Capítulo 1

Álgebra, medidas estadísticas básicas y distancias

1.1. ÁLGEBRA LINEAL Y MATRICIAL

1.1.1. VECTORES Y ÁLGEBRA LINEAL

Un conjunto de n números reales $(a_1, \ldots, a_n)$ se puede representar:

- ✓ como un punto en el espacio n-dimensional;
- ✓ como un vector con punto inicial el origen de coordenadas $(0, \ldots, 0)$, y punto final $(a_1, \ldots, a_n)$.

Punto de $\mathbb{R}^n$:

$$A = (a_1, \ldots, a_n), \quad O = (0, \ldots, 0)$$

Vector de $\mathbb{R}^n$:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{o} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Los a_i se denominan *componentes* del vector.

Suma de vectores:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \\ \vdots \\ \\ \end{pmatrix}$$

Producto de un vector por un escalar:

$$k \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \\ \vdots \\ \\ \end{pmatrix}$$

Vector traspuesto:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{a}' = ...$$

Combinación lineal (CL) de vectores: Sean $\mathbf{a}_1, \ldots, \mathbf{a}_m \in I\!R^n$ vectores y $k_1, \ldots, k_m \in I\!R$ escalares. El vector siguiente es una *combinación lineal* de $\mathbf{a}_1, \ldots, \mathbf{a}_m$:

$$\mathbf{a} = k_1\mathbf{a}_1 + \cdots + k_m\mathbf{a}_m$$

Vectores linealmente dependientes e independientes: Sea $m \in I\!N$. Se dice que un conjunto de vectores $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ de $I\!R^n$ es *linealmente dependiente*, si existen escalares $k_1, \ldots, k_m \in I\!R$ no todos nulos, tales que

$$k_1\mathbf{a}_1 + \cdots + k_m\mathbf{a}_m = \mathbf{o}\,, \tag{1.1}$$

Un conjunto de vectores se dice que es *linealmente independiente* si no es linealmente dependiente, es decir, si la única solución del sistema de ecuaciones (1.1) es

$$k_1 = \cdots = k_m = 0\,.$$

Ejemplo 1.1 Razonar si los siguientes grupos de vectores son linealmente dependientes o independientes.

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 4 \\ -6 \end{pmatrix}$$

$$\mathbf{a}_3 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} -5 \\ 15/2 \end{pmatrix}$$

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} -3 \\ 1 \\ -4 \\ 1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} -1 \\ 4 \\ -3 \\ 0 \end{pmatrix}$$

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$$

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Producto escalar de dos vectores: El producto escalar de $\mathbf{a} = (a_1, \ldots, a_n)'$ por $\mathbf{b} = (b_1, \ldots, b_n)'$ es el número real

$$\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} a_i b_i \, .$$

Ejemplo 1.2 Contesta a las siguientes cuestiones:

(a) Para un vector $\mathbf{a} \neq \mathbf{o}$, ¿puede ser $\mathbf{a}'\mathbf{a} = 0$?

...

(b) ¿Crees que el producto escalar verifica la propiedad conmutativa?

...

(c) Sea $k \in I\!R$. ¿Crees que se cumple $(k\mathbf{a})'\mathbf{b} = k\,\mathbf{a}'\mathbf{b}$?

...

(d) Sean $\mathbf{a} = (a_1, \ldots, a_n)'$, $\mathbf{b} = (b_1, \ldots, b_n)'$ y $\mathbf{c} = (c_1, \ldots, c_n)'$. ¿Crees que se verifica $(\mathbf{a} + \mathbf{b})'\mathbf{c} = \mathbf{a}'\mathbf{c} + \mathbf{b}'\mathbf{c}$?

...

Módulo o norma de un vector: El módulo de un vector $\mathbf{a} = (a_1, \ldots, a_n)$ es el número real

$$|\mathbf{a}| = \sqrt{\mathbf{a}'\mathbf{a}} = \sqrt{\sum_{i=1}^{n} a_i^2}\,.$$

El módulo de un vector mide su longitud.

Teorema 1.1 Sean $\mathbf{a}, \mathbf{b} \in I\!R^n$ y $k \in I\!R$. Entonces

(a) $|\mathbf{a}| = 0 \Leftrightarrow \mathbf{a} = 0$,

(b) $|\mathbf{a}| > 0 \Leftrightarrow \mathbf{a} \neq 0$,

(c) $|k\mathbf{a}| = |k||\mathbf{a}|$,

(d) $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$.

Vectores unitarios: Un vector $\mathbf{a} \in I\!R^n$ con módulo $|\mathbf{a}| = 1$ se llama vector *unitario*.

Normalización de una vector: Si un vector $\mathbf{a} \neq \mathbf{o}$ no es unitario, se puede construir el vector

$$\mathbf{a}^\circ = \frac{1}{|\mathbf{a}|}\,\mathbf{a}\,.$$

¿Es $\mathbf{a}^\circ$ unitario? ...

Ángulo entre dos vectores: Sean $\mathbf{a}, \mathbf{b} \in I\!R^n - \{\mathbf{o}\}$. El *ángulo entre a y b* es el número real contenido entre 0 y π definido por

$$\text{ang}(\mathbf{a}, \mathbf{b}) = \arccos \frac{\mathbf{a}'\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}.$$

Si $\mathbf{a} = \mathbf{o}$ ó $\mathbf{b} = \mathbf{o}$, entonces el ángulo no está definido .

Ejemplo 1.3 Sean los vectores de $I\!R^2$,

$$\mathbf{a} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad y \quad \mathbf{b} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}\,.$$

Calculamos el ángulo entre ellos:

$$\mathbf{a}'\mathbf{a} = \ldots \quad , \quad \mathbf{b}'\mathbf{b} = \ldots \quad ,$$
$$|\mathbf{a}| = \ldots \quad , \quad |\mathbf{b}| = \ldots \quad ,$$

$$\mathbf{a}'\mathbf{b} = \ldots \quad ,$$
$$\text{ang}(\mathbf{a}, \mathbf{b}) = \ldots \quad .$$

Vectores ortogonales: Dos vectores $\mathbf{a}$ y $\mathbf{b}$ son *ortogonales* si

$$\mathbf{a}'\mathbf{b} = 0\,.$$

Según esta definición, ¿a qué vectores es ortogonal el vector nulo?...

Vector proyección: Sean $\mathbf{a}$ y $\mathbf{c}$ dos vectores, donde $\mathbf{c} \neq \mathbf{o}$. El vector *proyección de* $\mathbf{a}$ *sobre* $\mathbf{c}$ es

$$\mathbf{c}_a = (\mathbf{a}'\mathbf{c}^\circ)\mathbf{c}^\circ\,.$$

El módulo del vector proyección es

$$|\mathbf{c}_a| = |\mathbf{a}'\mathbf{c}^\circ||\mathbf{c}^\circ| = |\mathbf{a}'\mathbf{c}^\circ|\,.$$

Ejemplo 1.4 Calculamos la proyección de los vectores $\mathbf{a}$ y $\mathbf{b}$ sobre $\mathbf{c}$, siendo

$$\mathbf{a} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\,.$$

Calcular asimismo la proyección de $\mathbf{a}+\mathbf{b}$ sobre $\mathbf{c}$. ¿Es igual a la suma de las proyecciones anteriores?

Vector que une dos puntos: Sean A, B dos puntos de $\mathbb{R}^n$. El vector con punto inicial A y punto final B es

$$\mathbf{v} = B - A.$$

Existe una correspondencia biunívoca entre los puntos de $\mathbb{R}^n$ y los vectores con punto inicial $O = (0, \ldots, 0)$: A cualquier punto A le corresponde un vector situado en el origen, $\mathbf{a} = A - O$, y viceversa.

Eje de coordenadas: Sea A un punto y $\mathbf{c}$ un vector. El *eje de coordenadas* con *punto origen* A y *vector director* $\mathbf{c}$ es la recta que pasa por A, con vector director $\mathbf{c}$.

Fijado un punto origen; por ejemplo, $O = (0, \ldots, 0)$, cada vector $\mathbf{c}$ determina un eje de coordenadas. Vamos a considerar ejes de coordenadas con origen en $O = (0, \ldots, 0)$ y determinados por vectores unitarios; es decir, con $|\mathbf{c}| = 1$.

Coordenada de un punto en un eje de coordenadas: Sea A un punto de $\mathbb{R}^n$ y $\mathbf{a} = A - O$ su vector asociado. Sea $\mathbf{c}$ un vector de $\mathbb{R}^n$ cuyo vector normalizado es $\mathbf{c}^\circ$. La *coordenada* del punto A en el eje de coordenadas determinado por $\mathbf{c}$ es el producto escalar

$$F(A, \mathbf{c}) = \mathbf{a}'\mathbf{c}^\circ \, .$$

Ejemplo 1.5 El eje X tiene punto origen $O = (0, \ldots, 0)$ y vector director $\mathbf{e}_1 = (1, 0)'$, y el eje Y tiene el mismo punto origen, pero el vector director es $\mathbf{e}_2 = (0, 1)'$. Sean los puntos $A = (2, 3)$ y $B = (-2, 1)$. Calcular:

✓ Coordenada de A sobre el eje X: ...

✓ Coordenada de A sobre el eje Y: ...

Considera ahora el eje de coordenadas con origen $O = (0, \ldots, 0)$ y vector director $\mathbf{u} = (1, 1)'$. Calcular:

✓ Coordenada de A sobre este eje: ...

✓ Coordenada de B sobre este eje: ...

Baricentro: Sea $(M_1, \ldots, M_k)$ una familia de puntos, y $(\alpha_1, \ldots, \alpha_k)$ una familia de escalares que suman uno. El *baricentro* o *centro de gravedad* de $(M_1, \ldots, M_k)$ con pesos $(\alpha_1, \ldots, \alpha_k)$ es el punto

$$M = \sum_{i=1}^{k} \alpha_i M_i \, .$$

Obsérvese que si todos los pesos son iguales, $\alpha_i = 1/k$, $i = 1, \ldots, k$, entonces M es la media aritmética de $M_1, \ldots, M_k$.

Ejemplo 1.6 La siguiente tabla de contingencia representa la distribución conjunta de 230 personas respecto a las variables *Estado Civil* y *Sexo*.

	Mujer	Hombre	Total
Soltero	112	8	120
Casado	107	3	110
Total	219	11	230

La siguiente tabla, llamada *tabla de perfiles fila*, representa la distribución de la variable *Sexo* por estado civil. Esta tabla permite comparar la distribución del sexo para los dos estados civiles.

	Mujer	Hombre	Total
Soltero	0.933	0.067	0.522
Casado	0.973	0.027	0.478
Total	0.952	0.048	1

Las filas de esta tabla son puntos en $\mathbb{R}^2$

$$M_1 = (0.933, 0.067), \quad M_2 = (0.973, 0.027)$$

Calcula el baricentro o centro de gravedad de M_1 y M_2 con pesos $\alpha_1 = 0{,}522$ y $\alpha_2 = 0{,}478$. Representar los puntos M_1, M_2 y el baricentro: ...

1.1.2. ÁLGEBRA MATRICIAL

Matriz: Sean $n, p \in \mathbb{N}$. Una *matriz de orden* $n \times p$ *sobre* $\mathbb{R}$ es un conjunto rectangular de np elementos de $\mathbb{R}$, representados en n filas y p columnas

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

El conjunto de todas las matrices reales de orden $n \times p$ se designa por $\mathbb{R}^{(n,p)}$.

Vectores fila: Las filas de A se pueden considerar como matrices de orden $1 \times p$ o como vectores de tamaño p:

$$\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{ip}), \ i = 1, \ldots, n\,.$$

Vectores columna: Las columnas de A se pueden considerar como matrices de orden $n \times 1$ o como vectores de tamaño n:

$$\mathbf{a}^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}, \; j = 1, \ldots, p\,.$$

Así, la matriz A puede escribirse como

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = (\mathbf{a}^1 \; \mathbf{a}^2 \; \ldots \; \mathbf{a}^p)\,.$$

Operaciones con matrices:

Operación	Restricciones	Definición
Suma	A, B del mismo orden	$A + B = (a_{ij} + b_{ij})$
Producto escalar	$c \in I\!R, \; A \in I\!R^{(n,p)}$	$cA = (ca_{ij})$
Multiplicación	$A \in I\!R^{(n,p)}, B \in I\!R^{(p,m)}$	$AB = (\mathbf{a}_i'\mathbf{b}^j)$
Traspuesta		$A' = (a_{ji})$
Traza	$n = p$	$\mathrm{tr}(A) = \sum_{i=1}^n a_{ii}$
Determinante	$n = p$	$\lvert A \rvert$
Inversa	$n = p$, $\lvert A \rvert \neq 0$	$AA^{-1} = A^{-1}A = \mathbf{I}$

Ejemplo 1.7

✓ ¿Se cumple la propiedad conmutativa para el producto de matrices? ...

✓ ¿Y la propiedad asociativa? ...

✓ ¿Y la propiedad distributiva? ...

Propiedades de la traspuesta:

$$
\begin{aligned}
&(A')' = \ldots \quad , \\
&(A+B)' = \ldots \qquad , \\
&(AB)' = \ldots \quad , \\
&A \text{ simétrica} \Leftrightarrow A' = \ldots \quad .
\end{aligned}
$$

Propiedades de la traza:

Sea $\alpha \in K$, y sean las matrices A_n, B_n, $C_{n\times p}$ y $D_{p\times n}$. La función traza cumple

$$
\begin{aligned}
&\mathrm{tr}(\alpha) = \ldots \qquad , \\
&\mathrm{tr}(\alpha A) = \ldots \qquad , \\
&\mathrm{tr}(A+B) = \ldots \qquad , \\
&\mathrm{tr}(CD) = \ldots \qquad .
\end{aligned}
$$

Determinantes: Fórmula de cálculo

(a) Matriz 2×2:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow |A| = a_{11}a_{22} - a_{21}a_{12}.$$

(b) Matriz $p \times p$:

$$A = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1j} & \cdots & a_{1p} \\ \vdots & & & \vdots & & \vdots \\ a_{i1} & \cdots & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & & & \vdots & & \vdots \\ \vdots & & & \vdots & & \vdots \\ a_{p1} & \cdots & \cdots & a_{pj} & \cdots & a_{pp} \end{pmatrix}$$

El *menor* de a_{ij} es el determinante de la matriz resultante al eliminar la fila i-ésima y la columna j-ésima. El *adjunto* de a_{ij} es $(-1)^{i+j}$ veces el menor de a_{ij}, y se denota por A_{ij}.

Tomamos una fila cualquiera, $\mathbf{a}_i = (a_{i1}, \ldots, a_{ip})$. El determinante de A es

$$|A| = a_{i1}A_{i1} + \cdots + a_{ip}A_{ip}.$$

Igualmente, tomamos una columna cualquiera, $\mathbf{a}_j = (a_{1j}, \ldots, a_{pj})'$. El determinante de A es

$$|A| = a_{1j}A_{1j} + \cdots + a_{pj}A_{pj}.$$

Ejemplo 1.8 Para la matriz A siguiente

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

escribir los adjuntos de los elementos a_{11} y a_{12}:

$$A_{11} = \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}, \quad A_{12} = -\begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}.$$

Propiedades:

(i) $|A| = |A'|$.

(ii) Si A es triangular o diagonal, $|A| = \prod_{i=1}^{p} a_{ii}$.

(iii) $|cA| = c^p|A|$, para $c \in K$.

(iv) $|AB| = |A||B|$.

Ejemplo 1.9 Calcular el determinante de la matriz

$$A = \begin{pmatrix} 1/4 & 3/4 & -1/4 \\ 3/4 & 1 & -1 \\ -1/4 & -1 & 1/4 \end{pmatrix}$$

Ejemplo 1.10 Comprobar (iv) para las matrices

$$A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & 0 \\ 1 & 5 \end{pmatrix}.$$

$$|A| = \ldots \quad , |B| = \ldots \quad , |AB| = \ldots \quad .$$

Matrices singulares o no singulares: Una matriz cuadrada es *singular* cuando $|A| = 0$; y es *no singular* si $|A| \neq 0$.

En una matriz singular, tanto las filas como las columnas son linealmente dependientes.

Ejemplo 1.11 Razonar si las siguientes matrices son singulares

$$A = \begin{pmatrix} 3 & 4 & 1 \\ 2 & 3 & 7 \\ 6 & 9 & 2 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 7 \\ 3 & 6 & 2 \end{pmatrix}, \; C = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}, \; a, b, c, d \neq 0 \, .$$

...

Matriz inversa: La *matriz inversa* de una matrix cuadrada A es la única matriz que verifica

$$AA^{-1} = A^{-1}A = \mathbf{I} \, .$$

A^{-1} existe si, y sólo si A es no singular.

Propiedades:

(i) $A^{-1} = (A_{ij})'/|A|$,

(ii) $(cA)^{-1} = \ldots$,

(iii) $(AB)^{-1} = \ldots$,

(iv) $(A^{-1})' = \ldots$,

(v) $(A^{-1})^{-1} = \ldots$.

Ejemplo 1.12 Calcular la inversa de las matrices del Ejemplo 1.11.

$$A^{-1} = \qquad \left(\qquad \right)$$

B no tiene inversa, ya que es singular.

$$C^{-1} = \left(\qquad \right)$$

Matrices ortogonales: Una matriz cuadrada A es *ortogonal* si cumple $AA' = \mathbf{I}$. Las propiedades más importantes son las siguientes:

(i) $A^{-1} = A'$

(ii) $A'A = \mathbf{I}$

(iii) Tanto las filas como las columnas de A son vectores ortogonales dos a dos.

(iv) Si A y B son ortogonales, entonces $C = AB$ es ortogonal.

Ejemplo 1.13 ¿Es A ortogonal?...

$$A = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Multiplícala por el vector $(1, 1)'$ y dibuja el resultado. ¿Qué transformación produce esta matriz?

$$\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \quad \end{pmatrix}.$$

Rango de una matriz: El *rango* de una matriz $A_{n\times p}$ es el número máximo de filas (o columnas) linealmente independientes.

Ejemplo 1.14 Calcular el rango de las matrices del Ejemplo 1.11.

$$\begin{aligned} \text{rang}(A) &= \ldots \quad , \\ \text{rang}(B) &= \ldots \quad , \\ \text{rang}(C) &= \ldots \quad . \end{aligned}$$

Valores propios: Se denominan *valores propios* de una matriz cuadrada A_p a las soluciones de la ecuación (con incógnita λ):

$$|A - \lambda \mathbf{I}| = 0.$$

El término $|A - \lambda \mathbf{I}|$ es un polinomio de grado p. Todo polinomio de grado p tiene p raíces, contando sus multiplicidades. Pero estas raíces pueden ser reales o complejas. Por tanto, hay p valores propios $\lambda_1, \ldots, \lambda_p$, donde algunos pueden ser iguales, y también pueden ser complejos.

✓ El rango de una matrix es igual al número de valores propios no nulos.

Vectores propios: Sea λ_i un valor propio de A. Un *vector propio* de A asociado a λ_i es un vector $\mathbf{v}_i = (v_{i1}, \ldots, v_{ip})$ que satisface

$$(A - \lambda_i \mathbf{I})\mathbf{v}_i = 0,$$

o equivalentemente,

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i.$$

Ejemplo 1.15 Sea la matriz

$$A = \begin{pmatrix} 3 & \sqrt{20} \\ \sqrt{20} & 4 \end{pmatrix}.$$

(a) Calcular su determinante y su traza.

(b) Calcular sus valores y vectores propios.

(c) Multiplica los valores propios. ¿Con qué coincide el resultado?

(d) Suma los valores propios. ¿con qué coincide el resultado?

(e) Haz el producto escalar de dos de los vectores propios asociados a los dos valores propios. ¿Qué ocurre?

Solución:...

Ejemplo 1.16 Sea $\mathbf{u}$ un vector propio de A asociado al valor propio λ.

(a) Sea c un escalar. ¿Es el vector $c\mathbf{u}$ vector propio de A? ¿Asociado a qué valor propio?

...

(b) Sea $\mathbf{v}$ otro vector propio de A asociado al mismo valor propio λ. ¿Es $\mathbf{u}+\mathbf{v}$ otro vector propio de A? ¿Asociado a qué valor propio?

...

(c) ¿Es $\mathbf{u}$ un vector propio de A^2? ¿Asociado a qué valor propio?

...

(d) Sea c un escalar. ¿Es $\mathbf{u}$ un vector propio de la matriz cA? ¿Asociado a qué valor propio?

...

Proposición 1.1 Si A es simétrica, entonces:

✓ todos sus valores propios son reales;

✓ los vectores propios asociados a distintos valores propios son ortogonales. Es decir, si $\mathbf{v}_i$ y $\mathbf{v}_j$ son dos vectores propios asociados a los valores propios $\lambda_i \neq \lambda_j$, entonces $\mathbf{v}_i'\mathbf{v}_j = 0$.

Descomposición espectral:
Sea A una matriz simétrica, con valores propios $\lambda_1, \ldots, \lambda_p$ y vectores propios asociados (normalizados) $\mathbf{v}_1, \ldots, \mathbf{v}_p$, es decir,

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i, \;\; i = 1, \ldots, p\,.$$

Escribimos estas igualdades en forma matricial. Para ello, definimos

$$P = (\mathbf{v}_1, \ldots, \mathbf{v}_p) \text{ y } \Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_p\}\,.$$

La matriz P es ortogonal, y se verifica

$$AP = P\Lambda \Leftrightarrow A = P\Lambda P'\,.$$

Teorema 1.2 (*Teorema de descomposición espectral*)
Cualquier matriz simétrica A puede expresarse como

$$A = P\Lambda P' = \lambda_1 \mathbf{v}_1 \mathbf{v}_1' + \cdots + \lambda_p \mathbf{v}_p \mathbf{v}_p' \,,$$

donde

- ✓ $\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_p\}$, donde $\lambda_1, \ldots, \lambda_p$ son los valores propios de A;
- ✓ $P = (\mathbf{v}_1, \ldots, \mathbf{v}_p)$, donde $\mathbf{v}_1, \ldots, \mathbf{v}_p$ son los vectores propios normalizados asociados a los valores propios de A, y P es ortogonal.

Además,

$$\text{rang}(A) = \text{número de valores propios no nulos.}$$

Ejemplo 1.17 Hacer la descomposición espectral de la matriz

$$\begin{pmatrix} 3 & \sqrt{20} \\ \sqrt{20} & 4 \end{pmatrix} .$$

1.2. MEDIDAS ESTADÍSTICAS BÁSICAS

1.2.1. MOMENTOS POBLACIONALES

Sea X una variable aleatoria con función de probabilidad $p(x)$ si es discreta, o función de densidad $f(x)$ si es continua.

Esperanza: La esperanza de una variable aleatoria X es

$$\mu_X = E(X) = \begin{cases} \sum_i x_i p(x_i) & \text{si } X \text{ es discreta;} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{si } X \text{ es continua.} \end{cases}$$

Propiedades: Sean X e Y dos variables aleatorias, y a, b, c constantes. Se verifica:

(i) $E(c) = ...$

(ii) $E(aX + b) = ...$

(iii) $E(X + Y) = ...$

(iv) $E(aX + bY) = ...$

Varianza:

$$\sigma_X^2 = V(X) = E\left[(X - E(X))^2\right] = E(X^2) - E^2(X).$$

Covarianza:

$$\begin{aligned} \sigma_{X,Y} &= Cov(X, Y) = E\left[(X - E(X))(Y - E(Y))\right] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Propiedades: Sean X, Y, U, V variables aleatorias y a, b, c, d constantes. Se verifica:

(i) $V(c) = ...$

(ii) $V(aX + b) = ...$

(iii) $V(X + Y) = ...$

(iv) $V(aX + bY) = ...$

(v) $Cov(aX + b, cY + d) = ...$

(vi) $Cov(X + Y, U + V) = ...$

(vii) $Cov(aX + bY, cU + dV) = ...$

Correlación:

$$\varrho_{X,Y} = Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}}.$$

Propiedades:

(i) $-1 \leq \varrho_{X,Y} \leq 1$;

(ii) La correlación lineal $\varrho_{X,Y}$ mide la fuerza de la asociación lineal entre X e Y. Si no existe dependencia lineal, $\varrho_{X,Y} = 0$; si la dependencia es lineal directa, $\varrho_{X,Y} > 0$, y si la dependencia es lineal inversa, $\varrho_{X,Y} < 0$.

Vector aleatorio: $\mathbf{X} = (X_1, \ldots, X_p)'$ es un vector cuyas componentes $X_1, \ldots, X_p$ son variables aleatorias.

Esperanza de un vector aleatorio: La esperanza de un vector aleatorio $\mathbf{X} = (X_1, \ldots, X_p)'$ es el vector compuesto por las esperanzas,

$$\boldsymbol{\mu}_X = E(\mathbf{X}) = (E(X_1), \ldots, E(X_p))'.$$

Matriz de varianzas-covarianzas: La matriz de varianzas-covarianzas de un vector aleatorio X es la matriz definida por:

$$\Sigma_X = V(\mathbf{X}) = E\left[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))'\right].$$

Equivalentemente:

$$\Sigma_X = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{pmatrix}$$

Matriz de correlaciones: La matriz de correlaciones de un vector aleatorio $\mathbf{X} = (X_1, \ldots, X_p)$ es

$$P_X = \begin{pmatrix} 1 & \varrho_{12} & \cdots & \varrho_{1p} \\ \varrho_{21} & 1 & \cdots & \varrho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varrho_{p1} & \varrho_{p2} & \cdots & 1 \end{pmatrix}$$

donde el elemento (i, j) es

$$\varrho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2 \sigma_j^2}}.$$

Momentos de CLs de variables aleatorias:

Sea $\mathbf{X} = (X_1, \ldots, X_p)'$ un vector aleatorio, y $\mathbf{a} = (a_1, \ldots, a_p)'$, $\mathbf{b} = (b_1, \ldots, b_p)'$ dos vectores de constantes. Definimos dos nuevas variables Z_1 y Z_2 que son combinación lineal de $\mathbf{X}$:

$$Z_1 = \mathbf{a}'\mathbf{X} = a_1 X_1 + \cdots + a_p X_p$$
$$Z_2 = \mathbf{b}'\mathbf{X} = b_1 X_1 + \cdots + b_p X_p$$

Se verifica:

$$E(Z_1) = \mathbf{a}' E(\mathbf{X}); \quad V(Z_1) = \mathbf{a}' V(\mathbf{X}) \mathbf{a}; \quad Cov(Z_1, Z_2) = \mathbf{a}' V(\mathbf{X}) \mathbf{b}.$$

1.2.2. MOMENTOS MUESTRALES

Media muestral: Sean $\{x_1, x_2, \ldots, x_n\}$ los valores observados de cierta variable X en n individuos. La *media muestral* de X es

$$\bar{x} = ...$$

Varianza muestral: La *varianza muestral* de X es

$$s_X^2 = ...$$

Este valor mide la dispersión de las observaciones alrededor de la media muestral.

Covarianza: Sean $\{x_{11}, x_{21}, \ldots, x_{n1}\}$, $\{x_{12}, x_{22}, \ldots, x_{n2}\}$ los valores observados de las variables X_1 y X_2 en n individuos. Se define la *covarianza muestral* entre X_1 y X_2 como

$$s_{X_1,X_2} = ...$$

La covarianza indica la asociación lineal que existe entre las variables X_1 y X_2. Si la relación entre las variables es lineal directa, la covarianza es positiva. Si la asociación es lineal inversa, entonces la covarianza es negativa.

$$\text{Si } X_2 = X_1, \text{ entonces } s_{X_1,X_2} = s_{X_1}^2.$$

La covarianza muestral tiene difícil interpretacion, ya que su magnitud depende de las unidades en las que estén medidas las variables. El coeficiente de correlación lineal evita este problema.

Coeficiente de correlación lineal: Se define el *coeficiente de correlación lineal* entre las variables X_1 y X_2 como

$$r_{X_1,X_2} = ... \qquad .$$

Ejemplo 1.18 Se dispone de los ingresos (X_1) y los gastos (X_2) anuales de 5 individuos. En la tabla de la izquierda están medidos en euros, mientras que en la de la derecha están medidos en miles de euros.

Ind.	Ingresos	Gastos
1	10000	9000
2	7000	8000
3	15000	13000
4	21000	20000
5	14000	13500
Media	13,400	12,700

Ind.	Ingresos	Gastos
1	10	9
2	7	8
3	15	13
4	21	20
5	14	13.5
Media	13.4	12.7

Para la tabla de la izquierda, la covarianza entre los ingresos y los gastos es $s_{X_1,X_2} = 19,820,000$ euros al cuadrado. Para la tabla de la derecha, la covarianza es de 19.82 miles de euros al cuadrado. El coeficiente de correlación es de 0.9829 en ambos casos.

Medidas multivariantes

Supongamos ahora que se han observado los valores de p variables $X_1, X_2, \ldots, X_p$ en n individuos, de manera que las observaciones se presentan en forma de una matriz $n \times p$,

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} . \tag{1.2}$$

Se define el *vector de medias muestrales* como

$$\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)', \quad \text{donde} \quad \bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}, \ j = 1, \ldots, p\,.$$

La *matriz de covarianzas muestral* de X_1, X_2,..., X_p es la matriz simétrica formada por las covarianzas entre cada par de variables,

$$S_X = \begin{pmatrix} s_1^2 & s_{12} & \cdots & s_{1p} \\ s_{21} & s_2^2 & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_p^2 \end{pmatrix} .$$

La *matriz de correlaciones muestral* de $X_1, X_2, \ldots, X_p$ es

$$R_X = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix} .$$

Ejemplo 1.19 Sean X_1 y X_2 el número de unidades vendidas de dos productos en 25 establecimientos.

X_1	X_2	X_1	X_2	X_1	X_2
191	155	179	158	192	154
195	149	183	147	174	143
181	148	174	150	176	139
183	153	190	159	197	167
176	144	188	151	190	163
208	157	163	137		
189	150	195	155		
197	159	186	153		
188	152	181	145		
192	150	175	140		

El vector de medias y la matriz de covarianzas son

$$\bar{\mathbf{x}} = \begin{pmatrix} 185{,}7 \\ 151{,}1 \end{pmatrix}, \quad S_X = \begin{pmatrix} 95{,}29 & 52{,}87 \\ 52{,}87 & 54{,}36 \end{pmatrix}$$

Sean dos nuevas variables Z_1 y Z_2 que se obtienen a partir de X_1 y X_2 de la forma

$$Z_1 = 0{,}2X_1 + 0{,}7X_2 \,,$$
$$Z_2 = -0{,}5X_1 + 0{,}1X_2 \,.$$

¿Cómo podemos obtener las medias muestrales de Z_1 y Z_2 a partir de las de X_1 y X_2? ¿y las varianzas?, ¿y la covarianza de Z_1 y Z_2?. ...

Momentos muestrales de CLs de variables:

Sean los vectores

$$\mathbf{a} = (a_1, a_2, \ldots, a_p)', \quad \mathbf{X} = (X_1, X_2, \ldots, X_p)'.$$

Definimos una variable Z que es combinación lineal de $\mathbf{X}$:

$$Z = \mathbf{a}'\mathbf{X} = a_1X_1 + a_2X_2 + \cdots + a_pX_p.$$

Se verifica:

$$\bar{z} = \mathbf{a}'\bar{\mathbf{x}}, \qquad s_Z^2 = \mathbf{a}'S_X\mathbf{a}.$$

Además, si tenemos dos combinaciones lineales de $\mathbf{X}$

$$Z_1 = \mathbf{a}_1'\mathbf{X} = a_{11}X_1 + a_{21}X_2 + \ldots + a_{p1}X_p\,,$$

$$Z_2 = \mathbf{a}_2'\mathbf{X} = a_{12}X_1 + a_{22}X_2 + \ldots + a_{p2}X_p\,.$$

entonces la covarianza entre ellas es

$$s_{Z_1,Z_2} = \mathbf{a}_1'S_X\mathbf{a}_2.$$

1.3. PROXIMIDADES

$X_1, \ldots, X_p$ variables
$A = (a_1, \ldots, a_p)$ valores de $X_1, \ldots, X_p$ para el individuo A
$B = (b_1, \ldots, b_p)$ valores de $X_1, \ldots, X_p$ para el individuo B

$$\text{Proximidades} \begin{cases} \text{Disimilaridades: } \delta(A,B) & \rightarrow \text{ Distancias: } d(A,B) \\ \text{Similaridades: } s(A,B) & \rightarrow \text{ Similitudes: } s(A,B) \end{cases}$$

1.3.1. VARIABLES CUANTITATIVAS

Distancia: Una *distancia* es una aplicación $d : I\!R^p \times I\!R^p \to I\!R$ que verifica:

(i) $d(A, B) = 0 \Leftrightarrow A = B$,

(ii) $d(A, B) + d(C, B) \geq d(A, C)$.

La distancia mide lo "lejanos" que están dos puntos.

Teorema 1.3 Para cualquier par $A, B \in I\!R^p$, se verifica

(iii) $d(A, B) \geq 0$,

(iv) $d(A, B) = d(B, A)$.

Por contra, una medida de similitud mide "lo cercanos" que están dos puntos.

Similitud: Una *similitud* es una aplicación $s : I\!R^p \times I\!R^p \to I\!R$ que verifica:

(i) $0 \leq s(A, B) \leq 1, \ \forall i, j$.

(ii) $s(A, B) = 1 \Leftrightarrow A$ y B son iguales.

(iii) $s(A, B) = s(B, A)$.

Distancia euclídea o L_2: La *distancia euclídea* entre dos puntos $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$ es el módulo del vector que une A con B, es decir,

$$d_e(A, B) = |B - A| = \sqrt{(B - A)'(B - A)} = \sqrt{\sum_{k=1}^{p} (b_k - a_k)^2} \, .$$

Distancia L_1, de Manhattan o city block:

$$d_M(A, B) = \sum_{k=1}^{p} |b_k - a_k| \, .$$

Distancia de Minkowski:

$$d_m(A,B) = \left(\sum_{k=1}^{p} |b_k - a_k|^m \right)^{1/m} .$$

Ejemplo 1.20 Para la tabla de perfiles fila del Ejemplo 1.6, calcular la distancia euclídea entre las dos filas

$$r_1 = (0.9332, 0.0667), \quad r_2 = (0.9726, 0.0273).$$

La distancia euclídea es:

$$d_e(r_1, r_2) = ...$$

Cada coordenada aporta la misma cantidad de distancia, a pesar de que intuitivamente parece que las primeras coordenadas están más cerca. La distancia ji-cuadrado permite ponderar cada coordenada, para tener en cuenta su magnitud.

Distancia chi-cuadrado: La *distancia ji-cuadrado* entre los puntos $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$, con pesos $\omega = (\omega_1, \ldots, \omega_p)$ es

$$d_{\mathcal{X}}(A,B) = \sum_{k=1}^{p} \omega_k (b_k - a_k)^2 .$$

Ejemplo 1.21 Si en el ejemplo anterior tomamos como pesos los inversos de la última fila de la tabla, es decir,

$$\omega_1 = \frac{230}{219}, \quad \omega_2 = \frac{230}{11} ,$$

obtenemos

$$d_{\mathcal{X}}(r_1, r_2) = ...$$

Observa que ahora es la segunda coordenada la que aporta mayor cantidad a la distancia.

Ninguna de las distancias anteriores tiene en cuenta la correlación entre las variables $X_1, \ldots, X_p$. La distancia de Mahalanobis tiene en cuenta las varianzas y la correlación que hay entre ellas.

Distancia de Mahalanobis: Se miden p variables $X_1, \ldots, X_p$ a n individuos, y se obtiene la matriz de varianzas-covarianzas muestral S_X. Sean $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$ los valores de las variables para dos individuos A y B. La *distancia de Mahalanobis* entre $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$ es

$$d_M(A, B) = (B - A)' S_X^{-1} (B - A)\,.$$

Obsérvese que esta distancia solo se puede calcular si se dispone de un conjunto de n mediciones de $X_1, \ldots, X_p$.

Ejemplo 1.22 Consideremos los datos económicos (en millones de dólares) de las 10 corporaciones industriales estadounidenses más importantes:

Compañía	Ventas	Beneficios	Bienes
General Motors	126.974	4.224	173.297
Ford	96.933	3.835	160.893
Exxon	86.656	3.510	83.219
IBM	63.438	3.758	77.734
General Electric	55.264	3.939	128.344
Mobil	50.976	1.809	39.08
Philip Morris	39.069	2.946	38.528
Chrisler	36.156	0.359	51.038
Du Pont	35.209	2.480	34.715
Texaco	32.416	2.413	25.636

Calcular la distancia de Mahalanobis entre Ford y Exxon.

La matriz de varianzas-covarianzas es

$$\begin{pmatrix} 1000{,}5090 & 23{,}0179 & 1360{,}6444 \\ 23{,}0179 & 1{,}4299 & 41{,}0892 \\ 1360{,}6444 & 41{,}0892 & 2980{,}4898 \end{pmatrix} .$$

Por tanto, la distancia de Mahalanobis entre Ford y Exxon es

$$\begin{pmatrix} ... & &) \times \\ \begin{pmatrix} 1000{,}5090 & 23{,}0179 & 1360{,}6444 \\ 23{,}0179 & 1{,}4299 & 41{,}0892 \\ 1360{,}6444 & 41{,}0892 & 2980{,}4898 \end{pmatrix}^{-1} & \begin{pmatrix} \\ \vdots \\ \\ \end{pmatrix} \end{pmatrix}$$

Similitud a partir de distancia:

A partir de una distancia se puede calcular una similitud:

$$s(A, B) = \frac{1}{1 + d(A, B)} .$$

1.3.2. VARIABLES BINARIAS

Una valor binario (0 ó 1) representa la presencia (1) o ausencia (0) de determinada característica. Ahora suponemos que $X_1, \ldots, X_p$ son variables binarias, con lo cual

$$A = (a_1 \ldots, a_p), \quad B = (b_1 \ldots, b_p), \quad a_k, b_k \in \{0, 1\}, \quad k = 1, \ldots, p.$$

Medidas de distancia

Distancia euclídea:

$$d_e(A, B) = \sqrt{\sum_{k=1}^{p} (b_k - a_k)^2} .$$

Se verifica

$$(b_k - a_k)^2 = \begin{cases} ..., & si\ b_k = a_k , \\ ..., & si\ b_k \neq a_k . \end{cases}$$

La distancia euclídea es el número de componentes de A y B que no coinciden.

Representamos los datos en una tabla de contingencia

		B		
		1	0	Total
A	1	a	b	$a+b$
	0	c	d	$c+d$
	Total	$a+c$	$b+d$	$t=a+b+c+d$

Según la tabla, la distancia euclídea es ... . Las medidas de distancia que aparecen en SPSS son:

Distancia euclídea	...
Distancia euclídea al cuadrado	...
Diferencia de tamaño	$\frac{(b-c)^2}{t^2}$
Diferencia de configuración	$\frac{bc}{t^2}$
Varianza	$\frac{b+c}{4t}$
Lance y Williams	$\frac{b+c}{2a+b+c}$

Medidas de similitud:

Coeficiente	Descripción
$\frac{a+d}{t}$	Igual peso a coincidencias 1-1 y 0-0.
$\frac{2(a+d)}{2(a+d)+b+c}$	Doble peso a las coincidencias.
$\frac{a+d}{a+d+2(b+c)}$	Doble peso a las no coincidencias.
$\frac{a}{t}$	No se tienen en cuenta las coincidencias 0-0 en el numerador.
$\frac{a}{a+b+c}$	No se tienen en cuenta las coincidencias 0-0.
$\frac{2a}{2a+b+c}$	No se tienen encuenta las coincidencias 0-0, y se da doble peso a las coincidencias 1-1.
$\frac{a}{a+2(b+c)}$	No se tienen encuenta las coincidencias 0-0, y se da doble peso a las no coincidencias.
$\frac{a}{b+c}$	Cociente de coincidencias y no coincidencias, con exclusión de coincidencias 0-0.

El primer coeficiente que aparece en la tabla se conoce como el *coeficiente de comparación simple.*

Ejemplo 1.23 Queremos medir la similaridad entre dos individuos según su actitud positiva (1) o negativa (0) hacia la compra de 10 artículos

	1	2	3	4	5	6	7	8	9	10
Indiv. 1	1	0	1	1	0	0	0	1	1	0
Indiv. 2	1	0	0	1	1	0	0	1	0	0

Construimos la tabla de contingencia de ambos individuos

		Indiv. 2	
		1 0	Total
Indiv. 1	1		
	0		
	Total		

El coeficiente de comparación simple vale ... .

1.3.3. VARIABLES NOMINALES O CATEGÓRICAS

Ejemplo 1.24 Medimos las variables X_1= "tipo de whisky", X_2= "tipo de botella" y X_3= "región de fabricación" a dos marcas de whisky escocés A y B. La variable X_1 toma valores m="puro de malta", b="mezclado" (en Inglés blended), y c="cereales diferentes de la cebada", y X_2 toma valores s="standard", cc="cilíndrica corta", cl="cilíndrica larga" y c="cuadrada". Por último, la procedencia del whisky (X_3) puede ser h="Highlands", l="lowlands" y wi="western islands". Supongamos que los valores obtenidos son

$$A = (m, s, h), \;\; B = (m, cc, wi)\,.$$

Una similaridad entre los whiskies A y B sería

$$s(A, B) = ... \quad .$$

Ahora suponemos que $X_1, \ldots, X_p$ son nominales, de manera que estas variables pueden tomar valores de un conjunto de categorías. Es decir, ahora deseamos comparar la similaridad de $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$, donde a_k y b_k son ciertas categorías de la variable X_k, $k = 1, \ldots, p$.

Una similaridad sencilla entre A y B es la proporción de coincidencias en las coordenadas de A y B.
Si para la coordenada k-ésima, definimos

$$s_k(A,B) = 1 \text{ si } a_k = b_k, \quad \text{y} \quad s_k(A,B) = 0 \text{ si } a_k \neq b_k,$$

entonces esta similaridad es

$$s(A,B) = \frac{1}{p}\sum_{k=1}^{p} s_k(A,B)\,.$$

En lugar de asignar el valor 0 a $s_k(A,B)$ si las coordenadas a_k y b_k no coinciden, se puede asignar un valor entre 0 y 1 en función del grado de semejanza entre a_k y b_k.

Ejemplo 1.25 En el Ejemplo 1.24, supongamos que la similitud entre las categorías de la variable X_2 se puede cuantificar mediante los coeficientes siguientes

	s	cc	cl	c
s	1	0.5	0.5	0
cc	0.5	1	0.3	0
cl	0.5	0.3	1	0
c	0	0	0	1

En este caso, para

$$A = (m, s, h) \text{ y } B = (m, cc, wi)$$

tenemos

$$s_1(A,B) = ..., \quad s_2(A,B) = ... \quad \text{y} \quad s_3(A,B) = ...,$$

con lo cual la similaridad entre A y B es

$$s(A,B) = ... \qquad .$$

1.3.4. VARIABLES ORDINALES

Supongamos ahora que las m categorías $c_1, c_2, \ldots, c_m$ de una de las variables (X_k) están ordenadas de forma natural. Entonces se construyen $m-1$ variables dummy, con los siguientes valores

Categoría	I_1	I_2	$\cdots$	I_{m-1}
c_1	0	0	$\cdots$	0
c_2	1	0	$\cdots$	0
c_3	1	1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$
c_m	1	1	$\cdots$	1

Para los sujetos $A = (a_1, \ldots, a_p)$ y $B = (b_1, \ldots, b_p)$, los elementos k-ésimos son ciertas categorías de la variable X_k; por ejemplo $a_k = c_i$ y $b_k = c_j$. Se construye una tabla de contingencia de ambas categorías para las variables I_1, I_2, ..., I_{m-1},

		c_j		
		1	0	Total
c_i	1	a	b	$a+b$
	0	c	d	$c+d$
	Total	$a+c$	$b+d$	$t=a+b+c+d$

Así, el coeficiente de comparación simple nos proporciona un coeficiente de similaridad para la variable X_k, llamado $s_k(A, B)$. Promediando para todas las variables, obtenemos la similaridad entre A y B.

Ejemplo 1.26 Supongamos que en el ejemplo anterior, la variable X_2 es "altura de la botella", con valores p="pequeña", s="standard", l="larga" y el="extra larga". Los whiskies A y B toman los valores

$$A = (m, s, h),\ B = (m, p, wi)\,.$$

Para la variable X_2 con cuatro categorías, construimos tres variables dummy de la forma

Categoría	I_1	I_2	I_3
p			
s			
l			
el			

Así, la tabla de contingencia sería

		s		
		1	0	Total
p	1			
	0			
	Total			

El coeficiente de la segunda componente es $s_2(A, B) = ...$. La similaridad entre A y B es

$$s(A, B) = ... \qquad .$$

Disimilaridades a partir de similaridades

Disimilaridades se pueden obtener a partir de similaridades de varias formas, entre ellas las siguientes

$$\delta(A, B) = 1 - s(A, B)\,,$$
$$\delta(A, B) = c - s(A, B), \text{ para alguna constante } c\,,$$
$$\delta(A, B) = \sqrt{2(1 - s(A, B))}\,.$$