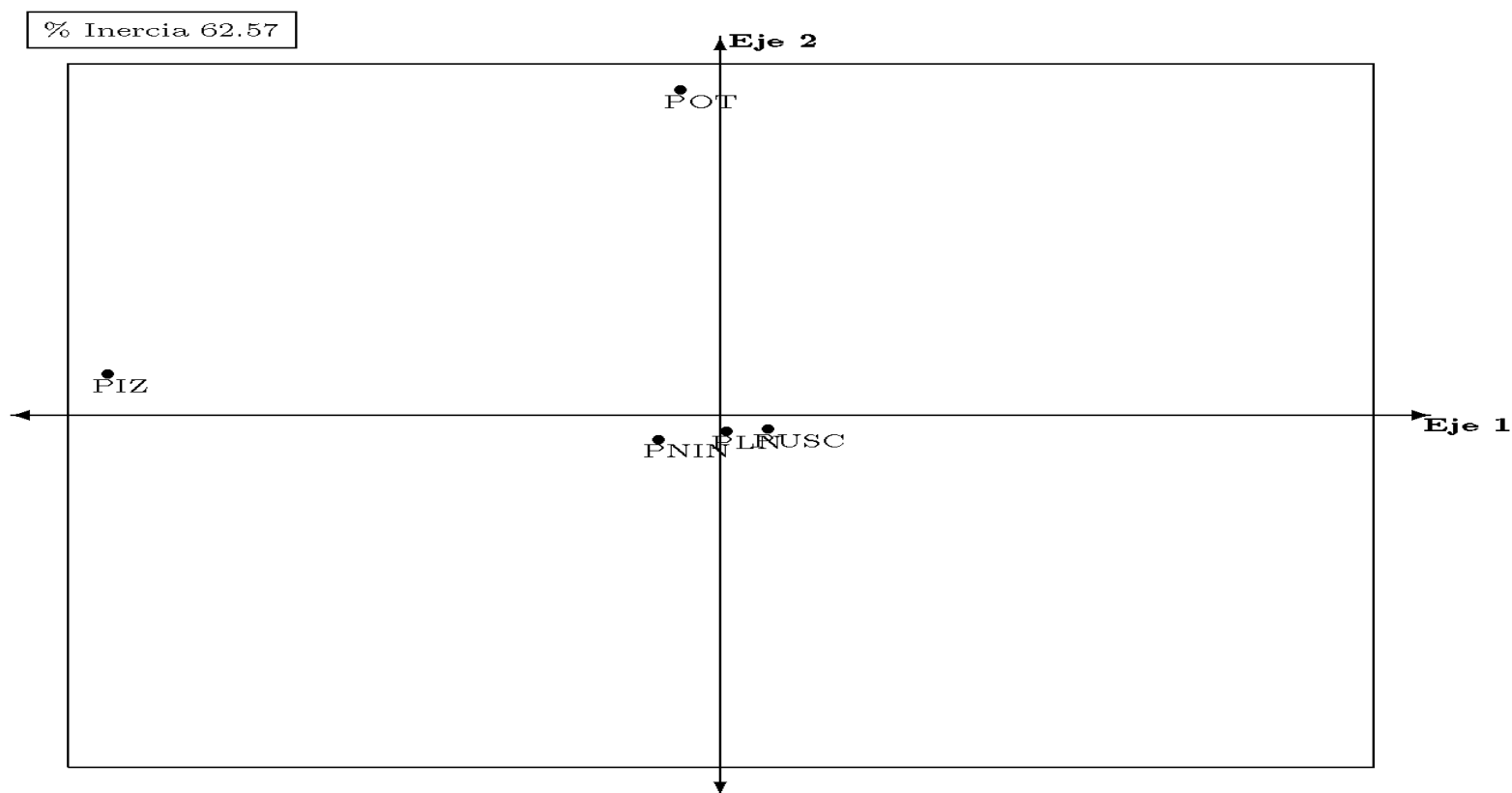


Figura 8.13: Grupos sociales promedios: Escolaridad, Edad, Provincia.



CIMPA-UCR

Simpatía Política





CIMPA-UCR

Grupos de Edad

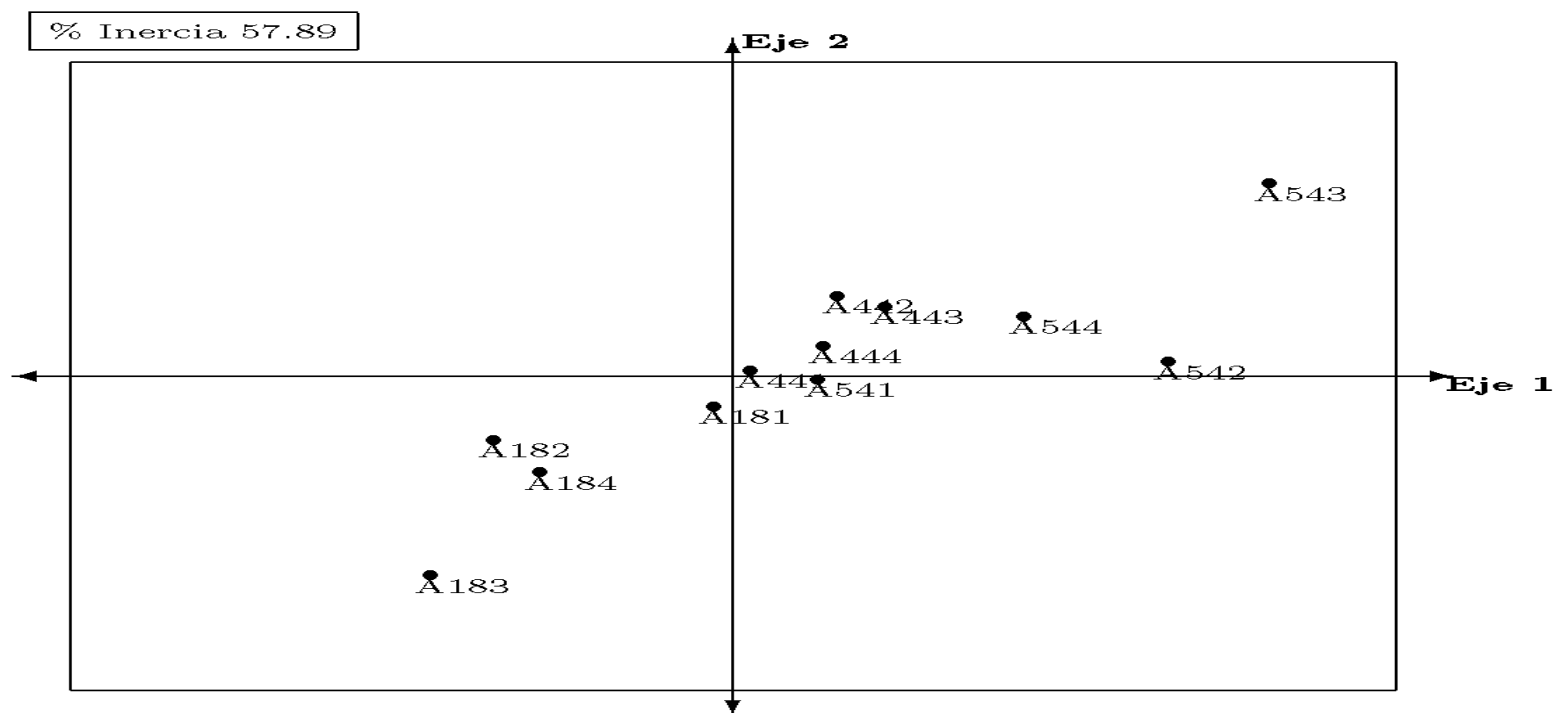


Figura 8.15: Trayectorias: grupos de edad.



CIMPA-UCR

Escolaridad

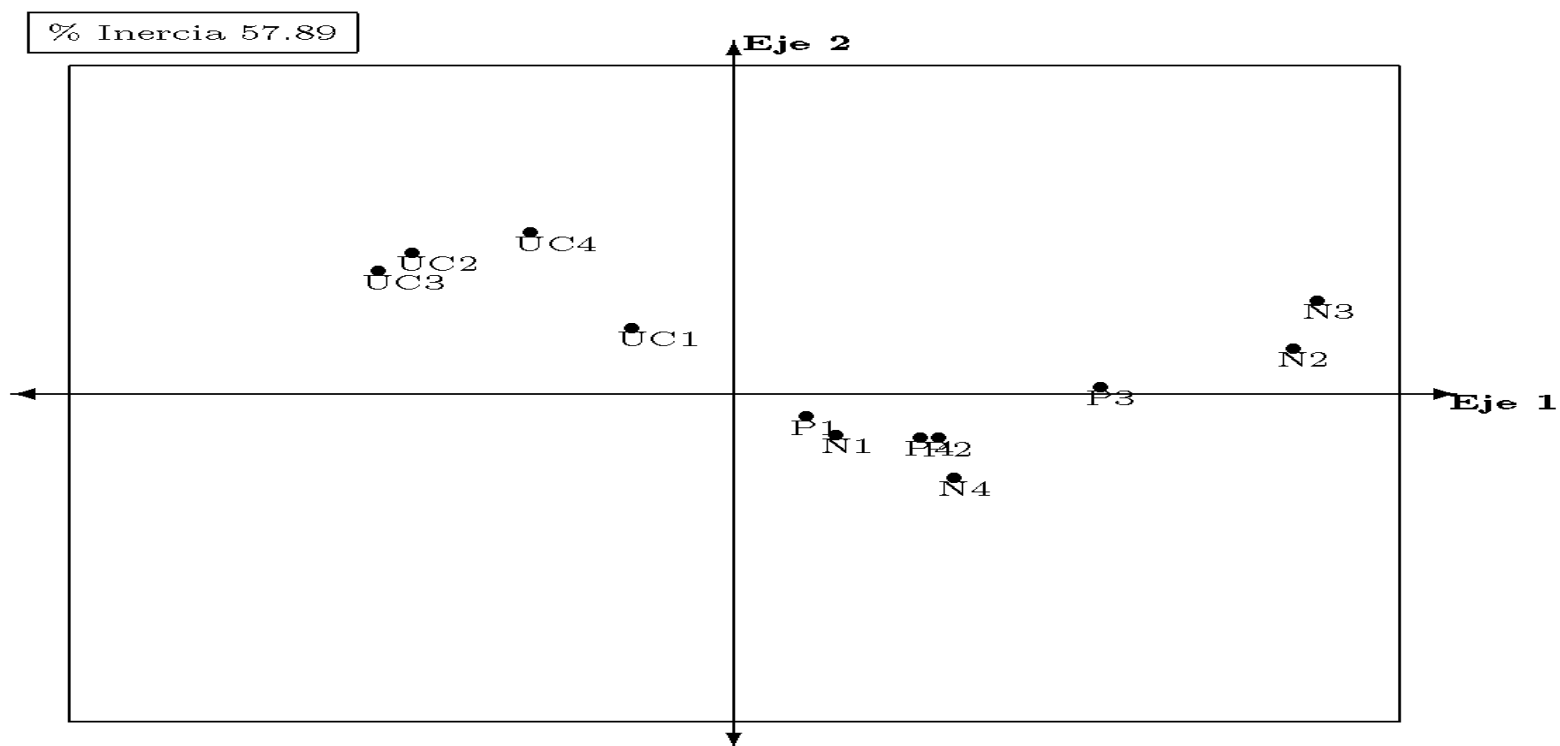


Figura 8.16: Trayectorias: escolaridad.



Análisis Factorial Múltiple (AFM)

CIMPA-UCR

	1	p_1	1	p_2		1	p_q
1	x_1 x_2 $\dots$ x_q						
n							

B. Escofier &
J. Pagès

$$P = \sum_{k=1}^q p_k : \# \text{ total de variables}$$

$$N_K = (x_k, M_k, D) \quad \text{Nube para cada grupo de variables}$$

$$M_K : \textit{diagonal}$$

$$N = (X, M, D) \quad \text{Nube total}$$

$$x = [x_1 : x_2 : \dots : x_q] \quad \text{Tabla concatenada}$$



CIMPA-UCR

AFM: Objetivos

Interestructura

Comparación global de los grupos de variables

Intraestructura:

Comparación de las nubes N_K que representan a los n individuos

→ mismo sistema de coordenadas

→ compromiso

Idea de base:

ponderación de los grupos de variables



CIMPA-UCR

AFM: Ponderación

$\lambda_1(k)$: Varianza de la primera componente principal del ACP de $N_k = (X_k, M_k, D)$ $\lambda_1(k) = I_{\Delta_{\frac{1}{u}}}(N_k)$

$\frac{1}{\sqrt{\lambda_1(k)}}$: Multiplica a cada variable de X_k

Ponderación $\rightarrow$ equilibrar influencia de los grupos X_k

Si se hace ACP de X , influyen

- P_k : # de variables por tabla (p_k : + influencia)
- estructura de x_k (+ correlación : + influencia)