ÍNDICE

3.6. Formas cuadráticas

Siendo **A** una matriz cuadrada y simétrica y **x** un vector de p elementos no nulos, se llama forma cuadrática a la expresión:

$$Q=\mathbf{x}'\mathbf{A}\mathbf{x}=\begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}=\sum a_{i1}x_i x_1+\sum a_{i2}x_i x_2+\ldots+\sum a_{ip}x_i x_p=$$

$$=\sum_i\sum_j a_{ij}x_i x_j=\sum_i a_{ii}x_{ii}^2+\sum_{i\neq j}a_{ij}x_i x_j=\sum_i a_{ii}x_{ii}^2+\sum_{i<j}(a_{ij}+a_{ji})x_i x_j \qquad (1.14)$$

Ejemplo para $\mathbf{A}_{2\times 2}$: $Q=\mathbf{x}'\mathbf{A}\mathbf{x}=\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=a_{11}x_1^2+(a_{12}+a_{21})x_1x_2+a_{22}x_2^2$

Como se observa, $\mathbf{x}'\mathbf{A}\mathbf{x}$ es una función cuadrática de las **x** e incluye todos los elementos posibles de segundo orden.

Propiedades:

1. Para $\mathbf{x} = 0$ todas las formas cuadráticas Q son 0.
2. Si $Q > 0$ para todo $\mathbf{x} \neq 0$, entonces $\mathbf{x}'\,\mathbf{A}\,\mathbf{x}$ (y por tanto **A**) es *definida positiva*. Donde **A** es *regular*, $r(\mathbf{A}) = p$ y todos sus *autovalores* son positivos.
3. Si $Q \geq 0$ para todo **x**, entonces **A** es *semidefinida positiva*. Donde **A** es *singular*, $r(\mathbf{A}) < p$ y sus *autovalores* son positivos con uno al menos nulo.
4. **A** sería *definida negativa* ($r(\mathbf{A}) = p$ y sus $\lambda_j < 0$) si $-Q$ es definida positiva y *semidefinida negativa* ($r(\mathbf{A}) < p$ y sus $\lambda_j \leq 0$) si $-Q$ es semidefinida positiva.
5. Si **A** es definida positiva con autovalores $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_p \geq 0$ y autovectores $a_1, a_2, \ldots, a_p$, entonces Q es máxima para el máximo valor de λ con la restricción $\mathbf{x}'\mathbf{x} = 1$. Es decir:

$$Q=\mathbf{x}'\,\mathbf{A}\,\mathbf{x}=\mathbf{x}'\,\lambda_i\,\mathbf{x}=\lambda_i\,\mathbf{x}'\mathbf{x}=\lambda_i \qquad (1.15)$$

Ejemplo 23: $\mathbf{A}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; $Q=\mathbf{x}'\mathbf{A}\mathbf{x}=x_1^2+x_2^2$; $Q>0$: *definida positiva*

Obsérvese que: $|\mathbf{A}| = 1$; $r(\mathbf{A}) = p = 2$.

$$|\mathbf{A}-\lambda\,\mathbf{I}|=\begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix}=(1-\lambda)(1-\lambda)=0;\quad \lambda_1=\lambda_2=1\,.$$

Ejemplo 24: $\mathbf{B}=\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$; $Q=x_1^2+x_2^2-2x_1x_2=(x_1-x_2)^2$; $Q\geq 0$: *semidefinida positiva*

Obsérvese que: $|\mathbf{B}| = 0$; $r(\mathbf{B}) = 1$.

$$|\mathbf{B}-\lambda\,\mathbf{I}| = \begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)(1-\lambda)\text{ - }1 = \lambda(\lambda-2) = 0; \quad \lambda_1 = 2;\ \lambda_2 = 0.$$

Las formas cuadráticas tienen muchos usos en el análisis multivariante.

4. Vectores y estadísticos

Los estadísticos descriptivos pueden expresarse mediante vectores. La siguiente tabla resume la forma matricial de algunos estadísticos y de las matrices que facilitan su cálculo:

	Estadístico	Forma matricial
Media	$\bar{\mathbf{x}} = \frac{1}{n}\Sigma X_i$	$\bar{\mathbf{x}}' = \frac{1}{n}\mathbf{1}'\mathbf{X}^*$
Puntuaciones diferenciales	$x_i = X_i - \bar{\mathrm{X}}$	$\mathbf{X} = \mathbf{X}^* - \mathbf{1}\bar{\mathbf{x}}'$
Varianza	$s^2 = \frac{1}{n-1}\Sigma x_i^2$	$\mathbf{s}^2 = \frac{1}{n-1}\mathbf{x}'\mathbf{x} = \frac{1}{n-1}\lvert\mathbf{x}\rvert$
Desviación típica	$s^2 = \sqrt{\frac{1}{n-1}\Sigma x_i^2}$	$\mathbf{s} = \left(\frac{1}{n-1}\right)^{1/2}\lVert\mathbf{x}\rVert$
Covarianza	$s_{xy} = \frac{1}{n-1}\Sigma\, x_i y_i$	$\mathbf{s}_{xy} = \frac{1}{n-1}\mathbf{x}'\mathbf{y} = \frac{1}{n-1}\lvert\mathbf{xy}\rvert$
Correlación	$r_{xy} = \frac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2 \Sigma y_i^2}}$	$r_{xy} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{(\mathbf{x}'\mathbf{x})(\mathbf{y}'\mathbf{y})}} = \frac{\lvert\mathbf{xy}\rvert}{\lVert\mathbf{x}\rVert\lVert\mathbf{y}\rVert} = cos\theta_{xy}$
Matriz de covarianzas		$\mathbf{S} = \frac{1}{n-1}\mathbf{X}'\,\mathbf{X}$
Matriz de varianzas		$\mathbf{D} = \begin{bmatrix} s_1^2 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & s_p^2 \end{bmatrix}$
Matriz de puntuaciones típicas		$\mathbf{Z} = \mathbf{X}\mathbf{D}^{-1/2} = \mathbf{X}\begin{bmatrix} 1/s_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1/s_p \end{bmatrix}$
Matriz de correlaciones		$\mathbf{R} = \frac{1}{n-1}\mathbf{Z}'\mathbf{Z} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2}$

S y **R** son matrices gramianas pues se basan en sumas de cuadrados y productos cruzados.

Las matrices **S** y **R** se relacionan mediante las siguientes fórmulas:

$$\mathbf{R} = \mathbf{D}^{-1/2}\,\mathbf{S}\,\mathbf{D}^{-1/2}$$
$$\mathbf{S} = \mathbf{D}^{1/2}\,\mathbf{R}\,\mathbf{D}^{1/2} \qquad (1.16)$$

Como **S** y **R** son cuadradas y simétricas y $\mathbf{D}^{1/2}$ es una matriz regular, las matrices **S** y **R** son *equivalentes*. Ello implica que $r(\mathbf{S}) = r(\mathbf{R})$.

Ejemplo 25: A continuación se presenta un ejemplo del cálculo de la covarianza y la correlación para dos variables medidas en tres sujetos:

Sujeto	X_1	X_2
1	3	6
2	5	6
3	10	12

Matriz de datos (en puntuaciones directas): $\mathbf{X}^* = \begin{bmatrix} 3 & 6 \\ 5 & 6 \\ 10 & 12 \end{bmatrix}$

Medias: $\bar{\mathbf{x}}' = \frac{1}{n}\mathbf{1}'\mathbf{X}^* = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 6 \\ 5 & 6 \\ 10 & 12 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 18 & 24 \end{bmatrix} = \begin{bmatrix} 6 & 8 \end{bmatrix}$

Puntuaciones diferenciales: $\mathbf{X} = \mathbf{X}^* - \mathbf{1}\bar{\mathbf{x}}' = \begin{bmatrix} 3 & 6 \\ 5 & 6 \\ 10 & 12 \end{bmatrix} - \begin{bmatrix} 6 & 8 \\ 6 & 8 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ -1 & -2 \\ 4 & 4 \end{bmatrix}$

Varianzas: $\mathbf{s}_1^2 = \frac{1}{n-1}|\mathbf{x}_1| = \frac{1}{2}\begin{bmatrix} -3 & -1 & 4 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \\ 4 \end{bmatrix} = 13$

$$\mathbf{s}_2^2 = \frac{1}{n-1}|\mathbf{x}_2| = \frac{1}{2}\begin{bmatrix} -2 & -2 & 4 \end{bmatrix}\begin{bmatrix} -2 \\ -2 \\ 4 \end{bmatrix} = 12$$

Covarianza: $\mathbf{s}_{xy} = \frac{1}{n-1}|\mathbf{x}_1\mathbf{x}_2| = \frac{1}{2}\begin{bmatrix} -3 & -1 & 4 \end{bmatrix}\begin{bmatrix} -2 \\ -2 \\ 4 \end{bmatrix} = 12$

Matriz de covarianzas: $\mathbf{S} = \frac{1}{n-1}\mathbf{X}'\mathbf{X} = \frac{1}{2}\begin{bmatrix} -3 & -1 & 4 \\ -2 & -2 & 4 \end{bmatrix}\begin{bmatrix} -3 & -2 \\ -1 & -2 \\ 4 & 4 \end{bmatrix} = \begin{bmatrix} 13 & 12 \\ 12 & 12 \end{bmatrix}$

Correlación: $r_{xy} = \frac{|\mathbf{x}_1\mathbf{x}_2|}{\|\mathbf{x}_1\|\|\mathbf{x}_2\|} = \frac{24}{(5.10)(4.90)} = 0.96$

Matriz de correlaciones:

$$\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} = \begin{bmatrix} 1/\sqrt{13} & 0 \\ 0 & 1/\sqrt{12} \end{bmatrix} \begin{bmatrix} 13 & 12 \\ 12 & 12 \end{bmatrix} \begin{bmatrix} 1/\sqrt{13} & 0 \\ 0 & 1/\sqrt{13} \end{bmatrix} = \begin{bmatrix} 1 & 0.96 \\ 0.96 & 1 \end{bmatrix}$$

5. Combinaciones lineales

Las técnicas multivariantes se formulan mediante combinaciones lineales por lo que es necesario comprender su definición y propiedades.

Considérese la siguiente combinación lineal:

$$\mathbf{y} = \mathbf{X}\,\mathbf{a}$$

La variable aleatoria **y** es una transformación o combinación lineal de **X** mediante **a**. Donde **a**' = $[a_1, ..., a_p]$ es un vector de constantes, **X** una matriz de puntuaciones de n sujetos en p variables (siendo $\boldsymbol{\mu}'$ su vector de medias). La media y varianza de **y** es:

$$\mathrm{E}\,(\mathbf{y}) = \boldsymbol{\mu}'\,\mathbf{a}$$

$$Var\,(\mathbf{y}) = \frac{1}{n-1}\mathbf{y}'\mathbf{y} = \frac{1}{n-1}(\mathbf{Xa})'(\mathbf{Xa}) = \frac{1}{n-1}\mathbf{a}'\mathbf{X}'\mathbf{Xa} =$$

$$= \mathbf{a}'\frac{1}{n-1}\mathbf{X}'\mathbf{Xa} = \mathbf{a}'\mathbf{Sa} \qquad (1.17)$$

A continuación se presenta un ejemplo para ilustrar la fórmula(1.17) en el caso en que p=2:

$$\mathrm{Var}\,(X_1\,a_1 + X_2\,a_2) = \mathrm{Var}\,(X_1\,a_1) + \mathrm{Var}\,(X_2\,a_2) + 2\,\mathrm{Cov}\,(X_1\,a_1, X_2\,a_2) =$$
$$= a_1^2\,\mathrm{Var}\,(X_1) + a_2^2\,\mathrm{Var}\,(X_2) + 2\,a_1\,a_2\,\mathrm{Cov}\,(X_1,\ X_2)$$

Como se observa, la varianza de una combinación lineal es una forma cuadrática. En el caso en que **a** fuese un vector normalizado (donde **a' a** = 1), la varianza de **y** queda como:

$$Var\,(\mathbf{y}) = \mathbf{a}'\,\mathbf{S}\,\mathbf{a} = \mathbf{a}'\,\lambda\,\mathbf{a} = \lambda$$

Las ecuaciones de (1.17) pueden generalizarse al caso $\boldsymbol{Y}$ = **X A**. Donde **A** es una matriz de constantes de orden $n \times p$, y la media y varianza de $\boldsymbol{Y}$ es:

$$\mathrm{E}\,(\boldsymbol{Y}) = \boldsymbol{\mu}'\,\mathbf{A}$$
$$\mathrm{Var}\,(\boldsymbol{Y}) = \mathbf{A}'\,\mathbf{S}\,\mathbf{A} \qquad (1.18)$$

A continuación se comentan algunas propiedades de las matrices **S** y **R**. En primer lugar ambas son *semidefinidas positivas*. Puesto que toda varianza ha de ser no negativa:

$$\mathrm{Var}\,(\mathbf{X}\,\mathbf{a}) \geq 0 \qquad \text{para todo } \mathbf{a}$$

Como Var (**X a**) = **a' S a**, entonces **S** tiene que ser, al menos, *semidefinida positiva*. **S** y **R** son matrices equivalentes pues en las fórmulas que las relacionan en (1.16) la matriz $\mathbf{D}^{1/2}$ es regular. Por tanto, **R** también es *semidefinida positiva*.

En segundo lugar, puesto que las matrices **S** y **R** son equivalentes, el rango de **S** es el mismo que el de **R**. Este rango puede ser menor o igual que p. Si $r(\mathbf{S}) = p$, entonces **S** y **R** serán *definidas positivas* pues Var (**X a**) = **a' S a** es mayor que cero para todo $\mathbf{a} \neq 0$. Sin embargo, si $r(\mathbf{S}) < p$ entonces **S** y **R** serán singulares y ello indicará una restricción de linealidad en los componentes de **X**. Esto implica que existe un vector $\mathbf{a} \neq 0$ tal que **X a** es igual a una constante. Entonces, $Var(\mathbf{X\,a}) = \mathbf{a'\,S\,a}$ será cero, indicando que la matriz **S** es *semidefinida positiva* en lugar de definida positiva.

Para ilustrar este último punto, supóngase que $p = 3$ y que existe una restricción de linealidad en las tres variables tal que $X_1 = X_2 + X_3$. Entonces, $\mathrm{Var}(X_1 - X_2 - X_3) = 0$ y el vector **a'** = [1, -1, -1]. En este caso, una de las tres variables es redundante y por tanto la dimensionalidad es 2 en lugar de 3. Esto se refleja en el rango de **S** que también será 2. Según esta propiedad, el rango de **S** es un indicador útil para establecer la dimensionalidad del problema, siendo $[p - r(\mathbf{S})]$ el número de restricciones lineales independientes en los componentes de **X**. De este modo, cuando $r(\mathbf{S}) < p$ se dice que los componentes de **X** son linealmente dependientes.

Ejemplo 26: Sean las variables X_1 y X_2 y su matriz de covarianzas $\mathbf{S} = \begin{bmatrix} 4 & 3 \\ 3 & 9 \end{bmatrix}$.

Si se forman las combinaciones lineales $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$, la matriz de covarianzas para **Y** es:

$$\mathrm{Var}(\mathbf{Y}) = \mathbf{A'SA} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 19 & -5 \\ -5 & 7 \end{bmatrix}$$

Y la matriz de correlaciones:

$$\mathbf{R} = \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} = \begin{bmatrix} 1/\sqrt{19} & 0 \\ 0 & 1/\sqrt{7} \end{bmatrix} \begin{bmatrix} 19 & -5 \\ -5 & 7 \end{bmatrix} \begin{bmatrix} 1/\sqrt{19} & 0 \\ 0 & 1/\sqrt{7} \end{bmatrix} = \begin{bmatrix} 1 & -0.43 \\ -0.43 & 1 \end{bmatrix}$$

6. El álgebra de matrices y el lenguaje MATRIX del SPSS

Las operaciones con matrices son complejas. Existen diversos paquetes informáticos que evitan su cálculo a mano. A continuación se introduce *el lenguaje MATRIX* del programa SPSS, uno de los más empleados en las ciencias sociales.

El lenguaje MATRIX no se encuentra en los menús desplegables del programa SPSS. Para utilizarlo es necesario acudir a ventanas de sintaxis desde donde se escribe la operación que se desea realizar y se ejecuta.