

Análisis de Correspondencias

Basado en las notas de Norberto Corral Blanco y Beatriz Sinova Fernández
Facultad de Ciencias, Universidad de Oviedo

1 de mayo de 2026

Índice

1. Introducción	2
2. Notación	2
3. Perfiles fila y distancia χ^2	2
3.1. Propiedad de equivalencia distribucional	3
4. Análisis de componentes principales sobre los perfiles	4
4.1. Centrado	4
4.2. Matriz de covarianzas	4
4.3. Valores y vectores propios	5
4.4. Equivalencia con la matriz de frecuencias relativas condicionadas	5
5. Puntuaciones factoriales (componentes principales)	6
5.1. Puntos fila en \mathbb{R}^p	6
5.2. Puntos columna en \mathbb{R}^n	6
6. Relaciones de transición y espacio conjunto	7
6.1. Acotación de los valores propios	7
7. Reconstrucción de la tabla de frecuencias	8
8. Contribuciones absolutas y relativas	8
8.1. Contribuciones absolutas	8
8.2. Contribuciones relativas (o calidades de representación)	8
9. Puntos suplementarios	9

1. Introducción

El análisis de correspondencias es una técnica descriptiva para representar tablas de contingencia, cuyo objetivo es analizar la razón por la cual dos variables cualitativas no son independientes. Se utiliza habitualmente en tablas de gran tamaño, y permite resumir una gran cantidad de datos en un número reducido de dimensiones, con la menor pérdida de información posible. La técnica se basa en el trabajo de H.O. Hirschfeld (1935) y fue desarrollada por J.P. Benzécri (1973).

A diferencia del test chi-cuadrado de Pearson, que simplemente detecta la existencia de asociación, el análisis de correspondencias busca describir la **estructura de dicha asociación**: qué categorías de una variable se relacionan con qué categorías de la otra y en qué sentido.

2. Notación

Consideramos una tabla de contingencia donde la variable X (filas) tiene n categorías y la variable Y (columnas) tiene p categorías.

$$\begin{aligned}k_{ij} &= \text{número de observaciones en la fila } i \text{ y columna } j, \\k_{i\cdot} &= \sum_{j=1}^p k_{ij} \quad (\text{total marginal de la fila } i), \\k_{\cdot j} &= \sum_{i=1}^n k_{ij} \quad (\text{total marginal de la columna } j), \\N &= \sum_{i=1}^n \sum_{j=1}^p k_{ij} \quad (\text{total de observaciones}).\end{aligned}$$

Las frecuencias relativas y marginales son:

$$f_{ij} = \frac{k_{ij}}{N}, \quad f_{i\cdot} = \frac{k_{i\cdot}}{N} = \sum_{j=1}^p f_{ij}, \quad f_{\cdot j} = \frac{k_{\cdot j}}{N} = \sum_{i=1}^n f_{ij}.$$

Se cumple $\sum_{i=1}^n f_{i\cdot} = \sum_{j=1}^p f_{\cdot j} = 1$.

3. Perfiles fila y distancia χ^2

Cada categoría fila i se caracteriza por su **perfil**, que es la distribución condicionada de las columnas dado i :

$$i \longrightarrow \left(\frac{f_{ij}}{f_{i\cdot}} \right)_{j=1}^p.$$

Bajo independencia ($f_{ij} = f_{i\cdot} f_{\cdot j}$), todos los perfiles fila serían idénticos al vector de marginales columna: $(f_{\cdot j})_{j=1}^p$. Por tanto, la asociación entre X e Y se traduce en **dispersión de los perfiles fila** respecto de este perfil medio.

Para medir dicha dispersión se introduce la **distancia** χ^2 entre dos filas i e i' :

$$d^2(i, i') = \sum_{j=1}^p \frac{1}{f_{\cdot j}} \left(\frac{f_{ij}}{f_{i\cdot}} - \frac{f_{i'j}}{f_{i'\cdot}} \right)^2.$$

El factor $1/f_{\cdot j}$ pondera cada columna de forma inversamente proporcional a su frecuencia, para que las diferencias en categorías raras tengan más importancia relativa que las mismas diferencias en categorías muy frecuentes.

Esta distancia puede reescribirse como una distancia euclídea estándar si redefinimos las coordenadas del punto fila:

$$i \longrightarrow \vec{x}_i = \left(\frac{f_{ij}}{f_{i\cdot} \sqrt{f_{\cdot j}}} \right)_{j=1}^p.$$

En efecto,

$$d^2(i, i') = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{i\cdot} \sqrt{f_{\cdot j}}} - \frac{f_{i'j}}{f_{i'\cdot} \sqrt{f_{\cdot j}}} \right)^2 = \|\vec{x}_i - \vec{x}_{i'}\|^2.$$

3.1. Propiedad de equivalencia distribucional

Si dos puntos fila i_1 e i_2 tienen idénticos perfiles, es decir,

$$\frac{f_{i_1 j}}{f_{i_1 \cdot}} = \frac{f_{i_2 j}}{f_{i_2 \cdot}}, \quad \forall j = 1, \dots, p,$$

y se agrupan en una única categoría i_0 (con $f_{i_0 j} = f_{i_1 j} + f_{i_2 j}$ y $f_{i_0 \cdot} = f_{i_1 \cdot} + f_{i_2 \cdot}$), entonces la distancia entre dos columnas j y j' no se altera.

Demostración. La distancia χ^2 entre columnas se define de manera análoga:

$$d^2(j, j') = \sum_{i=1}^n \frac{1}{f_{i\cdot}} \left(\frac{f_{ij}}{f_{\cdot j}} - \frac{f_{ij'}}{f_{\cdot j'}} \right)^2.$$

Separando los sumandos correspondientes a i_1 e i_2 y usando la igualdad de perfiles $\frac{f_{i_1 j}}{f_{i_1 \cdot}} = \frac{f_{i_2 j}}{f_{i_2 \cdot}} = \frac{f_{i_0 j}}{f_{i_0 \cdot}}$, se tiene:

$$\begin{aligned} \frac{1}{f_{i_1 \cdot}} \left(\frac{f_{i_1 j}}{f_{\cdot j}} - \frac{f_{i_1 j'}}{f_{\cdot j'}} \right)^2 &= f_{i_1 \cdot} \left(\frac{f_{i_0 j}/f_{i_0 \cdot}}{f_{\cdot j}} - \frac{f_{i_0 j'}/f_{i_0 \cdot}}{f_{\cdot j'}} \right)^2, \\ \frac{1}{f_{i_2 \cdot}} \left(\frac{f_{i_2 j}}{f_{\cdot j}} - \frac{f_{i_2 j'}}{f_{\cdot j'}} \right)^2 &= f_{i_2 \cdot} \left(\frac{f_{i_0 j}/f_{i_0 \cdot}}{f_{\cdot j}} - \frac{f_{i_0 j'}/f_{i_0 \cdot}}{f_{\cdot j'}} \right)^2. \end{aligned}$$

Sumando ambas contribuciones se obtiene

$$(f_{i_1 \cdot} + f_{i_2 \cdot}) \left(\frac{f_{i_0 j}/f_{i_0 \cdot}}{f_{\cdot j}} - \frac{f_{i_0 j'}/f_{i_0 \cdot}}{f_{\cdot j'}} \right)^2 = \frac{1}{f_{i_0 \cdot}} \left(\frac{f_{i_0 j}}{f_{\cdot j}} - \frac{f_{i_0 j'}}{f_{\cdot j'}} \right)^2,$$

que es exactamente el sumando que corresponde a la categoría agrupada i_0 . Por tanto, $d^2(j, j')$ queda invariante. \square

4. Análisis de componentes principales sobre los perfiles

El objetivo es analizar la variabilidad de la nube de puntos fila, ponderada por las masas $f_{i\cdot}$, en el espacio \mathbb{R}^p . Para ello se aplica un análisis de componentes principales a las variables

$$\text{variable } j \longrightarrow \left(\frac{f_{ij}}{f_{i\cdot} \sqrt{f_{\cdot j}}} \right)_{i=1}^n.$$

Las etapas son las habituales: centrar las variables, calcular la matriz de covarianzas, obtener valores y vectores propios, y proyectar.

4.1. Centrado

La esperanza (media ponderada por f_i) de la variable j es

$$E(j) = \sum_{i=1}^n \frac{f_{ij}}{f_i \sqrt{f_{\cdot j}}} f_i = \frac{1}{\sqrt{f_{\cdot j}}} \sum_{i=1}^n f_{ij} = \frac{f_{\cdot j}}{\sqrt{f_{\cdot j}}} = \sqrt{f_{\cdot j}}.$$

Las coordenadas centradas del punto fila i son, por tanto,

$$i \rightarrow \left(\frac{f_{ij}}{f_i \sqrt{f_{\cdot j}}} - \sqrt{f_{\cdot j}} \right)_{j=1}^p.$$

4.2. Matriz de covarianzas

La covarianza entre las variables j y j' se calcula como:

$$\text{Cov}(j, j') = \sum_{i=1}^n f_i \left(\frac{f_{ij}}{f_i \sqrt{f_{\cdot j}}} - \sqrt{f_{\cdot j}} \right) \left(\frac{f_{ij'}}{f_i \sqrt{f_{\cdot j'}}} - \sqrt{f_{\cdot j'}} \right).$$

Definimos los coeficientes

$$w_{ij} = \frac{f_{ij}}{\sqrt{f_i \cdot f_{\cdot j}}} - \sqrt{f_i \cdot f_{\cdot j}},$$

con lo cual la covarianza anterior se reescribe como

$$\text{Cov}(j, j') = \sum_{i=1}^n w_{ij} w_{ij'}.$$

Así, la matriz de varianzas-covarianzas de las p variables es

$$T = W^t W, \quad \text{donde } W = (w_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, p}}.$$

Nótese que cada w_{ij} puede interpretarse como el **residuo tipificado** de Pearson:

$$w_{ij} = \frac{f_{ij} - f_i \cdot f_{\cdot j}}{\sqrt{f_i \cdot f_{\cdot j}}}.$$

4.3. Valores y vectores propios

Sean $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ los valores propios de $T = W^t W$ y $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p$ los vectores propios ortonormales asociados.

Teorema 4.1 (Valor propio nulo). *La matriz T tiene un valor propio igual a 0, con vector propio $\vec{u}_p = (\sqrt{f_{\cdot j}})_{j=1}^p$.*

Demostración. Basta comprobar que $W\vec{u}_p = \vec{0}$, ya que entonces $W^t W\vec{u}_p = \vec{0}$. Para cada fila i ,

$$\begin{aligned} \sum_{j=1}^p w_{ij} \sqrt{f_{\cdot j}} &= \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_i \cdot f_{\cdot j}}} - \sqrt{f_i \cdot f_{\cdot j}} \right) \sqrt{f_{\cdot j}} \\ &= \sum_{j=1}^p \frac{f_{ij}}{\sqrt{f_i}} - \sqrt{f_i} \cdot \sum_{j=1}^p f_{\cdot j} \\ &= \frac{1}{\sqrt{f_i}} f_i - \sqrt{f_i} \cdot 1 = 0. \end{aligned}$$

Por tanto, $W\vec{u}_p = \vec{0}$ y \vec{u}_p es vector propio de T con valor propio 0. \square

La existencia de este valor propio nulo refleja que las frecuencias marginales suman 1 y, en consecuencia, la nube de puntos está contenida en un subespacio de dimensión a lo sumo $p - 1$.

4.4. Equivalencia con la matriz de frecuencias relativas condicionadas

En muchos textos se trabaja directamente con la matriz W^* de términos

$$w_{ij}^* = \frac{f_{ij}}{\sqrt{f_{i\cdot} f_{\cdot j}}}.$$

La relación con W es $w_{ij} = w_{ij}^* - \sqrt{f_{i\cdot} f_{\cdot j}}$.

Teorema 4.2 (Equivalencia de vectores propios). *Todo vector propio \vec{u}_α de $W^t W$ con $\alpha \neq p$ (asociado a λ_α) es también vector propio de $(W^*)^t W^*$ asociado al mismo valor propio λ_α .*

Demostración. Sea \vec{u}_α con $\alpha \neq p$ tal que $W^t W \vec{u}_\alpha = \lambda_\alpha \vec{u}_\alpha$. La componente j' -ésima de $W^t W \vec{u}_\alpha$ es

$$\sum_{j=1}^p \sum_{i=1}^n w_{ij'} w_{ij} u_{\alpha j}.$$

Sustituyendo $w_{ij} = w_{ij}^* - \sqrt{f_{i\cdot} f_{\cdot j}}$ y desarrollando:

$$\sum_{j=1}^p \sum_{i=1}^n w_{ij'} w_{ij} u_{\alpha j} = \sum_{i=1}^n w_{ij'} \sum_{j=1}^p w_{ij}^* u_{\alpha j} - \sum_{i=1}^n w_{ij'} \sqrt{f_{i\cdot}} \sum_{j=1}^p \sqrt{f_{\cdot j}} u_{\alpha j}.$$

Como \vec{u}_α es ortogonal a $\vec{u}_p = (\sqrt{f_{\cdot j}})_j$, se tiene $\sum_j \sqrt{f_{\cdot j}} u_{\alpha j} = 0$, luego el segundo sumando se anula. Análogamente, podemos reexpresar $\sum_i w_{ij'} w_{ij}^*$ y, tras un cálculo similar, se llega a que la componente j' -ésima de $(W^*)^t W^* \vec{u}_\alpha$ coincide con la de $W^t W \vec{u}_\alpha$. De manera concisa:

$$(W^*)^t W^* \vec{u}_\alpha = W^t W \vec{u}_\alpha + \text{términos que contienen } \sum_j \sqrt{f_{\cdot j}} u_{\alpha j} = 0,$$

por lo que $(W^*)^t W^* \vec{u}_\alpha = \lambda_\alpha \vec{u}_\alpha$. □

Teorema 4.3 (Valor propio unidad de W^*). *El vector $\vec{u}_p = (\sqrt{f_{\cdot j}})_{j=1}^p$ es vector propio de $(W^*)^t W^*$ asociado al valor propio 1.*

Demostración. Para cada j' , la componente j' -ésima de $(W^*)^t W^* \vec{u}_p$ es

$$\begin{aligned} \sum_{j=1}^p \sum_{i=1}^n \frac{f_{ij'}}{\sqrt{f_{i\cdot} f_{\cdot j'}}} \frac{f_{ij}}{\sqrt{f_{i\cdot} f_{\cdot j}}} \sqrt{f_{\cdot j}} &= \frac{1}{\sqrt{f_{\cdot j'}}} \sum_{i=1}^n \frac{f_{ij'}}{f_{i\cdot}} \sum_{j=1}^p f_{ij} \\ &= \frac{1}{\sqrt{f_{\cdot j'}}} \sum_{i=1}^n f_{ij'} = \frac{f_{\cdot j'}}{\sqrt{f_{\cdot j'}}} = \sqrt{f_{\cdot j'}}. \end{aligned}$$

Así, $(W^*)^t W^* \vec{u}_p = 1 \cdot \vec{u}_p$. □

En consecuencia, los valores propios no nulos de T coinciden con los de $(W^*)^t W^*$ excepto el valor propio 1 que reemplaza al 0 de T . En la práctica, se trabaja con los vectores propios de $W^t W$ (o de $(W^*)^t W^*$ omitiendo el asociado a 1).

5. Puntuaciones factoriales (componentes principales)

5.1. Puntos fila en \mathbb{R}^p

La puntuación del punto fila i en el eje α (con $\alpha \neq p$) es:

$$\psi_{\alpha i} = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{i\cdot} \sqrt{f_{\cdot j}}} \right) u_{\alpha j},$$

donde se ha omitido el término $-\sqrt{f_{\cdot j}}$ porque $\sum_j \sqrt{f_{\cdot j}} u_{\alpha j} = 0$ para $\alpha \neq p$. En notación matricial, $\Psi = W^* U$.

5.2. Puntos columna en \mathbb{R}^n

De manera simétrica, los puntos columna se representan mediante sus perfiles:

$$j \rightarrow \left(\frac{f_{ij}}{f_{\cdot j} \sqrt{f_{i\cdot}}} \right)_{i=1}^n,$$

cuya esperanza es $\sqrt{f_{i\cdot}}$. La puntuación de la columna j en el eje α es

$$\varphi_{\alpha j} = \sum_{i=1}^n \left(\frac{f_{ij}}{f_{\cdot j} \sqrt{f_{i\cdot}}} \right) v_{\alpha i},$$

donde \vec{v}_α es el vector propio unitario de WW^t asociado a λ_α .

6. Relaciones de transición y espacio conjunto

La descomposición en valores singulares de W (o de W^*) vincula los dos espacios. Como $W = \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} \vec{v}_\alpha \vec{u}_\alpha^t$, se verifican las **fórmulas de transición**:

$$\vec{v}_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} W \vec{u}_\alpha, \quad \vec{u}_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} W^t \vec{v}_\alpha.$$

Desarrollando la primera:

$$\begin{aligned} v_{\alpha i} &= \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_{i\cdot} f_{\cdot j}}} - \sqrt{f_{i\cdot} f_{\cdot j}} \right) u_{\alpha j} \\ &= \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \frac{f_{ij}}{\sqrt{f_{i\cdot} f_{\cdot j}}} u_{\alpha j} = \frac{\sqrt{f_{i\cdot}}}{\sqrt{\lambda_\alpha}} \psi_{\alpha i}, \end{aligned}$$

donde se ha usado que $\sum_j \sqrt{f_{\cdot j}} u_{\alpha j} = 0$. Por tanto,

$$\boxed{\psi_{\alpha i} = \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_{i\cdot}}} v_{\alpha i}, \quad \varphi_{\alpha j} = \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_{\cdot j}}} u_{\alpha j}.$$

Las relaciones entre las puntuaciones de filas y columnas se obtienen inmediatamente:

$$\varphi_{\alpha j} = \frac{1}{\lambda_\alpha} \sum_{i=1}^n \frac{f_{ij}}{f_{\cdot j}} \psi_{\alpha i}, \quad \psi_{\alpha i} = \frac{1}{\lambda_\alpha} \sum_{j=1}^p \frac{f_{ij}}{f_{i\cdot}} \varphi_{\alpha j}.$$

Estas expresiones muestran que, salvo un factor de escala, cada puntuación columna es un promedio ponderado de las puntuaciones fila, y viceversa. Esta **relación baricéntrica** permite representar puntos fila y columna en un mismo plano factorial, interpretando la proximidad entre una fila y una columna como indicio de asociación.

6.1. Acotación de los valores propios

Teorema 6.1. *Todos los valores propios λ_α de $W^t W$ son menores o iguales que 1.*

Demostración. Supongamos que existe $\lambda_\alpha > 1$. Usando la relación $\varphi_{\alpha j} = \frac{1}{\lambda_\alpha} \sum_i \frac{f_{ij}}{f_{.j}} \psi_{\alpha i}$, y teniendo en cuenta que $\frac{f_{ij}}{f_{.j}} \geq 0$ y $\sum_i \frac{f_{ij}}{f_{.j}} = 1$, $\varphi_{\alpha j}$ es una combinación lineal convexa de los $\psi_{\alpha i}$. Por tanto,

$$|\varphi_{\alpha j}| \leq \max_i |\psi_{\alpha i}|.$$

Como $\lambda_\alpha > 1$, el factor $1/\lambda_\alpha$ hace que la desigualdad sea estricta para algún j , obteniéndose $\max_j |\varphi_{\alpha j}| < \max_i |\psi_{\alpha i}|$. Aplicando el mismo razonamiento a la relación inversa $\psi_{\alpha i} = \frac{1}{\lambda_\alpha} \sum_j \frac{f_{ij}}{f_{i.}} \varphi_{\alpha j}$, se deduce que $\max_i |\psi_{\alpha i}| < \max_j |\varphi_{\alpha j}|$. Ambas desigualdades son contradictorias, a menos que $\max_i |\psi_{\alpha i}| = \max_j |\varphi_{\alpha j}| = 0$. Pero en ese caso el eje α tendría varianza nula, lo que es imposible para $\alpha \neq p$ por la definición de los valores propios. En consecuencia, $\lambda_\alpha \leq 1$. \square

7. Reconstrucción de la tabla de frecuencias

Las frecuencias relativas f_{ij} pueden recuperarse completamente a partir de los marginales y de las puntuaciones factoriales.

Teorema 7.1 (Fórmula de reconstrucción).

$$f_{ij} = f_{i.} f_{.j} \left(1 + \sum_{\alpha=1}^{p-1} \frac{1}{\sqrt{\lambda_\alpha}} \psi_{\alpha i} \varphi_{\alpha j} \right).$$

(El sumatorio se extiende hasta $p-1$ porque $\lambda_p = 0$ y el término correspondiente se anula; en la versión con W^* se suma sobre $\alpha = 2, \dots, p$ con $\lambda_1 = 1$ y $\psi_{1i} = \varphi_{1j} = 1$.)

Demostración. Partimos de la descomposición espectral de W^* . Como $W^* = \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} \vec{v}_\alpha \vec{u}_\alpha^t$, el elemento (i, j) de W^* es

$$\frac{f_{ij}}{\sqrt{f_{i.} f_{.j}}} = \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} v_{\alpha i} u_{\alpha j}.$$

Sustituyendo $v_{\alpha i} = \frac{\sqrt{f_{i.}}}{\sqrt{\lambda_\alpha}} \psi_{\alpha i}$ y $u_{\alpha j} = \frac{\sqrt{f_{.j}}}{\sqrt{\lambda_\alpha}} \varphi_{\alpha j}$, se obtiene

$$\begin{aligned} \frac{f_{ij}}{\sqrt{f_{i.} f_{.j}}} &= \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} \frac{\sqrt{f_{i.}}}{\sqrt{\lambda_\alpha}} \psi_{\alpha i} \frac{\sqrt{f_{.j}}}{\sqrt{\lambda_\alpha}} \varphi_{\alpha j} \\ &= \sqrt{f_{i.} f_{.j}} \sum_{\alpha=1}^p \frac{1}{\sqrt{\lambda_\alpha}} \psi_{\alpha i} \varphi_{\alpha j}. \end{aligned}$$

Multiplicando ambos lados por $\sqrt{f_{i.} f_{.j}}$ y notando que para $\alpha = p$ se tiene $\lambda_p = 0$ (ese término se anula) o, si se prefiere, teniendo en cuenta que el eje trivial ($\lambda = 1$) tiene $\psi_{1i} = \varphi_{1j} = 1$, se llega a

$$f_{ij} = f_{i.} f_{.j} \left(1 + \sum_{\alpha=2}^p \frac{1}{\sqrt{\lambda_\alpha}} \psi_{\alpha i} \varphi_{\alpha j} \right),$$

que es la expresión deseada (con el sumatorio sobre los ejes no triviales). \square

8. Contribuciones absolutas y relativas

8.1. Contribuciones absolutas

Indican la parte de variabilidad del eje α explicada por cada punto.

$$\text{Fila } i : \quad \text{Ca}_\alpha(i) = \frac{f_{i.} \psi_{\alpha i}^2}{\lambda_\alpha}, \quad \text{Columna } j : \quad \text{Ca}_\alpha(j) = \frac{f_{.j} \varphi_{\alpha j}^2}{\lambda_\alpha}.$$

Se cumple $\sum_i \text{Ca}_\alpha(i) = \sum_j \text{Ca}_\alpha(j) = 1$.

8.2. Contribuciones relativas (o calidades de representación)

Representan la aportación del eje α a la distancia del punto al centroide (origen).

$$\text{Cr}_i(\alpha) = \frac{\psi_{\alpha i}^2}{d^2(i, G)}, \quad \text{Cr}_j(\alpha) = \frac{\varphi_{\alpha j}^2}{d^2(j, G)},$$

donde $d^2(i, G) = \sum_{\alpha=1}^{p-1} \psi_{\alpha i}^2$ es la distancia al cuadrado del punto i al centro de la nube.

9. Puntos suplementarios

En ocasiones, ciertas categorías fila o columna no deben intervenir en la construcción de los ejes (por ser *outliers*, tener una masa muy grande o ser categorías ilustrativas). Estos **puntos suplementarios** se proyectan después, usando las relaciones de transición, sin afectar a la inercia de los ejes.

Si las frecuencias del punto suplementario son k_{ij}^+ , con totales $k_{i\cdot}^+$ y $k_{\cdot j}^+$, sus puntuaciones son:

$$\psi_{\alpha i}^+ = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \frac{k_{ij}^+}{k_{i\cdot}^+} \varphi_{\alpha j}, \quad \varphi_{\alpha j}^+ = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \frac{k_{ij}^+}{k_{\cdot j}^+} \psi_{\alpha i}.$$