CdL in Scienze Statistiche ed Economiche - Università degli Studi di Milano-Bicocca

Esercitazione : Dati centrati e standardizzati

Esercitatrice: Chiara Gaia Magnani

Example 0.1. *(a) Calcolare la traccia della matrice di centramento* $\underset{n\times n}{H}$
(b) Calcolare $\underset{n\times n}{H}\underset{n\times 1}{1}$
(c) Si supponga che $\underset{n\times 1}{a}$ *è un vettore i cui elementi sommano 0. Calcolare* $\underset{n\times n}{H}\underset{n\times 1}{a}$

Dimostrazione. (a)

$$\mathrm{tr}(H) = \mathrm{tr}(I - \frac{1}{n}11') = \mathrm{tr}(I) - \frac{1}{n}\mathrm{tr}(11') = n - \frac{1}{n}n = n - 1$$

(b)

$$H1 = (I - \frac{1}{n}11')1 = 1 - \frac{1}{n}11'1 = 1 - \frac{1}{n}1n = \underset{n\times 1}{0}$$

(c)

$$Ha = (I - \frac{1}{n}11')a = a - \frac{1}{n}11'a = a - \frac{1}{n}1\sum_{i=1}^{n} a_i = \underset{n\times 1}{a}$$

□

Example 0.2. *Sia* $\underset{n\times n}{J} = \frac{1}{n}11'$*, quindi* $H = I - J$*.*
(a) Calcolare $\underset{n\times n}{J}\underset{n\times 1}{a}$ *per un generico vettore* a*.*
(b) Si dimostri che $\underset{n\times n}{J}$ *è una matrice idempotente.*

Dimostrazione. (a)

$$Ja = \frac{1}{n}11'a = \frac{1}{n}1\sum_{i=1}^{n} a_i = \begin{bmatrix} \bar{a} \\ \bar{a} \\ \dots \\ \bar{a} \end{bmatrix}$$

(b) J è una matrice simmetrica, i.e. $J' = (\frac{1}{n}11')' = \frac{1}{n}11' = J$. Inoltre

$$JJ = \frac{1}{n}11'\frac{1}{n}11' = \frac{1}{n^2}11'11' = \frac{1}{n^2}1n1' = \frac{1}{n}11' = J$$

□

Example 0.3. *Si consideri una matrice dei dati* $\underset{n\times p}{X}$ *con vettore delle medie* $\bar{x} = [3, 2, -2, 0]'$ *e matrice di varianza e covarianza*

$$S = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \tag{1}$$

Si consideri la matrice

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \tag{2}$$

1. Calcolare il vettore delle medie di $Y = XA'$*.*
2. Calcolare la matrice di varianza e covarianza di $Y = XA'$*.*
3. Quali coppie di colonne della matrice Y *hanno correlazione pari a zero?.*

Dimostrazione. 1 . Abbiamo

$$\bar{y} = (1/n)Y'1 = (1/n)(XA')'1 = (1/n)AX'1 = A\bar{x}$$

$$\bar{y} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix} \tag{3}$$

2. Abbiamo

$$S^Y = (1/n)Y'HY = (1/n)(XA')'HXA' = (1/n)AX'HXA' = ASA'$$

quindi

$$S^Y = ASA' = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 36 \end{bmatrix} \tag{4}$$

3.

Dalla matrice di varianza e covarianza si può dedurre facilmente che tutte le possibili coppie di colonne di Y hanno covarianza nulla, e quindi correlazione nulla. □

Example 0.4. *Si consideri una matrice dei dati* $\underset{n\times p}{X}$ *con vettore delle medie* $\bar{x} = [2, 4, -1, 3, 0]'$ *e matrice di varianza e covarianza*

$$S = \begin{bmatrix} 4 & -1 & 1/2 & -1/2 & 0 \\ -1 & 3 & 1 & -1 & 0 \\ 1/2 & 1 & 6 & 1 & -1 \\ -1/2 & -1 & 1 & 4 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix} \tag{5}$$

si consideri la partizione della matrice X *in due sottomatrici* $A = [x_1\, x_2]$ *e* $B = [x_3\, x_4\, x_5]$*, dove* A *è la matrice corrispondente alle prime due colonne di* X *mentre* B *è la matrice corrispondente alle ultime tre colonne di* X.

Calcolare le seguenti quantità:
1. vettore delle medie di A *e vettore delle medie di* B
2. matrice di varianza e covarianza di A *e matrice di varianza e covarianza di* B
3. La covarianza tra la prima colonna di A *e la prima colonna di* B.

Dimostrazione. 1. I vettori delle medie di A e B sono i corrispondenti sottovettori del vettore delle medie di X:

$$\bar{x}^A = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \;\; e \;\; \bar{x}^B = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \tag{6}$$

2. Le matrici di varianza e covarianza di A e B sono i corrispondenti blocchi di dimensioni rispettivamente 2x2 e 3x3 sulla diagonale di S:

$$S^A = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \;\; e \;\; S^B = \begin{bmatrix} 6 & 1 & -1 \\ 1 & 4 & 0 \\ -1 & 0 & 2 \end{bmatrix} \tag{7}$$

La covarianza tra la prima colonna di A e la prima colonna di B è pari a $\{S\}_{1,3} = 1/2$

□