Matrice dei dati centrati e standardizzati

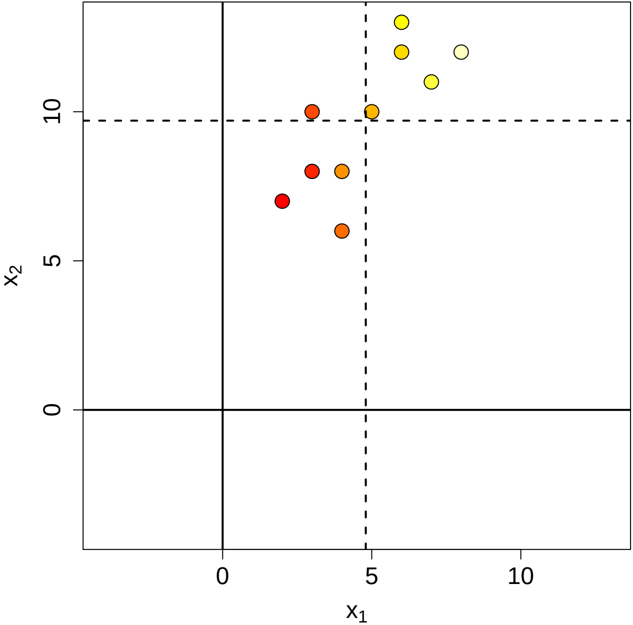
Analisi Esplorativa

Aldo Solari

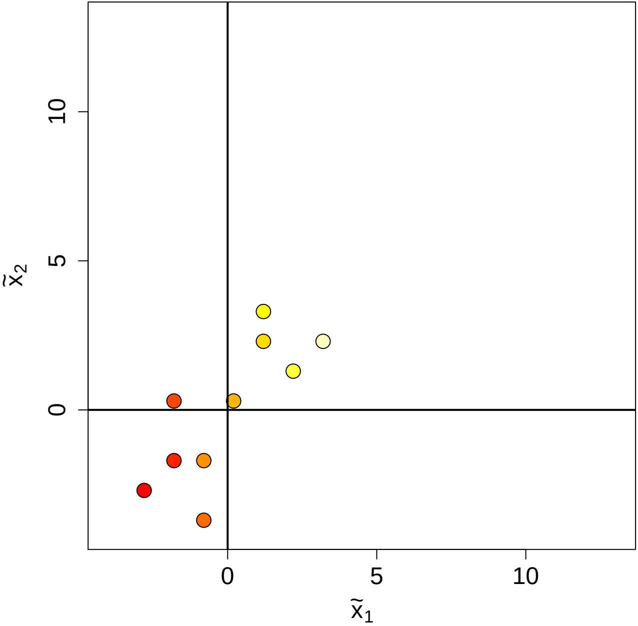
Esempio

$$\underset{10\times 2}{X} = \begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 3 & 10 \\ 4 & 6 \\ 4 & 8 \\ 5 & 10 \\ 6 & 12 \\ 6 & 13 \\ 7 & 11 \\ 8 & 12 \end{bmatrix}$$

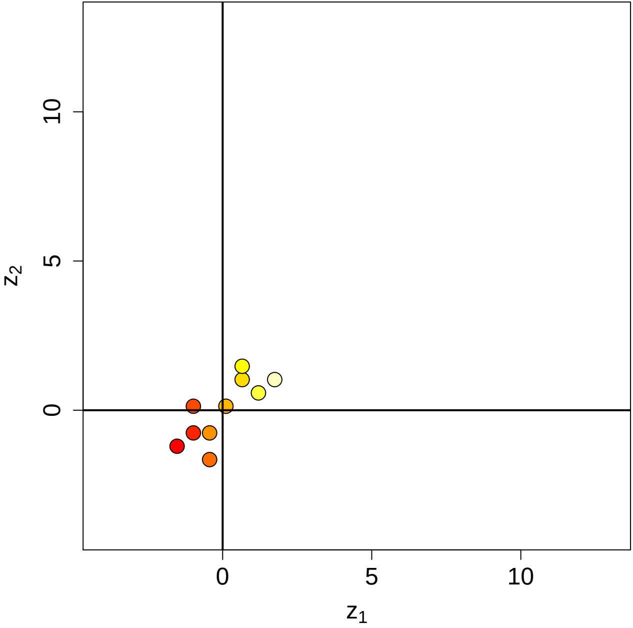
Matrice originale X



Matrice centrata $\tilde{X}$ (traslazione)



Matr. stand. Z (compressione/dilatazione)



Dati centrati e dati standardizzati

Abbiamo appena visto che possiamo trasformare (linearmente) la matrice dei dati originali $\underset{n\times p}{X}$ per ottenere

- La matrice dei dati centrati

$$\underset{n\times p}{\tilde{X}} = \left(\underset{n\times n}{I} - \frac{1}{n} \underset{n\times 1}{1} \underset{1\times n}{1'} \right) \underset{n\times p}{X}$$

- La matrice dei dati standardizzati

$$\underset{n\times p}{Z} = \underset{n\times p}{\tilde{X}} \operatorname{diag}\left(\frac{1}{\sqrt{s_{11}}}, \ldots, \frac{1}{\sqrt{s_{pp}}} \right)$$

Come sono

- il vettore delle medie
- la matrice di varianza/covarianza
- la matrice di correlazione

dei dati centrati e dei dati standardizzati?

Dati centrati e dati standardizzati

Matrice dei dati	Vettore delle medie	Matrice di varianze/covarianze	Matrice di correlazione
$\underset{n\times p}{X}$	$\underset{p\times 1}{\bar{x}}$	$\underset{p\times p}{S}$	$\underset{p\times p}{R}$
$\underset{n\times p}{\tilde{X}}$	$\underset{p\times 1}{0}$	$\underset{p\times p}{S^{\tilde{X}}} = \underset{p\times p}{S}$	$\underset{p\times p}{R^{\tilde{X}}} = \underset{p\times p}{R}$
$\underset{n\times p}{Z}$	$\underset{p\times 1}{0}$	$\underset{p\times p}{S^{Z}} = \underset{p\times p}{R}$	$\underset{p\times p}{R^{Z}} = \underset{p\times p}{R}$

$\bar{x}$, S e R in forma matriciale

- $\underset{p\times 1}{\bar{x}} = \frac{1}{n} \underset{p\times n}{X'} \underset{n\times 1}{1}$
- $\underset{p\times p}{S} = \frac{1}{n} \underset{p\times n}{\tilde{X}'} \underset{n\times p}{\tilde{X}}$
- $\underset{p\times p}{R} = \frac{1}{n} \underset{p\times n}{Z'} \underset{n\times p}{Z}$

Proprietà di H, S e R

- La matrice di centramento $\underset{n\times n}{H} = \underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'}$ è simmetrica e idempotente
- La matrice di varianze/covarianze $\underset{p\times p}{S}$ e la matrice di correlazione $\underset{p\times p}{R}$ sono semidefinite positive

Per la dimostrazione di tutti questi risultati, si vedano le prossime slides.

Outline

Vettore delle medie in forma matriciale

$$\boxed{\underset{p\times 1}{\bar{x}} = \frac{1}{n} \underset{p\times n}{X'} \underset{n\times 1}{1}}$$

$$
\underset{p\times 1}{\bar{x}} = \begin{bmatrix} \bar{x}_1 \\ \cdots \\ \bar{x}_j \\ \cdots \\ \bar{x}_p \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n x_{i1} \\ \cdots \\ \sum_{i=1}^n x_{ij} \\ \cdots \\ \sum_{i=1}^n x_{ip} \end{bmatrix}
$$

$$
= \frac{1}{n} \begin{bmatrix} \underset{1\times n}{1'} \underset{n\times 1}{x_1} \\ \cdots \\ \underset{1\times n}{1'} \underset{n\times 1}{x_j} \\ \cdots \\ \underset{1\times n}{1'} \underset{n\times 1}{x_p} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \underset{1\times n}{(\, x_1 \,)'} \underset{n\times 1}{1} \\ \cdots \\ \underset{1\times n}{(\, x_j \,)'} \underset{n\times 1}{1} \\ \cdots \\ \underset{1\times n}{(\, x_p \,)'} \underset{n\times 1}{1} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \underset{1\times n}{(\, x_1 \,)'} \\ \cdots \\ \underset{1\times n}{(\, x_j \,)'} \\ \cdots \\ \underset{1\times n}{(\, x_p \,)'} \end{bmatrix} \underset{n\times 1}{1}
$$

$$
= \frac{1}{n} \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{i1} & \cdots & x_{n1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ x_{1j} & x_{2j} & \cdots & x_{ij} & \cdots & x_{nj} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ x_{1p} & x_{2p} & \cdots & x_{ip} & \cdots & x_{np} \end{bmatrix} \underset{n\times 1}{1} = \frac{1}{n} \underset{p\times n}{X'} \underset{n\times 1}{1}
$$

Outline

Matrice dei dati centrati

$$\boxed{\underset{n\times p}{\tilde{X}} = \left(\underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'}\right)\underset{n\times p}{X} = \underset{n\times n}{H}\underset{n\times p}{X}}$$

dove

- $\underset{n\times n}{H}$ è la matrice di centramento
- $\underset{n\times n}{I}$ è la matrice identità (vedi Appendice)

$$
\begin{aligned}
\underset{n\times p}{\tilde{X}} &= \begin{bmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1j}-\bar{x}_j & \cdots & x_{1p}-\bar{x}_p \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2j}-\bar{x}_j & \cdots & x_{2p}-\bar{x}_p \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ x_{i1}-\bar{x}_1 & x_{i2}-\bar{x}_2 & \cdots & x_{ij}-\bar{x}_j & \cdots & x_{ip}-\bar{x}_p \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{nj}-\bar{x}_j & \cdots & x_{np}-\bar{x}_p \end{bmatrix} \\
&= \underset{n\times p}{X} - \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \vdots & \vdots & \vdots & \vdots \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \end{bmatrix} \\
&= \underset{n\times p}{X} - \underset{n\times 1}{1}\,\underset{1\times p}{\bar{x}'} \\
&= \underset{n\times p}{X} - \frac{1}{n}\underset{n\times 1}{1}\,\underset{1\times n}{1'}\,\underset{n\times p}{X} \\
&= \underset{n\times n}{I}\,\underset{n\times p}{X} - \frac{1}{n}\underset{n\times 1}{1}\,\underset{1\times n}{1'}\,\underset{n\times p}{X} \\
&= \left(\underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\,\underset{1\times n}{1'}\right)X \\
&= \underset{n\times n}{H}\,\underset{n\times p}{X}
\end{aligned}
$$

Outline

Matrice di centramento

$$\boxed{\underset{n\times n}{H} = \underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'}}$$

Matrice di centramento: proprietà

$\underset{n\times n}{H}$ è una matrice simmetrica

$$\underset{n\times n}{H} = \underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'} = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{bmatrix}$$

Una matrice quadrata $\underset{n\times n}{A}$ è simmetrica se $A = A'$; ovvero se $a_{ij} = a_{ji}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$; vedi Appendice

Matrice di centramento: proprietà

$\underset{n\times n}{H}$ è una matrice idempotente

Dimostrazione:

$$\begin{aligned}
\underset{n\times n}{H}\underset{n\times n}{H} &= (\underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'})(\underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'}) \\
&= \underset{n\times n}{I} - 2\frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'} + \frac{1}{n^2}(\underset{n\times 1}{1}\underset{1\times n}{1'})(\underset{n\times 1}{1}\underset{1\times n}{1'}) \\
&= \underset{n\times n}{I} - 2\frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'} + \frac{1}{n^2}\underset{n\times 1}{1}\, n\, \underset{1\times n}{1'} \\
&= \underset{n\times n}{I} - \frac{1}{n}\underset{n\times 1}{1}\underset{1\times n}{1'} = \underset{n\times n}{H}
\end{aligned}$$

$\square$

Una matrice quadrata $\underset{n\times n}{A}$ è detta idempotente se vale $\underset{n\times n}{A}\underset{n\times n}{A} = \underset{n\times n}{A}$; vedi Appendice.

Centrare la matrice dei dati centrati

Non produce alcun effetto:

Dimostrazione:

$$\underset{n\times n}{H}\underset{n\times p}{\tilde{X}} = \underset{n\times n}{H}\underset{n\times n}{H}\underset{n\times p}{X} = \underset{n\times p}{\tilde{X}}$$

$\square$

Outline

Matrice di varianze/covarianze

$$\boxed{\underset{p\times p}{S} = \frac{1}{n}\underset{p\times n}{\tilde{X}'}\underset{n\times p}{\tilde{X}} = \frac{1}{n}\underset{p\times n}{X'}\underset{n\times n}{H}\underset{n\times p}{X}}$$

$$
\begin{aligned}
n \underset{p\times p}{S} &= \begin{bmatrix} \underset{1\times n}{\tilde{x}_1'} \\ \underset{1\times n}{\tilde{x}_2'} \\ \vdots \\ \underset{1\times n}{\tilde{x}_j'} \\ \vdots \\ \underset{1\times n}{\tilde{x}_p'} \end{bmatrix} \begin{bmatrix} \underset{n\times 1}{\tilde{x}_1} & \underset{n\times 1}{\tilde{x}_2} & \cdots & \underset{n\times 1}{\tilde{x}_j} & \cdots & \underset{n\times 1}{\tilde{x}_p} \end{bmatrix} \\
&= \underset{p\times n}{\tilde{X}'} \underset{n\times p}{\tilde{X}} \\
&= \underset{n\times p}{X'} \underset{n\times n}{H'} \underset{n\times n}{H} \underset{n\times p}{X} \\
&= \underset{n\times p}{X'} \underset{n\times n}{H} \underset{n\times p}{X}
\end{aligned}
$$

Matrice di varianze/covarianze: proprietà

$\underset{p\times p}{S}$ è una matrice semidefinita positiva

Dimostrazione:

$$
\begin{aligned}
n \underset{1\times p}{a'} \underset{p\times p}{S} \underset{p\times 1}{a} &= \underset{1\times p}{a'} \underset{p\times n}{\tilde{X}'} \underset{n\times p}{\tilde{X}} \underset{p\times 1}{a} \\
&= \left(\underset{1\times p}{a'} \underset{p\times n}{X'} \underset{n\times n}{H'} \right) \left(\underset{n\times n}{H} \underset{n\times p}{X} \underset{p\times 1}{a} \right) \\
&= \underset{1\times n}{b'} \underset{n\times 1}{b} \\
&= \sum_{i=1}^{n} b_i^2 \geq 0 \quad \forall \underset{p\times 1}{a}
\end{aligned}
$$

dove $\underset{n\times 1}{b} = \underset{n\times n}{H} \underset{n\times p}{X} \underset{p\times 1}{a}$.

□

Una matrice simmetrica $\underset{p\times p}{B}$ è detta semidefinita positiva se vale $\underset{1\times p}{a'} \underset{p\times p}{B} \underset{p\times 1}{a} \geq 0 \quad \forall \underset{p\times 1}{a}$;
vedi Appendice

Matrice di varianze/covarianze: proprietà

La matrice di varianze/covarianze calcolata per $\underset{n\times p}{\tilde{X}}$ risulta uguale alla varianze/covarianze calcolata per per $\underset{n\times p}{X}$.

<u>Dimostrazione</u>:

$$\underset{p\times p}{S^{\tilde{X}}} = \frac{1}{n}\left(\underset{p\times n}{\tilde{X}'}\underset{n\times n}{H'}\right)\left(\underset{n\times n}{H}\underset{n\times p}{\tilde{X}}\right) = \frac{1}{n}\underset{p\times n}{\tilde{X}'}\underset{n\times p}{\tilde{X}} = \underset{p\times p}{S^{X}}$$

$\square$

La trasformazione centrante $X \mapsto \tilde{X}$ è una trasformazione lineare che fornisce una matrice $\underset{n\times p}{\tilde{X}}$ con vettore delle medie nullo e matrice di varianze/covarianze pari a quella dei dati originali $\underset{n\times p}{X}$.

Outline

Matrice dei dati standardizzati

$$\boxed{\underset{n\times p}{Z} = \underset{n\times p}{\tilde{X}}\,\underset{p\times p}{D}^{-1/2}}$$

dove $\underset{p\times p}{D}^{1/2} = \mathrm{diag}(\sqrt{s_{11}}, \ldots, \sqrt{s_{pp}})$ con

$$\mathrm{diag}(\sqrt{s_{11}}, \ldots, \sqrt{s_{pp}}) = \begin{bmatrix} \sqrt{s_{11}} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sqrt{s_{22}} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{s_{jj}} & \cdots 0 & \\ \vdots & \vdots & \ldots & \vdots & \ddots & \ldots \\ 0 & 0 & \cdots & 0 & \cdots & \sqrt{s_{pp}} \end{bmatrix}$$

Per la definizione e proprietà di una matrice diagonale, vedi Appendice

La matrice

$$\underset{p\times p}{D}^{-1/2} = \mathrm{diag}\left(\frac{1}{\sqrt{s_{11}}}, \ldots, \frac{1}{\sqrt{s_{pp}}}\right) = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ldots & \vdots & \ddots & \ldots \\ 0 & \cdots & 0 & \cdots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix}$$

è la matrice inversa di $\underset{p\times p}{D}^{1/2}$

Questo richiede che $s_{11}, \ldots, s_{pp}$ siano tutti diversi da 0.

Moltiplicare $\underset{n\times p}{\tilde{X}}$ da destra per $\underset{p\times p}{D^{-1/2}}$ equivale a moltiplicare la j-sima colonna di $\underset{n\times p}{\tilde{X}}$ per $\frac{1}{\sqrt{s_{jj}}}$;

$$
\begin{aligned}
\underset{n\times p}{Z} &= \begin{bmatrix}
\frac{x_{11}-\bar{x}_1}{\sqrt{s_{11}}} & \frac{x_{12}-\bar{x}_2}{\sqrt{s_{22}}} & \dots & \frac{x_{1j}-\bar{x}_j}{\sqrt{s_{jj}}} & \dots & \frac{x_{1p}-\bar{x}_p}{\sqrt{s_{pp}}} \\
\frac{x_{21}-\bar{x}_1}{\sqrt{s_{11}}} & \frac{x_{22}-\bar{x}_2}{\sqrt{s_{22}}} & \dots & \frac{x_{2j}-\bar{x}_j}{\sqrt{s_{jj}}} & \dots & \frac{x_{2p}-\bar{x}_p}{\sqrt{s_{pp}}} \\
\dots & \dots & \dots & \dots & \dots & \\
\frac{x_{n1}-\bar{x}_1}{\sqrt{s_{11}}} & \frac{x_{n2}-\bar{x}_2}{\sqrt{s_{22}}} & \dots & \frac{x_{nj}-\bar{x}_j}{\sqrt{s_{jj}}} & \dots & \frac{x_{np}-\bar{x}_p}{\sqrt{s_{pp}}}
\end{bmatrix} \\
&= \begin{bmatrix}
x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1j}-\bar{x}_j & \cdots & x_{1p}-\bar{x}_p \\
x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2j}-\bar{x}_j & \cdots & x_{2p}-\bar{x}_p \\
\dots & \dots & \dots & \dots & \dots & \\
x_{i1}-\bar{x}_1 & x_{i2}-\bar{x}_2 & \cdots & x_{ij}-\bar{x}_j & \cdots & x_{ip}-\bar{x}_p \\
\dots & \dots & \dots & \dots & \dots & \\
x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{nj}-\bar{x}_j & \cdots & x_{np}-\bar{x}_p
\end{bmatrix} D^{-1/2} \\
&= \tilde{X}D^{-1/2}
\end{aligned}
$$

Outline

Matrice di correlazione

$$\boxed{\underset{p\times p}{R} = \underset{p\times p}{D^{-1/2}}\,\underset{p\times p}{S}\,\underset{p\times p}{D^{-1/2}}}$$

- Moltiplicare $\underset{p\times p}{S}$ da sinistra per $\underset{p\times p}{D^{-1/2}}$ equivale a moltiplicare l'i-sima riga di $\underset{p\times p}{S}$ per $\frac{1}{\sqrt{s_{ii}}}$;
- Moltiplicare $\underset{p\times p}{S}$ da destra per $\underset{p\times p}{D^{-1/2}}$ equivale a moltiplicare la j-sima colonna di $\underset{p\times p}{S}$ per $\frac{1}{\sqrt{s_{jj}}}$;

Quindi

$$\underset{p\times p}{R} = D^{-1/2} S D^{-1/2} = \begin{bmatrix} \frac{s_{11}}{\sqrt{s_{11}}\sqrt{s_{11}}} & \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & \cdots & \frac{s_{1j}}{\sqrt{s_{11}}\sqrt{s_{jj}}} & \cdots & \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots \\ \frac{s_{i1}}{\sqrt{s_{ii}}\sqrt{s_{11}}} & \frac{s_{i2}}{\sqrt{s_{ii}}\sqrt{s_{22}}} & \cdots & \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}} & \cdots & \frac{s_{ip}}{\sqrt{s_{ii}}\sqrt{s_{pp}}} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots \\ \frac{s_{p1}}{\sqrt{s_{pp}}\sqrt{s_{11}}} & \frac{s_{p2}}{\sqrt{s_{pp}}\sqrt{s_{22}}} & \cdots & \frac{s_{pj}}{\sqrt{s_{pp}}\sqrt{s_{jj}}} & \cdots & \frac{s_{pp}}{\sqrt{s_{pp}}\sqrt{s_{pp}}} \end{bmatrix}$$

Matrice di correlazione: proprietà

$$\boxed{\underset{p\times p}{S} = \underset{p\times p}{D^{1/2}}\,\underset{p\times p}{R}\,\underset{p\times p}{D^{1/2}}}$$

<u>Dimostrazione</u>:

$$R = D^{-1/2} S D^{-1/2}$$
$$D^{1/2} R D^{1/2} = D^{1/2} D^{-1/2} S D^{-1/2} D^{1/2}$$

e visto che $D^{-1/2}$ è la matrice inversa di $D^{1/2}$, per definizione

$$\underset{p\times p}{D^{1/2}}\underset{p\times p}{D^{-1/2}} = \underset{p\times p}{D^{-1/2}}\underset{p\times p}{D^{1/2}} = \underset{p\times p}{I}$$

e quindi $\underset{p\times p}{D^{1/2}}\,\underset{p\times p}{R}\,\underset{p\times p}{D^{1/2}} = \underset{p\times p}{S}$

$\square$

Per la definizione di matrice inversa, vedi Appendice

Matrice di correlazione: proprietà

La matrice di varianze/covarianze calcolata per Z risulta uguale alla matrice di correlazione calcolata per X.

Dimostrazione:

$$\underset{p\times p}{S^Z} = \frac{1}{n}(Z'H')(HZ)$$
$$= \frac{1}{n}Z'Z$$
$$= \frac{1}{n}D^{-1/2}\tilde{X}'\tilde{X}D^{-1/2}$$
$$= D^{-1/2}SD^{-1/2} = \underset{p\times p}{R}^X$$

$\square$

La trasformazione scalante $X \mapsto Z$ è una trasformazione lineare che fornisce una matrice Z con vettore delle medie nullo e matrice di varianze e covarianze pari alla matrice di correlazione dei dati originali X.

Esempio

Matrice

$$\underset{4\times 2}{X} = \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix}$$

Vettore delle medie

$$\underset{2\times 1}{\bar{x}} = \frac{1}{4}\begin{bmatrix} 42 & 52 & 48 & 58 \\ 4 & 5 & 4 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 50 \\ 4 \end{bmatrix}$$

Matrice di centramento

$$\begin{aligned} \underset{4\times 4}{H} &= \underset{4\times 4}{I} - \frac{1}{4}\underset{1\times 4}{1}\,\underset{4\times 1}{1'} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1-1/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 1-1/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 1-1/4 & -1/4 \\ -1/4 & -1/4 & -1/4 & 1-1/4 \end{bmatrix} \end{aligned}$$

Esempio

Matrice dei dati centrati:

$$
\begin{aligned}
\underset{4\times 2}{\tilde{X}} &= HX \\
&= \begin{bmatrix} 1-1/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 1-1/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 1-1/4 & -1/4 \\ -1/4 & -1/4 & -1/4 & 1-1/4 \end{bmatrix} \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix} \\
&= \begin{bmatrix} -8 & 0 \\ 2 & 1 \\ -2 & 0 \\ 8 & -1 \end{bmatrix}
\end{aligned}
$$

Matrice di varianze/covarianze

$$
\begin{aligned}
\underset{2\times 2}{S} &= \frac{1}{4}\tilde{X}'\tilde{X} \\
&= \frac{1}{4} \begin{bmatrix} -8 & 2 & -2 & 8 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} -8 & 0 \\ 2 & 1 \\ -2 & 0 \\ 8 & -1 \end{bmatrix} = \begin{bmatrix} 34 & -1.5 \\ -1.5 & 0.5 \end{bmatrix}
\end{aligned}
$$

Esempio

Matrice di correlazione:

$$\underset{2\times 2}{R} = \underset{2\times 2}{D}^{-1/2} \underset{2\times 2}{S} \underset{2\times 2}{D}^{-1/2}$$

$$= \begin{bmatrix} 1/\sqrt{34} & 0 \\ 0 & 1/\sqrt{0.5} \end{bmatrix} \begin{bmatrix} 34 & -1.5 \\ -1.5 & 0.5 \end{bmatrix} \begin{bmatrix} 1/\sqrt{34} & 0 \\ 0 & 1/\sqrt{0.5} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1.5/(\sqrt{34}\sqrt{5}) \\ -1.5/(\sqrt{34}\sqrt{5}) & 1 \end{bmatrix}$$

Matrice dei dati standardizzati:

$$\underset{4\times 2}{Z} = \begin{bmatrix} -8 & 0 \\ 2 & 1 \\ -2 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{34} & 0 \\ 0 & 1/\sqrt{0.5} \end{bmatrix} = \begin{bmatrix} -8/\sqrt{34} & 0 \\ 2/\sqrt{34} & 1/\sqrt{0.5} \\ -2/\sqrt{34} & 0 \\ 8/\sqrt{34} & -1/\sqrt{0.5} \end{bmatrix}$$

Outline

Matrice trasposta

Data una matrice $\underset{n\times p}{A}$

$$A = \underset{n\times p}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ip} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{np} \end{bmatrix}$$

la matrice trasposta $\underset{p\times n}{A'}$ è

$$A' = \underset{p\times n}{A'} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{i1} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{i2} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ a_{1j} & a_{2j} & \cdots & a_{ij} & \cdots & a_{nj} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ a_{1p} & a_{2p} & \cdots & a_{ip} & \cdots & a_{np} \end{bmatrix}$$

dove l'operatore trasposizione $'$ fa in modo che le righe vengono invertite con le colonne, ovvero la prima riga diventa la prima colonna, la seconda riga la seconda colonna etc.

Prodotto fra due matrici

Date due matrici $\underset{n\times p}{A}$ e $\underset{p\times q}{B}$, il loro prodotto è dato da

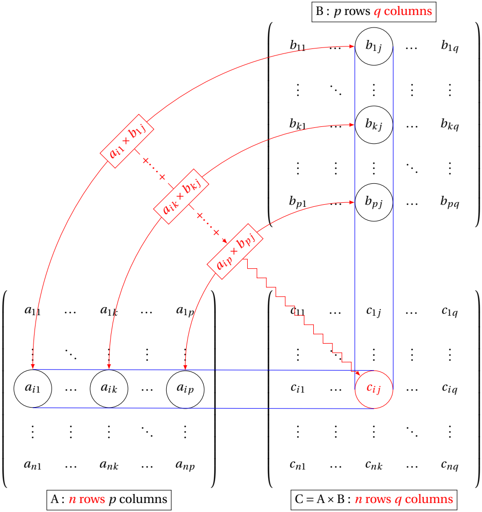
$$\underset{n\times p}{A}\underset{p\times q}{B} = \underset{n\times q}{C}$$

dove l'elemento di posizione (i, j) della matrice C è definito come

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + \ldots + a_{ip}b_{pj}$$

Si noti che il prodotto è possibile fra matrici di dimensioni opportune. Due matrici possono essere moltiplicate fra loro solo se il numero di colonne della prima è uguale al numero di righe della seconda.

Prodotto fra due matrici



Alcune proprietà

Date le matrici A, B e C (di dimensione opportune per definire l'eventuale prodotto) e una costante c

- $c(AB) = (cA)B$
- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$
- $(B + C)A = BA + CA$
- $(AB)' = B'A'$

Matrice quadrata e matrice simmetrica

Matrice quadrata

Una matrice è quadrata se il numero delle righe è uguale al numero delle colonne.

Matrice simmetrica

Una matrice quadrata $\underset{n\times n}{B}$ è detta simmetrica se $B = B'$, ovvero se $b_{ij} = b_{ji}, i = 1, \ldots, n, j = 1, \ldots, n$.

Matrice identità

E' una matrice simmetrica con valore 1 sulla diagonale e 0 altrove:

$$\underset{n\times n}{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \dots & \vdots & \ddots & \dots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

Data una matrice $\underset{n\times p}{A}$, vale

$$\underset{n\times n}{I}\underset{n\times p}{A} = \underset{n\times p}{A}$$

e

$$\underset{n\times p}{A}\underset{p\times p}{I} = \underset{n\times p}{A}$$

Matrice invertibile

Una matrice quadrata $\underset{n\times n}{A}$ è detta invertibile se esiste una matrice $\underset{n\times n}{B}$ tale che

$$\underset{n\times n}{A}\underset{n\times n}{B} = \underset{n\times n}{B}\underset{n\times n}{A} = \underset{n\times n}{I}$$

Se è questo il caso, allora la matrice $\underset{n\times n}{B}$ è univocamente determinata da $\underset{n\times n}{A}$ ed è chiamata l'inversa di $\underset{n\times n}{A}$, indicata con $\underset{n\times n}{A^{-1}}$

Matrice inversa

Sia $\underset{q\times q}{A}$ e $\underset{q\times q}{B}$ tali che le rispettive matrici inverse esistano;

- $(\underset{q\times q}{A}^{-1})' = (\underset{q\times q}{A'})^{-1}$
- $(\underset{q\times q}{A}\underset{q\times q}{B})^{-1} = \underset{q\times q}{B}^{-1}\underset{q\times q}{A}^{-1}$

Matrice diagonale

E' una matrice simmetrica con valori $d_1, \ldots, d_n$ sulla diagonale e 0 altrove:

$$\operatorname{diag}(d_1, \ldots, d_n) = \begin{bmatrix} d_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_i & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots & \ddots & \ldots \\ 0 & 0 & \cdots & 0 & \cdots & d_n \end{bmatrix}$$

Moltiplicare una matrice $\underset{n\times p}{A}$ da sinistra per $\operatorname{diag}(d_1, \ldots, d_p)$ equivale, per ogni i, a moltiplicare l'i-sima riga di $\underset{n\times p}{A}$ per d_i; moltiplicare una matrice $\underset{n\times p}{A}$ da destra per $\operatorname{diag}(d_1, \ldots, d_n)$ equivale, per ogni j, a moltiplicare la j-sima colonna di $\underset{n\times p}{A}$ per d_j;

Matrice diagonale invertibile

Una matrice diagonale $\mathrm{diag}(d_1, \ldots, d_n)$ è invertibile se e solo se i valori $d_1, \ldots, d_n$ sono diversi da 0. In questo caso si ha:

$$(\mathrm{diag}(d_1, \ldots, d_n))^{-1} = \mathrm{diag}(1/d_1, \ldots, 1/d_n)$$

Matrice idempotente

Una matrice quadrata $\underset{n\times n}{B}$ è detta idempotente se vale

$$\underset{n\times n}{B}\underset{n\times n}{B} = \underset{n\times n}{B}$$

.

Matrice (semi)definita positiva

Una matrice simmetrica $\underset{p\times p}{B}$ è detta semidefinita positiva se vale

$$\underset{1\times p}{a'}\underset{p\times p}{B}\underset{p\times 1}{a} \geq 0 \quad \forall \underset{p\times 1}{a}$$

Una matrice simmetrica $\underset{p\times p}{B}$ è detta definita positiva se vale

$$\underset{1\times p}{a'}\underset{p\times p}{B}\underset{p\times 1}{a} > 0 \quad \forall \underset{p\times 1}{a}$$