Capítulo 4

Análisis de Correspondencias Simple

4.1. INTRODUCCIÓN

El Análisis de Correspondencias Simple permite describir las relaciones entre dos variables categóricas dispuestas en una tabla de contingencia.

Variables categóricas:

$$\mathcal{I} \in \{1, 2, \ldots, n\}, \quad \mathcal{J} \in \{1, 2, \ldots, p\}\,.$$

Distribución conjunta:

$(\mathcal{I},\ \mathcal{J})$	1	2	$\cdots$	p	Total
1	π_{11}	π_{12}	$\cdots$	π_{1p}	π_{1+}
2	π_{21}	π_{22}	$\cdots$	π_{2p}	π_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	π_{n1}	π_{n2}	$\cdots$	π_{np}	π_{n+}
Total	π_{+1}	π_{+2}	$\cdots$	π_{+p}	1

Distribuciones condicionadas:

$\mathcal{J} \ \mid \mathcal{I}$	1	2	$\cdots$	p	Total
1	π_{11}/π_{1+}	π_{12}/π_{1+}	$\cdots$	π_{1p}/π_{1+}	1
2	π_{21}/π_{2+}	π_{22}/π_{2+}	$\cdots$	π_{2p}/π_{2+}	1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	π_{n1}/π_{n+}	π_{n2}/π_{n+}	$\cdots$	π_{np}/π_{n+}	1
Total	π_{+1}	π_{+2}	$\cdots$	π_{+p}	1

$\mathcal{I} \quad \vert \mathcal{J}$	1	2	$\cdots$	p	Total
1	π_{11}/π_{+1}	π_{12}/π_{+2}	$\cdots$	π_{1p}/π_{+p}	π_{1+}
2	π_{21}/π_{+1}	π_{22}/π_{+2}	$\cdots$	π_{2p}/π_{+p}	π_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	π_{n1}/π_{+1}	π_{n2}/π_{+2}	$\cdots$	π_{np}/π_{+p}	π_{n+}
Total	1	1	$\cdots$	1	1

Tabla de contingencia: Se toma una muestra aleatoria simple de m individuos y se apuntan sus valores de las variables $\mathcal{I}$ y $\mathcal{J}$. La tabla de contingencia se obtiene contando el número de individuos en cada combinación de categorías de $\mathcal{I}$ y de $\mathcal{J}$.

$(\mathcal{I},\ \mathcal{J})$	1	2	$\cdots$	p	Total
1	k_{11}	k_{12}	$\cdots$	k_{1p}	k_{1+}
2	k_{21}	k_{22}	$\cdots$	k_{2p}	k_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	k_{n1}	k_{n2}	$\cdots$	k_{np}	k_{n+}
Total	k_{+1}	k_{+2}	$\cdots$	k_{+p}	$k_{++} = m$

k_{ij}: Número de individuos con $\mathcal{I} = i$ y $\mathcal{J} = j$.

Tabla de frecuencias relativas: Representa la distribución conjunta muestral de las variables $\mathcal{I}$ y $\mathcal{J}$, y las distribuciones marginales.

$(\mathcal{I},\ \mathcal{J})$	1	2	$\cdots$	p	Total
1	f_{11}	f_{12}	$\cdots$	f_{1p}	f_{1+}
2	f_{21}	f_{22}	$\cdots$	f_{2p}	f_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	f_{n1}	f_{n2}	$\cdots$	f_{np}	f_{n+}
Total	f_{+1}	f_{+2}	$\cdots$	f_{+p}	1

donde

$$f_{ij} = \frac{k_{ij}}{k_{++}}, \quad f_{i+} = \frac{k_{i+}}{k_{++}}, \quad f_{+j} = \frac{k_{+j}}{k_{++}}\,.$$

Ejemplo 4.1 Se dispone de datos sobre el consumo de cuatro marcas en tres segmentos de consumidores. En la tabla siguiente tenemos el número de individuos que compran habitualmente la marca i y que pertenecen al segmento j. Se desea estudiar qué relaciones hay entre las variables “Marca” y “Segmento”.

Tabla de contingencia:

	Segmento			
Marca	1	2	3	Total
A	30	30	155	215
B	30	130	30	190
C	80	30	30	140
D	80	30	5	115
Total	220	220	220	660

Tabla de frecuencias relativas:

	1	2	3	Total
A	0.045	0.045	0.235	0.326
B	0.045	0.197	0.045	0.288
C	0.121	0.045	0.045	0.212
D	0.121	0.045	0.008	0.174
Total	0.333	0.333	0.333	1.000

Tabla de perfiles fila: Representa las distribuciones muestrales de la variable $\mathcal{J}$ condicionadas a cada categoría de la variable $\mathcal{I}$; es decir, las distribuciones de $\mathcal{J}|\mathcal{I} = i$, $i = 1, \ldots, n$.

$\mathcal{J} \mid \mathcal{I}$	1	2	$\cdots$	p	Total
1	f_{11}/f_{1+}	f_{12}/f_{1+}	$\cdots$	f_{1p}/f_{1+}	1
2	f_{21}/f_{2+}	f_{22}/f_{2+}	$\cdots$	f_{2p}/f_{2+}	1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	f_{n1}/f_{n+}	f_{n2}/f_{n+}	$\cdots$	f_{np}/f_{n+}	1
	f_{+1}	f_{+2}	$\cdots$	f_{+p}	1

Esta tabla se puede obtener a partir de la tabla de frecuencias relativas o de la tabla de contingencia inicial:

$$\frac{f_{ij}}{f_{i+}} = \frac{\dfrac{k_{ij}}{k_{++}}}{\dfrac{k_{i+}}{k_{++}}} = \frac{k_{ij}}{k_{i+}} \, .$$

Tabla de perfiles columna: Representa las distribuciones muestrales de la variable $\mathcal{I}$ condicionadas a cada categoría de la variable $\mathcal{J}$; es decir, las distribuciones de $\mathcal{I}|\mathcal{J} = j$, $j = 1, \ldots, p$.

$\mathcal{I} \mid \mathcal{J}$	1	2	$\cdots$	p	Total
1	f_{11}/f_{+1}	f_{12}/f_{+2}	$\cdots$	f_{1p}/f_{+p}	f_{1+}
2	f_{21}/f_{+1}	f_{22}/f_{+2}	$\cdots$	f_{2p}/f_{+p}	f_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	f_{n1}/f_{+1}	f_{n2}/f_{+2}	$\cdots$	f_{np}/f_{+p}	f_{n+}
	1	1	$\cdots$	1	1

Ejemplo 4.2 Para la tabla de contingencia dada en el Ejemplo 4.1, contestar a las preguntas:

- De entre los individuos del Segmento 1, ¿qué proporción compran la Marca A?
- De entre los individuos que compran la Marca A, ¿qué proporción están en el Segmento 1?

Para contestar necesitamos obtener las tablas de perfiles fila y de perfiles columna.

Tabla de perfiles fila:

	1	2	3	Total
A	0.140	0.140	0.721	1.000
B	0.158	0.684	0.158	1.000
C	0.571	0.214	0.214	1.000
D	0.696	0.261	0.043	1.000
Total	0.333	0.333	0.333	1.000

Tabla de perfiles columna:

	1	2	3	Total
A	0.136	0.136	0.705	0.326
B	0.136	0.591	0.136	0.288
C	0.364	0.136	0.136	0.212
D	0.364	0.136	0.023	0.174
Total	1.000	1.000	1.000	1.000

4.2. TEST $\mathcal{X}^2$ DE INDEPENDENCIA

Supongamos que deseamos contrastar

H_0: $\mathcal{I}$ y $\mathcal{J}$ son independientes
H_1: $\mathcal{I}$ y $\mathcal{J}$ son dependientes.

Ejemplo 4.3 La variable $\mathcal{I}$=Marca$\in\{$A,B,C,D$\}$ es independiente de $\mathcal{J}$=Segmento$\in\{1,2,3\}$, si:

- la proporción de personas que compra la marca A es la misma dentro de los tres segmentos.
- Idem para las marcas B, C y D.

Es decir, debe ocurrir ...

En general, la variable $\mathcal{I}$ es independiente de $\mathcal{J}$ si la proporción de individuos que pertenecen a una categoría i de $\mathcal{I}$ es la misma dentro de cualquier categoría j de $\mathcal{J}$, es decir, si ...

En forma más resumida, $\mathcal{I}$ es independiente de $\mathcal{J}$ si

$$\frac{\pi_{ij}}{\pi_{+j}} = \pi_{i+}, \quad i \in \{1,2,\ldots,n\},\ j \in \{1,2,\ldots,p\}\,,$$

o lo que es lo mismo,

$$\pi_{ij} = \pi_{i+}\pi_{+j}, \quad i \in \{1,2,\ldots,n\},\ \{1,2,\ldots,p\}\,.$$

Estadístico $\mathcal{X}^2$: El estadístico $\mathcal{X}^2$ mide las distancias entre las frecuencias observadas f_{ij} y las que deberíamos observar si las variables $\mathcal{I}$ y $\mathcal{J}$ fuesen independientes, $f_{i+}f_{+j}$,

$$\mathcal{X}^2 = k_{++}\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{(f_{ij}-f_{i+}f_{+j})^2}{f_{i+}f_{+j}}\,.$$

En términos de las frecuencias absolutas, el estadístico $\mathcal{X}^2$ es

$$\mathcal{X}^2 = \sum_{i=1}^{n}\sum_{j=1}^{p}\frac{\left(k_{ij}-\dfrac{k_{i+}k_{+j}}{k_{++}}\right)^2}{\dfrac{k_{i+}k_{+j}}{k_{++}}}\,.$$

Teorema 4.1 (Distribución del estadístico $\mathcal{X}^2$ de independencia)

Si $\mathcal{I}$ y $\mathcal{J}$ son independientes, entonces $\quad \mathcal{X}^2 \sim \mathcal{X}^2_{(n-1)(p-1)}$.

Por el Teorema anterior, podemos tomar un valor extremo de la distribución; es decir, un valor $\mathcal{X}^2_{(n-1)(p-1),\alpha}$ que deja a su derecha una probabilidad α pequeña, y se rechaza la hipótesis de independencia al nivel de significación α si el valor obtenido del estadístico supera este valor

$$\mathcal{X}^2 > \mathcal{X}^2_{(n-1)(p-1),\alpha} \Leftrightarrow \text{p-valor} = P\left(Y > \mathcal{X}^2\right) < \alpha\,,$$

donde Y es una variable aleatoria con distribución $\mathcal{X}^2_{(n-1)(p-1)}$. Normalmente se toma $\alpha = 0{,}05$.

Ejemplo 4.4 Considera la tabla de contingencia del Ejemplo 4.1. Contrastamos la hipótesis H_0 frente a H_1:

$$\begin{aligned} H_0 &: \ \textit{Marca y Segmento son independientes} \\ H_1 &: \ \textit{Marca y Segmento son dependientes} \end{aligned}$$

$\mathcal{X}^2 = ...$

Sea $Y \sim \mathcal{X}^2_{3\times 2}$. El p-valor se puede calcular, por ejemplo, mediante Excel:

$$\text{p-valor} = P(Y > 362{,}413) = 3{,}33 \times 10^{-75} < 0{,}05\,.$$

Se rechaza la hipótesis nula de independencia. Esto significa que:

- bien, existen al menos dos segmentos en los que las proporciones de consumidores de las marcas es diferente;
- bien, existen al menos dos marcas en las que la proporción de individuos en los segmentos es diferente.

4.3. NUBES DE PUNTOS FILA Y COLUMNA

Puntos fila: Las filas de la tabla de perfiles fila son n puntos con p coordenadas, y los llamaremos *puntos fila*:

$$R_i = (r_{i1}, r_{i2}, \ldots, r_{ip}) = \left(\frac{\quad}{\quad}, \frac{\quad}{\quad}, \ldots, \frac{\quad}{\quad} \right), \quad i = 1, \ldots, n.$$

Cada punto fila va a tener asociado un *peso* (llamado también *ponderación*), dado por su frecuencia relativa marginal.

Nube de puntos fila: La *nube de puntos fila* es el conjunto de puntos fila asociados a sus correspondientes pesos, es decir,

$$N(\mathcal{I}) = \left\{ (R_i, f_{i+}),\ i = 1, 2, \ldots, n \right\}.$$

Centro de gravedad de la nube de puntos fila: El *baricentro* o *centro de gravedad* de los puntos fila es una media ponderada de los n puntos:

$$\overline{R} = (\overline{r}_1, \ldots, \overline{r}_p),$$

donde

$$\overline{r}_j = \sum_{i=1}^{n} f_{i+} r_{ij}, \quad j = 1, \ldots, p.$$

El centro de gravedad coincide con la última fila de la tabla de perfiles fila, es decir

$$\overline{R} = (f_{+1}, \ldots, f_{+p}).$$

Efectivamente,

$$\overline{r}_j = \sum_{i=1}^{n} f_{i+} \frac{f_{ij}}{f_{i+}} = \sum_{i=1}^{n} f_{ij} = f_{+j},\ j = 1, 2, \ldots, p.$$

Puntos columna: Cada perfil columna es un punto con n coordenadas y se denomina *punto columna*,

$$C_j = (c_{1j}, c_{2j}, \ldots, c_{nj}) = \left(\text{---}, \text{---}, \ldots, \text{---} \right), \; j = 1, 2, \ldots, p\,.$$

Nube de puntos columna: La *nube de puntos columna* es el conjunto de puntos columna con sus respectivos pesos,

$$N(\mathcal{J}) = \{(C_j, f_{+j}), j = 1, 2, \ldots, p\}\,.$$

Centro de gravedad de la nube de puntos columna: El baricentro o centro de gravedad de la nube de puntos columna es una media ponderada de los p puntos columna:

$$\overline{C} = (\overline{c}_1, \ldots, \overline{c}_n)\,,$$

donde

$$\overline{c}_i = \sum_{j=1}^{p} f_{+j} c_{ij}, \quad i = 1, 2, \ldots, n\,.$$

El centro de gravedad coincide con la última columna de la tabla de perfiles columna, es decir,

$$\overline{C} = (f_{1+}, \ldots, f_{n+})\,.$$

Efectivamente,

$$\overline{c}_i = \sum_{j=1}^{p} f_{+j} \frac{f_{ij}}{f_{+j}} = \sum_{j=1}^{p} f_{ij} = f_{i+}, \quad i = 1, 2, \ldots, n\,.$$

4.4. INERCIA DE LAS NUBES DE PUNTOS

Distancia $\mathcal{X}^2$ entre dos puntos fila R_i y $R_{i'}$:

$$d_{\mathcal{X}}(R_i, R_{i'}) = \sum_{j=1}^{p} \frac{1}{f_{+j}} (r_{ij} - r_{i'j})^2 = \ldots$$

Distancia $\mathcal{X}^2$ entre dos puntos columna C_j y C_j':

$$d_{\mathcal{X}}(C_j, C_{j'}) = \sum_{i=1}^{n} \frac{1}{f_{i+}}(c_{ij} - c_{ij'})^2 = ...$$

Ejemplo 4.5 Para la tabla de contingencia del Ejemplo 4.1, calcular la distancia $\mathcal{X}^2$ entre los puntos fila R_1 y R_2, y entre los puntos columna C_1 y C_2.

Distancia $\mathcal{X}^2$ entre R_1 y R_2:

$d_{\mathcal{X}}(R_1, R_2) = ...$.

Distancia $\mathcal{X}^2$ entre C_1 y C_2:

$d_{\mathcal{X}}(C_1, C_2) = ...$.

Inercia de un punto fila: La *inercia* del punto fila R_i es el producto de su peso por la distancia $\mathcal{X}^2$ de dicho punto al centro de gravedad,

$$I(R_i) = f_{i+} d_{\mathcal{X}}(R_i, \overline{R}) \,.$$

Inercia de un punto columna: La *inercia* del punto columna C_j es

$$I(C_j) = f_{+j} d_{\mathcal{X}}(C_j, \overline{C}) \,.$$

Ejemplo 4.6 Para la tabla de contingencia del Ejemplo 4.1, obtener las inercias de cada punto fila.

Distancias al centro de gravedad: ...

Inercias de los puntos fila: ...

Inercia de la nube de puntos fila: La *inercia de la nube de puntos fila* $N(\mathcal{I})$ es la suma de las inercias de los puntos fila,

$$I\left[N(\mathcal{I})\right] = \sum_{i=1}^{n} I(R_i) = \sum_{i=1}^{n} f_{i+} d_{\mathcal{X}}(R_i, \overline{R}) = ...$$

✓ La inercia de los puntos fila es una medida de la variabilidad de los puntos fila.

Inercia de la nube de puntos columna:

$$I\left[N(\mathcal{J})\right] = \sum_{j=1}^{p} I(C_j) = \sum_{j=1}^{p} f_{+j} d_{\mathcal{X}}(C_j, \overline{C}) \, .$$

Proposición 4.1 Se verifica

$$I\left[N(\mathcal{I})\right] = I\left[N(\mathcal{J})\right] = \frac{\mathcal{X}^2}{k_{++}}\,.$$

Prueba:...

✓ La inercia total de la nube de puntos fila coincide con la inercia de la nube de puntos columna.

✓ La cantidad de inercia o variabilidad de la nube de puntos fila (o de puntos columna) mide con el grado de dependencia entre las variables $\mathcal{I}$ y $\mathcal{J}$.

Ejemplo 4.7 Calcular la inercia total de la nube de puntos fila para la tabla de contingencia del Ejemplo 4.1. Comprobar que coincide con $\mathcal{X}^2/k_{++}$.

$$I[N(\mathcal{I})] = ... \qquad .$$

Por otro lado,

$$\frac{\mathcal{X}^2}{k_{++}} = ... \qquad .$$

4.5. ANÁLISIS DE LA NUBE DE PUNTOS FILA

Se desea representar los puntos fila en un espacio de menor dimensión, con ejes ortogonales, y de manera que los puntos proyectados en los nuevos ejes tengan la máxima inercia (o máxima $\mathcal{X}^2$, es decir, máxima dependencia entre las variables $\mathcal{I}$ y $\mathcal{J}$).

Como podemos observar, el problema es similar al Análisis de Componentes Principales, reemplazando la varianza por la inercia de los puntos fila.

Sin embargo, en el Análisis de Componentes Principales la distancia natural entre puntos es la distancia euclídea, mientras que aquí es la distancia $\mathcal{X}^2$.

4.5.1. TRANSFORMACIONES DE LA NUBE DE PUNTOS

Nube de puntos transformada: La siguiente transformación de la nube de puntos fila permitirá utilizar la distancia euclídea al cuadrado d_e^2 en lugar de la distancia $d_{\mathcal{X}}$.

$$N^*(\mathcal{I}) = \left\{(R_i^*, f_{i+}),\ i = 1, 2, \ldots, n\right\},$$

donde

$$R_i^* = (r_{i1}^*, \ldots, r_{ip}^*) = \left(\frac{}{}, \ldots, \frac{}{} \right) = \left(\frac{}{}, \ldots, \frac{}{} \right).$$

El centro de gravedad de $N^*(\mathcal{I})$ es

$$\overline{R}^* = \left(\sqrt{f_{+1}}, \ldots, \sqrt{f_{+p}} \right).$$

Ejemplo 4.8 Para la tabla de contingencia del Ejemplo 4.1, obtener la nube de puntos transformada, representando sus puntos fila en una tabla.

	1	2	3
A	0.242	0.242	1.249
B	0.273	1.185	0.273
C	0.990	0.371	0.371
D	1.205	0.452	0.075
Total	0.577	0.577	0.577

Proposición 4.2 La distancia euclídea al cuadrado entre los puntos fila transformados R_i^* e $R_{i'}^*$ es igual a la distancia $\mathcal{X}^2$ entre los puntos fila sin transformar R_i y $R_{i'}$.

Prueba:...

Por tanto, la inercia de la nube de puntos fila se puede escribir utilizando la distancia euclídea al cuadrado, de la forma

$$I[N(\mathcal{I})] = \sum_{i=1}^{n} f_{i+} d_e^2 \left(R_i^*, \overline{R}^* \right) .$$

Ejemplo 4.9 Para la tabla de contingencia dada en el Ejemplo 4.1, calcular la distancia euclídea al cuadrado entre los puntos fila transformados R_1^* y R_2^*. Compara el resultado con el obtenido en el Ejemplo 4.5.

...

Nube de puntos fila centrada: Se realiza una nueva transformación de la nube de puntos fila, restándoles su centro de gravedad,

$$N^{**}(\mathcal{I}) = \{(X_i, f_{i+}),\ i = 1, 2, \ldots, n\}\,,$$

donde los puntos son $X_i = R_i^* - \bar{R}^*$; es decir,

$$X_i = (x_{i1}, \ldots, x_{ip}) = \left(\qquad\qquad , \ldots , \qquad\qquad \right),\ i = 1, \ldots, p\,.$$

El centro de gravedad es ahora

$$\overline{X} = (0, \ldots, 0) = O\,.$$

La inercia total, en función de los puntos fila centrados, es:

$$I[N(\mathcal{I})] = \sum_{i=1}^{n} f_{i+} d_e^2(X_i, O)\,.$$

Llamaremos $\boldsymbol{X}$ a la matrix cuyas filas son los puntos fila centrados $X_1, \ldots, X_n$.

Ejemplo 4.10 Para la tabla de contingencia del Ejemplo 4.1, obtener la matriz $\boldsymbol{X} = (x_{ij})$ de puntos fila centrados

$$\boldsymbol{X} = \begin{pmatrix} -0{,}336 & -0{,}336 & 0{,}671 \\ -0{,}304 & 0{,}608 & -0{,}304 \\ 0{,}412 & -0{,}206 & -0{,}206 \\ 0{,}628 & -0{,}126 & -0{,}502 \end{pmatrix}$$

4.5.2. DETERMINACIÓN DE LOS EJES PRINCIPALES

La nube de puntos fila centrados consta de n puntos,

$$\begin{gathered} X_1 = (x_{11}, \ldots, x_{1p}) \\ \vdots \\ X_n = (x_{n1}, \ldots, x_{np}) \end{gathered}$$

Los puntos tienen p coordenadas, pero en realidad están en un espacio de dimensión $p-1$.

Buscamos otros $r < p-1$ ejes en los que representar los puntos.

Sea $\mathbf{a}_\alpha = (a_{1\alpha}, \ldots, a_{p\alpha})'$ el vector director de un nuevo eje. La coordenada de un punto X_i en el eje determinado por $\mathbf{a}_\alpha$ es:

$$g_{i\alpha} = \mathbf{x}_i'\mathbf{a}_\alpha = \ldots \qquad .$$

Inercia de un eje α: Es la suma de las inercias de los puntos proyectados en el eje α,

$$I_\alpha = \sum_{i=1}^{n} f_{i+} \left(g_{i\alpha}\right)^2 = \sum_{i=1}^{n} f_{i+} \left(\mathbf{x}_i'\mathbf{a}_\alpha\right)^2 .$$

Matriz de inercia: Consideremos la matriz de puntos fila centrados $\boldsymbol{X} = (x_{ij})$, y la matriz diagonal con los pesos de los puntos fila, $D_{\mathcal{I}} = \mathrm{diag}\{f_{1+}, \ldots, f_{n+}\}$. La *matriz de inercia de los puntos fila* es

$$\mathbf{T} = \boldsymbol{X}' D_{\mathcal{I}} \boldsymbol{X} .$$

El elemento (j,k) de esta matriz es

$$t_{jk} = \sum_{i=1}^{n} f_{i+}\, x_{ij}\, x_{ik} .$$

La matriz de inercia $\mathbf{T}$ es análoga a una matriz de varianzas-covarianzas de $\boldsymbol{X}$, ponderando las filas de $\boldsymbol{X}$ por sus correspondientes pesos.

Ejemplo 4.11 Obtener la matriz de inercia $\mathbf{T}$ para la tabla de contingencia del Ejemplo 4.1.

$$\mathbf{T} = \begin{pmatrix} 0{,}168 & -0{,}048 & 0{,}226 \\ -0{,}048 & 0{,}155 & 0{,}022 \\ 0{,}226 & 0{,}022 & 0{,}226 \end{pmatrix} .$$

La inercia del eje α se puede escribir en función de la matriz de inercia $\mathbf{T}$ de la forma

$$I_\alpha = \mathbf{a}'_\alpha \mathbf{T} \mathbf{a}_\alpha \,.$$

Primer eje principal de correspondencias: El *primer eje principal* es el eje con máxima inercia; es decir, el eje con vector director $\mathbf{a}_1$ que es solución del problema:

$$\begin{array}{ll} \text{Max} & I_1 = \mathbf{a}'_1 \mathbf{T} \mathbf{a}_1 \\ \text{s.a.} & \mathbf{a}'_1 \mathbf{a}_1 = 1 \,. \end{array} \tag{4.1}$$

Solución: La inercia máxima I_1 es el mayor valor propio λ_1 de la matriz de inercia $\mathbf{T}$, y la dirección del primer eje principal $\mathbf{a}_1$ es el vector propio unitario asociado a λ_1.

Eje principal α-ésimo de correspondencias: Es el eje ortogonal a los ejes $1, 2, \ldots, \alpha - 1$, con máxima inercia; es decir, es el eje cuyo vector director $\mathbf{a}_\alpha$ es solución del problema

$$\begin{array}{ll} \text{Max} & I_\alpha = \mathbf{a}'_\alpha \mathbf{T} \mathbf{a}_\alpha \\ \text{s.a.} & \mathbf{a}'_\alpha \mathbf{a}_\alpha = 1 \,. \\ & \mathbf{a}'_\alpha \mathbf{a}_j = 0, \; j = 1, \ldots, \alpha - 1 \,. \end{array}$$

Solución: Si λ_α es el α-ésimo mayor valor propio de T, con vector propio $\mathbf{a}_\alpha$, entonces el eje principal α-ésimo es la recta con vector director $\mathbf{a}_\alpha$, y λ_α es la inercia de dicho eje principal, es decir, la inercia máxima es $I_\alpha = \lambda_\alpha$.

Coordenadas de los puntos fila en los ejes principales: Los vectores directores de los ejes principales son

$$\begin{aligned}\mathbf{a}_1 &= (a_{11}, \ldots, a_{p1})' \\ &\vdots \\ \mathbf{a}_r &= (a_{1r}, \ldots, a_{pr})'\end{aligned}$$

Llamamos $\mathbf{x}_i$ al vector que va desde el origen hasta el punto X_i. Las coordenadas de X_i en los ejes determinados por $\mathbf{a}_1, \ldots, \mathbf{a}_r$ son:

$$\begin{aligned}g_{i1} &= \ldots \qquad = \ \ldots \qquad\qquad\qquad , \\ &\vdots \\ g_{ir} &= \ldots \qquad = \ \ldots \qquad\qquad\qquad .\end{aligned}$$

Las coordenadas del punto fila $X_i = (x_{i1}, \ldots, x_{ip})$ en el nuevo sistema de coordenadas formado por los ejes principales son

$$X_i \longrightarrow G_i = (g_{i1}, \ldots, g_{ir}) \ ,$$

Coordenadas de los n puntos fila en los ejes principales:

$$\begin{aligned}X_1 &\longrightarrow G_1 = (g_{11}, \ldots, g_{1r}) \\ \vdots \quad &\qquad\ \vdots \\ X_n &\longrightarrow G_n = (g_{n1}, \ldots, g_{nr})\end{aligned}$$

Observemos que el centro de gravedad de las nuevas coordenadas G_i, $i = 1, 2, \ldots, n$ es

$$\overline{G} = (0, \ldots, 0) = O \, .$$

Ejemplo 4.12 Los valores y vectores propios de la matriz $\mathbf{T}$ obtenida en el Ejemplo 4.11 son

$$\begin{aligned}\lambda_1 &= 0.3405, & \mathbf{a}_1 &= (0.4682, 0.3452, -0.8134)' \\ \lambda_2 &= 0.2086, & \mathbf{a}_2 &= (-0.6689, 0.7399, -0.0710)' \\ \lambda_3 &= 0, & \mathbf{a}_3 &= (0.5773, 0.5773, 0.5773)'\end{aligned}$$

Hay $r = 2$ ejes principales, con vectores directores $\mathbf{a}_1$ y $\mathbf{a}_2$,. Las inercias de dichos ejes son

$$I_1 = ... \qquad\qquad\qquad \text{e} \quad I_2 = ... \qquad\qquad\qquad .$$

Las coordenadas del punto fila "Marca A" en los ejes principales son $G_1 = (g_{11}, g_{12})$, donde

$$g_{11} = ...$$
$$g_{12} = ...$$

Para el punto fila "Marca B" son $G_2 = (g_{21}, g_{22})$, donde

$$g_{21} = ...$$
$$g_{22} = ...$$

Las coordenadas de "Marca C" son $G_3 = (g_{31}, g_{32})$, donde

$$g_{31} = ...$$
$$g_{32} = ...$$

Por último, las coordenadas de "Marca D" son $G_4 = (F_{41}, F_{42})$, siendo

$$g_{41} = ...$$
$$g_{42} = ...$$

Proposición 4.3 El vector $\overline{R}^* = \left(\sqrt{f_{+1}}, \ldots, \sqrt{f_{+p}}\right)$ es un vector propio de $\mathbf{T}$ asociado al valor propio 0.

Esta proposición implica que hay como máximo $p - 1$ ejes principales.

4.6. ANÁLISIS DE LA NUBE DE PUNTOS COLUMNA

Nube de puntos columna:

$$N(\mathcal{J}) = \{(C_j, f_{+j}), j = 1, 2, \ldots, p\},$$

donde

$$C_j = (c_{1j}, c_{2j}, \ldots, c_{nj}) = \left(\text{---}, \text{---}, \ldots, \text{---} \right), \; j = 1, 2, \ldots, p\,.$$

Distancia $\mathcal{X}^2$ entre dos puntos columna:

$d_{\mathcal{X}}(C_j, C_{j'}) = ...$

Nube de puntos columna transformada: La nube de puntos columna transformada va a permitir utilizar la distancia euclídea

$$N^*(\mathcal{J}) = \{(C_j^*, f_{+j}), j = 1, 2, \ldots, p\}\,,$$

donde

$C_j^* = (c_{1j}^*, \ldots, c_{nj}^*) = ...$.

Centro de gravedad de $N^*(\mathcal{J})$:

$$\overline{C}^* = \left(\sqrt{f_{1+}}, \ldots, \sqrt{f_{n+}} \right)\,.$$

Nube de puntos centrada:

$$N^{**}(\mathcal{J}) = \{(Y_j, f_{+j}), j = 1, 2, \ldots, p\}\,,$$

donde

$Y_j = (y_{1j}, \ldots, y_{nj}) = ...$

Puntos columna centrados:

$$\begin{gathered} Y_1 = (y_{11}, \ldots, y_{n1}) \\ \vdots \\ Y_p = (y_{1p}, \ldots, y_{np}) \end{gathered}$$

La coordenada de un punto Y_j en el eje con vector director $\mathbf{b}_\beta$ es

$$h_{j\beta} = ...$$

Inercia del eje β:

$$I_\beta = ...$$

Matriz de inercia de los puntos columna: Consideremos la matriz de puntos columna $\mathbf{Y} = (y_{ij})$, y la matriz diagonal con los pesos de los puntos columna, $D_\mathcal{J} = \mathrm{diag}\{f_{+1}, \ldots, f_{+p}\}$. La *matriz de inercia de los puntos columna* es la matriz $\mathbf{S} = \mathbf{Y} D_\mathcal{J} \mathbf{Y}'$. El elemento (i, j) de esta matriz es

$$s_{ij} = \sum_{k=1}^{p} f_{+k}\, y_{ik}\, y_{jk}\, .$$

Se verifica que

$$I_\beta = \mathbf{b}'_\beta\, \mathbf{S}\, \mathbf{b}_\beta\, .$$

Eje principal de la nube de puntos columna: Es el eje que maximiza la inercia de los puntos columna. Se encuentra entonces resolviendo el problema

$$\begin{array}{ll} \text{Max} & I_1 = \mathbf{b}'_1\, \mathbf{S}\, \mathbf{b}_1 \\ \text{s.a.} & \mathbf{b}'_1 \mathbf{b}_1 = 1\, . \end{array}$$

La inercia máxima es el mayor valor propio de la matriz $\mathbf{S}$, llamémoslo λ_1, y $\mathbf{b}_1$ es el vector propio unitario asociado a λ_1.

Eje principal β-ésimo de la nube de puntos columna: Es el eje ortogonal a los ejes principales $1, 2, \ldots, \beta - 1$ con inercia máxima.

El eje principal β-ésimo es la recta cuyo vector director es $\mathbf{b}_\beta$, el vector propio asociado al β-ésimo mayor valor propio λ_β de $\mathbf{S}$.

Además, λ_β es la inercia del eje principal β-ésimo.

Coordenadas de los puntos columna en los ejes principales: Los vectores directores de los ejes principales son

$$\mathbf{b}_1, \ldots, \mathbf{b}_r$$

Coordenadas de los p puntos columna en los r ejes principales:

$$\begin{array}{c} Y_1 \longrightarrow H_1 = (h_{11}, \ldots, h_{1r}), \quad h_{1\beta} = \mathbf{y}_1' \mathbf{b}_\beta \\ \vdots \\ Y_p \longrightarrow H_p = (h_{p1}, \ldots, h_{pr}), \quad h_{p\beta} = \mathbf{y}_p' \mathbf{b}_\beta \end{array}$$

Proposición 4.4 El vector $\overline{C}^* = \left(\sqrt{f_{1+}}, \ldots, \sqrt{f_{p+}}\right)$ es un vector propio de $\mathbf{S}$ asociado al valor propio 0.

Por tanto, hay como máximo $n-1$ ejes principales para la nube de puntos columna.

Proposición 4.5 Las matrices $\mathbf{T}$ y $\mathbf{S}$ tienen los mismos valores propios no nulos.

✓ Para cada valor propio λ_α positivo, tenemos asociado un vector propio $\mathbf{a}_\alpha$ de $\mathbf{T}$, y un vector propio $\mathbf{b}_\alpha$ de $\mathbf{S}$.

✓ Tendremos $r = \text{mín}\{n-1, p-1\}$ ejes principales.

Proposición 4.6 (*Fórmulas de transición*)
Para cada uno de los ejes principales α, se verifica

$$h_{j\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^{n} c_{ij} g_{i\alpha}, \quad g_{i\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^{p} r_{ij} h_{j\alpha} \, .$$

Ejemplo 4.13 Para los datos del Ejemplo 4.1, obtener las coordenadas de los puntos columna en sus ejes principales.

Las coordenadas del primer punto columna, $H_1 = (h_{11}, h_{12})$, se obtienen como una media ponderada de las coordenas de los puntos fila, donde las ponderaciones son los elementos del primer perfil columna

$$h_{11} = ... \qquad ,$$
$$h_{12} = ... \qquad .$$

Las coordenadas del segundo punto columna son $H_2 = (h_{21}, h_{22})$, donde

$$h_{21} = ... \qquad ,$$
$$h_{22} = ... \qquad .$$

Por último, las coordenadas del tercer punto fila son $H_3 = (h_{31}, h_{32})$, donde

$$h_{31} = ... \qquad ,$$
$$h_{32} = ... \qquad .$$

4.7. INTERPRETACIÓN DE LOS RESULTADOS

4.7.1. SELECCIÓN DEL NÚMERO DE EJES PRINCIPALES

Los criterios de selección del número de ejes principales son:

- Proporción de inercia explicada $\geq 75\,\%$: Como la inercia del eje α es λ_α, la inercia total es

  $$I[N(\mathcal{I})] = \sum_{\alpha=1}^{r} \lambda_\alpha \,.$$

  Por tanto, la proporción de inercia explicada por el eje α es

  $$\frac{\lambda_\alpha}{\sum_{\alpha=1}^{r} \lambda_\alpha} \,.$$

La proporción de inercia explicada por los r_1 primeros ejes será

$$\frac{\sum_{\alpha=1}^{r_1} \lambda_\alpha}{\sum_{\alpha=1}^{r} \lambda_\alpha}$$

- Inercia superior a la inercia media.
- Gráfico de sedimentación.

4.7.2. CONTRIBUCIÓN ABSOLUTA Y RELATIVA

La inercia de un eje α es la suma de las inercias de los puntos fila proyectados en dicho eje,

$$\lambda_\alpha = ...$$

o de los puntos columna proyectados en dicho eje,

$$\lambda_\alpha = ...$$

Contribución absoluta de un punto fila a un eje: La *contribución absoluta* del punto fila i-ésimo a la inercia del eje α es

$$ca_\alpha(G_i) = \frac{f_{i+} g_{i\alpha}^2}{\lambda_\alpha}\,.$$

La contribución absoluta del punto columna j-ésimo al eje α es

$$ca_\alpha(H_j) = \frac{f_{+j} h_{j\alpha}^2}{\lambda_\alpha}\,.$$

✓ Las contribuciones absolutas miden la cantidad de inercia que aporta una categoría (un punto fila o columna) a la inercia de un eje.

✓ Las categorías con contribuciones absolutas más altas son las protagonistas en la construcción del eje, y nos van a servir para interpretar el sentido de los ejes principales.

Ejemplo 4.14 Calcular las contribuciones absolutas de los puntos fila y de los puntos columna a la inercia de los ejes principales obtenidos para el Ejemplo 4.1.

		Coordenadas		Contrib. absol.	
Punto fila	Peso	Dim 1	Dim 2	Dim 1	Dim 2
A	0.326	-0.819	-0.071		
B	0.288	0.315	0.674		
C	0.212	0.290	-0.414		
D	0.174	0.659	-0.477		

		Coordenadas		Contrib. absol.	
Punto columna	Peso	Dim 1	Dim 2	Dim 1	Dim 2
1	0.333	0.473	-0.529		
2	0.333	0.349	0.585		
3	0.333	-0.822	-0.056		

Ahora vamos a definir unas medidas de la cantidad de inercia que aporta un eje a un punto.

La inercia del punto fila G_i en el nuevo sistema de coordenadas es

$$I(G_i) = f_{i+} d_e^2(G_i, O), \quad d_e^2(G_i, O) = \sum_{\alpha=1}^{r} g_{i\alpha}^2 \, .$$

Contribución relativa de un eje a la inercia de un punto: La *contribución relativa* del eje α a la inercia del punto fila G_i es

$$cr_\alpha(G_i) = \frac{g_{i\alpha}^2}{d_e^2(G_i, O)} \, .$$

De la misma forma, la *contribución relativa* del eje α a la inercia del punto columna H_j es

$$cr_\alpha(H_j) = \frac{h_{j\alpha}^2}{d_e^2(H_j, O)} \, .$$

Las contribuciones relativas nos indican si los puntos están bien representados en los nuevos ejes.

Ejemplo 4.15 Calcular las contribuciones relativas de los puntos fila para los datos del Ejemplo 4.1.

	Contrib. rel.		
Punto fila	Dim 1	Dim 2	Total
A			
B			
C			
D			

Un punto se considera aceptablemente representado por los nuevos ejes si la suma de las contribuciones relativas de los ejes seleccionados a la inercia de este punto es al menos 0.6, es decir, si

$$\sum_{\alpha=1}^{r} cr_\alpha(G_i) \geq 0{,}6\,.$$

Si existe algún punto fila o columna que no esté bien representado, entonces no se pueden extraer conclusiones sobre sus relaciones con el resto de puntos.

Un punto fila G_i se considera representado por el eje α si

$$cr_\alpha(G_i) \geq 0{,}2\,.$$

Se podrán extraer conclusiones acerca de dicho punto en relación al eje α sólo cuando su contribución relativa supere esta cantidad.

4.7.3. REPRESENTACIONES GRÁFICAS

- Nube de puntos fila respecto a sus ejes principales.
- Nube de puntos columna respecto a sus ejes principales.
- Representación simultánea de puntos fila y columna en un espacio conjunto.